SCHOLARLY PUBLICATIONS

A CURRENT AWARENESS BULLETIN
OF RESEARCH OUTPUT

@DTU

(82nd Edition)

OCTOBER 2019

BY: CENTRAL LIBRARY

DELHI TECHNOLOGICAL UNIVERSITY
(FORMERLY *DELHI COLLEGE OF ENGINEERING)*
GOVT. OF N.C.T. OF DELHI
SHAHBAD DAULATPUR, MAIN BAWANA ROAD
DELHI 110042

PREFACE

This is the **Eighty Second** Issue of Current Awareness Bulletin started by Delhi Technological University, Central Library. The aim of the bulletin is to compile, preserve and disseminate information published by the faculty, students and alumni for mutual benefits. The bulletin also aims to propagate the intellectual contribution of Delhi Technological University (DTU) as a whole to the academia.

The bulletin contains information resources available in the internet in the form of articles, reports, presentations published in international journals, websites, etc. by the faculty and students of DTU. The publications of faculty and student which are not covered in this bulletin may be because of the reason that the full text either was not accessible or could not be searched by the search engine used by the library for this purpose.

The learned faculty and students are requested to provide their uncovered publications to the library either through email or in CD, etc. to make the bulletin more comprehensive.

This issue contains the information published during **October, 2019**. The arrangement of the contents is alphabetical. The full text of the article which is either subscribed by the university or available in the web is provided in this bulletin.

Central Library

CONTENTS

1. A Hybrid Intelligent Approach for Solar Photovoltaic Power Forecasting: Impact of Aerosol Data, ***6.Astitva Kumar***, ***3.M. Rizwan*** and ***3.Uma Nangia***, Electrical, DTU

2. An ANFIS-based compatibility scorecard for IoT integration in websites, ***3.Akshi Kumar*** and ***6.Anshika Arora***, Computer Science, DTU

3. Analytical modeling for static and dynamic response of organic pseudo all-p inverter circuits, Neha Mishra, ***3.Poornima Mittal*** and Brijesh Kumar, ECE, DTU

4. Cross-Domain Ambiguity Detection using Linear Transformation of Word Embedding Spaces, ***8.Vaibhav Jain, 8.Sanskar Jain and 8.Nishant Tanwar***, CSE, DTU

5. Design, synthesis and evaluation of Coumarin-Phenylthiazole conjugates as cholinesterase inhibitors, ***6.Deepak Mishra***, ***6.Atiya Fatima***, ***3.Ram Singh***, Nupur S Munja, Vineet Mehta and Udayabanu Malairaman, Applied Chemistry, DTU

6. Do Women on Boards affect Firm's Financial Performance? Evidence from Indian IPO Firms, Amit Kumar Singh, ***8.Shubham Singhania*** and Varda Sardana, DTU

7. Edge detection based on type-1 fuzzy logic and guided smoothening, ***6.Sahil Raheja*** and ***3.Akshi Kumar***, IT and Computer, DTU

8. Efficient uorescence quenching of 5-aminoquinoline: ffi fl Silver ion recognition study, Nupur Pandey, ***3.Mohan Singh Mehata***, Nisha Fatma and Sanjay Pant, Applied Physics, DTU

9. ENHANCEMENT OF SURFACE INTEGRITY USING SPARK ASSISTED ABRASIVE FLOW MACHINING PROCESS, ***6.PARVESH ALI***, ***3.R. S. WALIA***, ***3.QASIM MURTAZA*** and ***3.RANGANATH M. S,*** DTU

10. Experimental and Computational Investigation on Mechanical Properties of Reinforced Additive Manufactured Component, ***6.Nagendra Kumar Maurya***, ***3.Vikas Rastogi*** and ***3.Pushpendra Singh***, Mechanical, DTU

11. Experimental investigation of joining of aluminum alloy 5083 by friction stir welding (FSW), ***3.R.S. Mishra*** and ***6.Preety Rani***, Mechanical, Production, Industrial and Automobiles, DTU

12. Experimentation and Modelling of CNT additive abrasive media for micro finishing, ***6.Parvesh Ali***, ***3.R S Walia***, ***3.Q Murtaza*** and ***3.Ranganath M.S***, Mechanical, DTU

13. Exploring Start-up Ecosystem and its Structural Impact : A Review, ***6.Surabhi Kapoor*** and ***3.Seema Singh***, Humanities, DTU

14. Fabrication and optimization of AA7075 matrix surface composites using Taguchi technique via friction stir processing (FSP), ***7.Ravi Butola*** , ***3.Ranganath M S*** and ***3.Qasim Murtaza***, Mechanical, DTU

15. Graphene as charge transport layers in lead free perovskite solar cell, Gagandeep, ***3.Mukhtiyar Singh***, Ramesh Kumar, and Vinamrita Singh, Applied Physics, DTU

16. High Diversity Gain Super-Wideband Single Band-Notch MIMO Antenna for Multiple Wireless Applications, ***6.Gaurav Saxena***, ***3.Priyanka Jain*** and Y K Awasthi, Electronics and Communication, DTU

17. Impedance Spectroscopy and Conduction Behavior in CoFe2O4-BaTiO3 Composites, ***6.S. SHANKAR***, O.P. THAKUR, and ***3.M. JAYASIMHADRI***, Applied Physics, DTU

18. Improving system reliability by optimal allocation of resources for discovering software vulnerabilities, Navneet Bhatt, Adarsh Anand and ***3.Deepti Aggrawal***, Management, DTU

19. Kinematic and Dynamic Analysis of an Industrial Six-Axis Robotic Manipulator, ***6.Ashwin Misra*** and ***8.Ghanvir Singh***, Mechanical, DTU

20. Managerial Paradox Toward Flexibility: Emergent Views Using Thematic Analysis of Literature, Sanjai Kumar Shukla, Sushil and ***3.Manoj Kumar Sharma,*** DTU

21. Modelling Of Viscoelastic Fluid In Finishing Processes, ***3.Krovvidi Srinivas***, ***3.Q. Murtaza***, ***3.A.K Agarwal***, Mechanical, DTU

22. New Failure Rate Model for Iterative Software Development Life Cycle Process, ***6.Sangeeta***, ***3.Kapil Sharma*** and Manju Bala, IT, DTU

23. On the Physical Layer Security over Beaulieu-Xie Fading Channel, Puspraj Singh Chauhana, ***6.1.Sandeep Kumar*** and Sanjay Kumar Sonia, Electronics and Communication, DTU

24. Performance and Emission Studies of a Variable Compression Ratio CI Engine using Bio Diesel Made from Waste Cooking Oil, Jatinder Kataria, S K Mohapatra, K Kundu and ***3.R Singh***, Mechanical, DTU

25. Preparation of Low Cost Magnetorheological Fluid and its Testing, ***3.R. C. Singh***, ***3.Ranganath M.S*** , ***8.1.Akshay Bhagat***, ***8.1.Kaushik K.***, ***8.1.Vaibhav Batra***, ***8.1.Shiven Mittal***, ***8.1.Gaurav Jain***, Mechanical and Production, DTU

26. Short-term PV power forecasting based on sky-conditions using intelligent modelling techniques, ***6.G. Perveen[1]***, ***3.M. Rizwan[2]*** and N. Goel, [1]Electronics and [2]Electrical, DTU

27. Sliding Mode Control with RBF Neural Network for Two Link Robot Manipulator, ***6.Ankita Yadav***, ***6.Ajit Kumar Sharma*** and ***3.Bharat Bhushan***, Electrical, DTU

28. Stability behavior of a nonlinear mathematical epidemic transmission model with time delay, ***6.Kanica Goel*** and ***3.Nilam***, Applied Mathematics, DTU

29. Thermodynamic analysis of Organic Rankine cycle driven by reversed absorber hybrid photovoltaic thermal compound parabolic concentrator system, Deepak Tiwari, Ahmad Faizan Sherwani, Deepali Atheaya, ***3.Anil Kumar*** and Nishant Kumar, Mechanical, DTU

1. *Chancellor*

2. *Pro Vice Chancellor* ***2.1.*** *Ex Pro Vice Chancellor*

3. *Faculty* ***3.1.*** *Ex Faculty*

4. *Teaching-cum-Research Fellow* ***4.1.****Asst. Librarian*

5. *Alumni*

6. *Research Scholar* ***6.1.*** *Ex Research Scholar*

7. *PG Scholar* ***7.1.*** *Ex PG Scholar*

8. *Undergraduate Student* ***8.1.*** *Ex Undergraduate Student*

Arabian Journal for Science and Engineering
https://doi.org/10.1007/s13369-019-04183-0

RESEARCH ARTICLE - ELECTRICAL ENGINEERING

A Hybrid Intelligent Approach for Solar Photovoltaic Power Forecasting: Impact of Aerosol Data

Astitva Kumar[1] · **M. Rizwan**[1] · **Uma Nangia**[1]

Received: 9 May 2019 / Accepted: 25 September 2019
© King Fahd University of Petroleum & Minerals 2019

Abstract
The penetration of solar photovoltaic (PV) power in distributed generating system is increasing rapidly. The increased level of PV penetration causes various issues like grid stability, reliable power generation and power quality; therefore, it becomes utmost important to forecast the PV power using the meteorological parameters. The proposed model is developed on the basis of meteorological data as input parameters, and the impacts of these parameters have been analyzed with respect to forecasted PV power. The main focus of this research is to explore the performance of optimization-based PV power forecasting models with varying aerosol particles and other meteorological parameters. A newly developed intelligent approach based on grey wolf optimization (GWO) using multilayer perceptron (MLP) has been used to forecast the PV power. The performance of the GWO-based MLP model is evaluated on the basis of statistical indicators such as normalized mean bias error (NMBE), normalized mean absolute error (NMAE), normalized root-mean-square error (NRMSE) and training error. The results of the developed model show the values of NMBE, NMAE and NRMSE as 2.267%, 4.681% and 6.67% respectively. To validate the results, a comparison has been made with particle swarm optimization, Levenberg–Marquardt algorithm and adaptive neuro-fuzzy approach. The performance of the model is found better as compared to other intelligent techniques. The obtained results may be used for demand response applications in smart grid environment.

Keywords Grey wolf optimization · Artificial neural network · Solar power forecasting · Solar PV · Distributed power generation

1 Introduction

The demand of electrical energy is increasing drastically because of increasing population, urbanization and increased comfort level. Due to greenhouse gases emission and depleting nature of conventional energy sources, the adoption of renewable energy sources such as solar, wind and biomass has been gaining momentum. Among the renewable energy sources, the adoption of solar photovoltaic is increasing day by day, because of various advantages such as clean, silent, availability and low running cost. However, the rapid development of solar PV system also brings many challenges in operation and control of solar PV-based power systems [1, 2]. Further, the high penetration of solar PV may cause the grid stability and reliability issues in smart grid environment. Therefore, the solar PV power forecasting is one of the key tools for this paradigm and plays a vital role in efficient integration of renewable power generation and reduction in volatility to distributed generation systems. Precise and accurate forecasting of PV power helps massively in planning and scheduling of power demand supply for the consumers as well as various distribution companies (DISCOMs). It further enhances system reliability and power quality issues due to uncertainty in PV power generation.

Further, the forecasting is an important tool for grid operators since they can schedule generation and distribution according to the load profile, and also define generation units to act as operating reserve generators [3, 4]. It is also an essential tool for creating a modern transactive energy framework in a grid-connected microgrid system [5–8]. PV power prediction not only reduces the number of units required in hot standby but also minimizes the operational costs. Therefore, prediction of PV power has become vital and very crucial for grid stability, optimal unit commitment and economical dispatch.

✉ Astitva Kumar
astitvakumar_phd2k16@dtu.ac.in

[1] Department of Electrical Engineering, Delhi Technological University, Delhi 110042, India

Published online: 12 October 2019

Recently, many studies have been performed to obtain accurate forecasts using different approaches such as mathematical modeling, regression models and intelligent technique-based approaches. There are two major approaches to predict solar PV output such as direct and indirect approach [9]. Direct approach is a technique where it directly calculates the power output of the PV system, whereas an indirect approach, firstly calculates the solar irradiation, and then calculates the PV power produced using a PV performance model of the plant. The indirect method is basically modeled using cloud imagery and numerical weather prediction techniques. Direct approach involves modeling using machine learning technique.

There are studies that have focused only on prediction of solar irradiation as it is the main element on which performance of the PV plant depends massively. Wan et al. [10] provided a comprehensive review for both the methods, direct and indirect, and also discussed the merits and demerits of various types of estimation methods. Antonanzas et al. presented an extensive review of solar PV power forecasting techniques focusing on the latest advancement and future trends. It is concluded that mostly forecasting horizon considered is day ahead, and statistical approach is superior in most occasions [11]. Chaturvedi et al. [12] also presented a review of solar forecasting based on various time horizon and classified as intra-hour, intra-day and day ahead. Figure 1 shows the different methodologies used for varying spatial and temporal resolutions, and Fig. 2 gives brief explanation about the different forecasting horizons (time horizons) and its various applications. Martinez-Anido et al. [13] discussed the impact of PV power forecasting

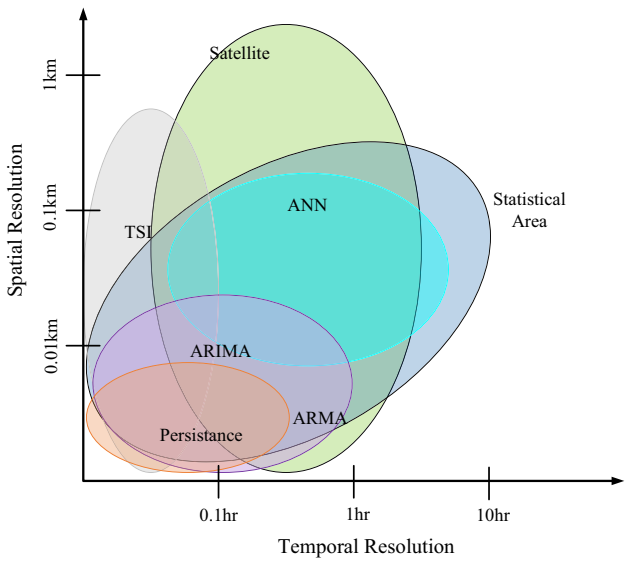


Fig. 1 Classification of short-term forecasting model

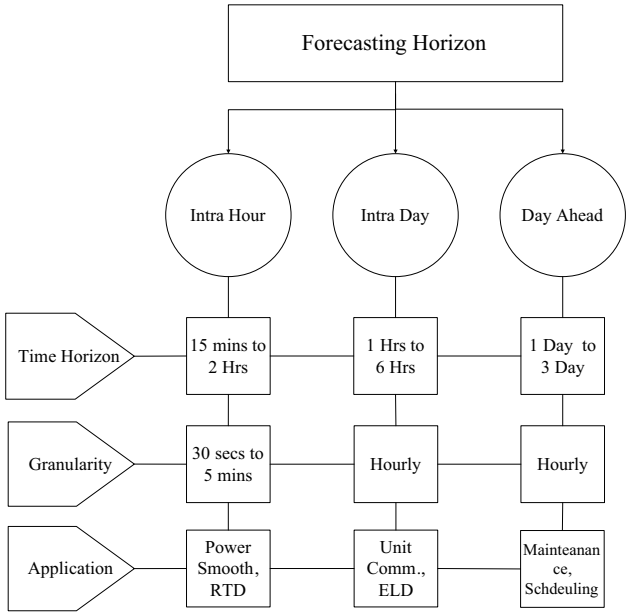


Fig. 2 Different forecasting horizons and their relation

in improving the power system operations and the cost of electricity generation.

Time horizon (t_h) of forecasted PV power output plays an important role in management of the PV systems and its control, for different locations like schools, hospitals, etc. Sanusi et al. performed a study considering clearness index (K_T) and diffuse ratio (D_R) for feasibility of solar energy utilization in tropical stations of Nigeria. In the study, 11-year data of solar radiation and maximum and minimum temperatures were collected and compared with monthly variations of K_T and D_R. The authors therefore suggested that the output of the solar PV system needs evaluation under different conditions, especially considering spatiotemporal variations, thus necessitating the need to monitor grid-connected PV plants by installers and investors [14]. Chu et al. [15] has proposed a PV power reforecasting model using ANN model based on three baseline model, namely ARMA, cloud tracking and kNN for 5, 10 and 15 min time horizons. The effect of aerosols over northern and eastern China was studied by Lia et al. [16]; their study showed that 1.5 kW h/m^2 per day was an annual reduction in solar radiation when compared with a pollution-free environment. The correlation between atmospheric aerosol and reduction of PV power is to be studied. Thus, there is a need for a solar PV power forecasting model that adapts to this change in the atmosphere.

To help in study of solar power output characteristics, authors have used multiple linear regression analysis with various weather data and solar irradiance data. Some other machine learning algorithms like *k*th nearest neighbor (kNN) regression algorithm, support vector machine (SVM), adaptive fuzzy algorithm, Markov chain, cluster evaluation are

present in different literatures for solar forecasting [17]. De Giorgi et al. designed Elmann ANN for the estimation of 960 kWp grid-connected PV plant. The performance indices considered were normalized error (NE), normalized mean absolute error (NMAE) and normalized root-mean-square error (NRMSE) [18]. In Ref. [19], authors have studied the prediction of global solar irradiance using relative humidity, temperature and modified sunshine duration factor. Amrouche et al. has proposed MLP-based ANN model to forecast global horizontal irradiance (GHI) for two different locations. Adaptability of the proposed ANN–MLP model by predicting GHI using unfamiliar data is highlighted by them [20]. Mellit et al. [21] discussed various artificial intelligent techniques for the sizing of various PV systems, PV–wind hybrid systems and also modeling of 24-h solar forecast system in Trieste, Italy. Ehsan et al. [22] proposed a 24-h ahead PV power forecasting model using MLP and discussed the use of different training algorithms and transfer function for different layers to map the input–output database. A fuzzy logic-based neural network model is designed in [23] to estimate solar radiation forecast using time-series data of radiation and temperature for different sky conditions. The mean absolute percentage error (MAPE) based on the fuzzy logic algorithm is between 13.87 and 20.22% for the solar radiation forecast. The MAPE of the neural network prediction model ranges from 10.85 to 20.33%. The results obtained from ARIMA models show MAPE is about 30% during the forecast studies of solar radiation [24]. NE ranges from [− 10%; 10%] for the various empirical and persistent models to estimate PV power; the use of optimized training algorithms along with MLP tool helps in further reducing it [25–27]. In Ref. [28], authors have proposed a new approach of ensemble tree learning algorithm to improve the SVM, ANN and DT techniques in predicting solar radiation with RMSE between 4.6 and 14.6%. Douiri modeled a self-tuning solar PV power model based on Takagi–Sugeno-type neuro-fuzzy technique. The model simulates the *I–V* characteristics to predict global maximum power point with 1.4% average error [29]. The above discussed models do not consider atmospheric aerosol content in modeling of the ANN or MLP forecasting models, and the errors range from 5 to 20%.

From the literature survey, it can be found that there is a need to find a new technique to forecast solar power for the growing smart grid environment with minimal error. The literature mainly focuses on solar energy or irradiance forecasting, and limited papers actually focused on PV power forecasting. The neural network techniques used in the previous PV power prediction model require large amount of dataset for training, but still are unable to predict accurately during hazy and high-pollution condition. The neural network-based prediction models with short-term time horizon were rarely found in the literature. The relationship between performance of PV forecasting models and aerosol attenuation in the atmosphere is rarely found in the literature. Authors of various literatures have tried to study the effect of rising pollution on the solar photovoltaics across different parts of the world and tried to develop a relation between aerosol particles and the solar energy. The paper develops a hybrid intelligent model to predict the solar PV power generation for varying aerosol content. A brief introduction about ANN and its various learning techniques, namely Levenberg–Marquardt (LM) and optimization-based (PSO and GWO) learning methods, have been discussed in the paper. In this paper, a hybrid direct solar PV power forecasting technique is proposed based on GWO approach. GWO is used to train weights of the feedforward ANN for mapping meteorological inputs and PV power as output. In this work, the solar irradiance (W/m^2) incident on the PV panel, cell temperature (^oC), Linke turbidity factor (τ_L) and wind speed (m/s) has been considered as input parameters and PV power as output of the model. The solar irradiance and cell temperature have been measured using solar system analyzer, and wind speed is collected from Global Forecast System (GFS) model. The highlights of the proposed model are use of GWO technique for training of the neural network. Study for selection of the various inputs and their relationship with solar PV power is performed, giving an insight of dependence of PV power generation on each input considered. The proposed model has been formulated considering the linear correlation factor of meteorological data to the PV power generation. The training and prediction of the solar PV power is done considering τ_L for varying aerosol content in the atmosphere. Extensive comparative analysis of the proposed model for hourly forecasting of PV power is presented using different performance indices for forecasting such as training error (TE), normalized root-mean-square error (NRSME), normalized mean bias error (NMBE), normalized mean absolute error (NMAE) and complexity and configuration of the system.

This paper is organized as follows. Section 2 presents a brief outline of the structure of the neural networks and MLP; the section also helps in explaining the learning techniques used to effectively train the artificial neural networks, techniques such as Levenberg–Marquardt, PSO and GWO. These techniques have been modeled using mathematical equations in Sect. 2. The effect of Linke turbidity and atmospheric aerosol variability on the solar irradiance is discussed in Sect. 3. The description of the site selected for data collection and testing along with the data processing is described in Sect. 4. Results and discussions of the hybrid intelligent model and its comparative study are provided in Sect. 5. A conclusion followed by the references is present in Sect. 6.

2 Artificial Neural Network

ANN can be described in general terms as a network to process information, and this network model is inspired by the architecture of a human brain. This is a parallel distributed type of architecture, comprising data processing elements called 'neurons.' ANN is generally characterized on the basis of three key factors such as architecture, training or learning algorithm and activation function. ANN is structured with input layers, hidden layers and output layers. Input layer is the data receiver; data are sent after computation through the output layer [30, 31]. Hidden layer is intermediate layer present between input and output layers; the number of hidden layers depends on convergence speed, generalization capability, neurons present in each layer and training data. These layers in the network are interconnected with help of communication links; these links are associated with specific weights that drive the amount of information passing from them. The transfer function in ANN is chosen according to the properties required. These can be linear or nonlinear to form an equivalent single-layer or multilayer network, respectively. Activation function is actuated when net_j exceeds the threshold to produce the output o_j. ANN working is divided into two steps: firstly, a training process, and secondly, network operating process. Schematic diagram of ANN architecture is shown in Fig. 3.

In this work, MLP model has been used because of various advantages such as functional approximation, fast learning capability to learn complex relationship for easy mapping of inputs to outputs. MLP is a layered feedforward ANN model that maps input to the output using historical data so that model can produce output when the desired output is unknown. Different training algorithms are employed on neural network model to update the weights of the layers; these training and updating of weights occur using pattern or batch [32]. The training cycle is completed when each training pattern is presented once to the MLP. Consider a two-layer MLP neural network as shown in Fig. 4.

The net input (n) of the unit i in layer $k+1$ is given by

$$n^{k+1}(i) = \sum_{j=1}^{sk} w^{k+1}(i,j)a^k(j) + b^{k+1}(i) \tag{1}$$

The output (a) of unit i will be

$$a^{k+1}(i) = f^{k+1}\left(n^{k+1}(i)\right) \tag{2}$$

For an M layer network, the system equations in matrix form are given by

$$a^{k+1}(i) = f^{k+1}\left(w^{k+1}a^k + b^{k+1}\right) \tag{3}$$

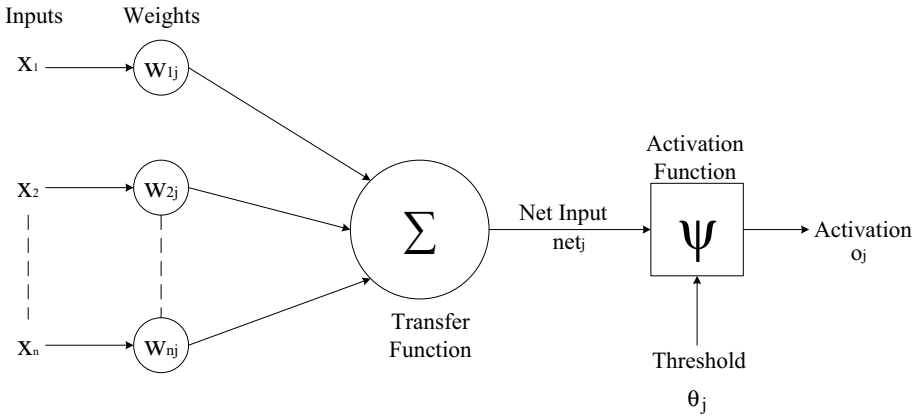


Fig. 3 Schematic diagram of ANN

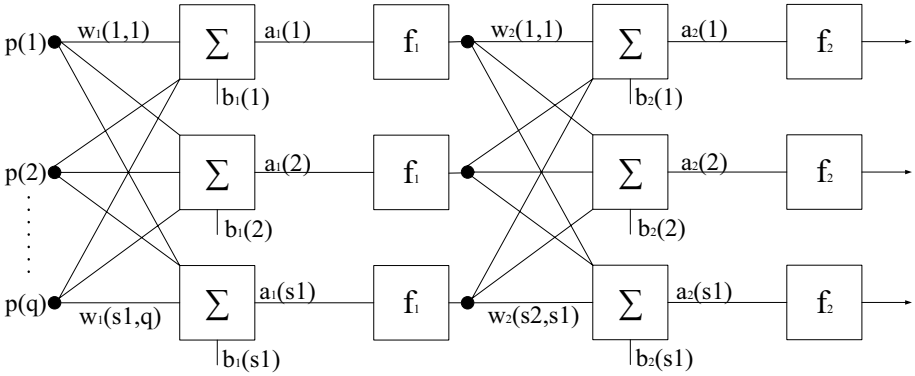


Fig. 4 Two-layer feedforward MLP neural network

Here, $k=0, 1, \ldots, M-1$, f^{k+1} is activation function of $(k+1)$th layer, w^{k+1} is the weight of $(k+1)$th layer and b^{k+1} is bias of $(k+1)$th layer.

From Fig. 4, it is observed that the task of the network is to learn associations between a specified pair of input and target pairs $\{(p_1,t_1),(p_2,t_2),\ldots,(p_q,t_q)\}$, so the performance index for the network is

$$V=\frac{1}{2}\sum_{q=1}^{Q}\left(t_q-a_q^M\right)^{\mathrm{T}}\left(t_q-a_q^M\right)=\frac{1}{2}\sum_{q=1}^{Q}e_q^{\mathrm{T}}e_q \tag{4}$$

where a_q^M is the output of the network when the p_q input is presented, and the error for the qth input is given by $e_q=t_q-a_q^M$. The performance index is approximated by

$$\hat{V}=\frac{1}{2}\sum_{q=1}^{Q}e_q^{\mathrm{T}}e_q \tag{5}$$

where the total sum of squares is replaced by squared errors for a single input/output pair.

2.1 Learning/Training Algorithms

Various learning techniques or training algorithms are used for modeling of the neural network-based MLP. These algorithms are namely gradient descent, Newton's method, conjugate gradient and Bayesian regularization. These learning methods help in creating a decision function for weight updation of the neural network. These techniques have few shortcomings of using large memory, high complexity and slow convergence. These learning methods define rules, and these rules imply a specific function to update the weights of the neurons after each epoch. The term gradient and BP can be used interchangeably; BP is a method where error between the network output and desired output is propagated backwardly. In the BP networks, data are fed forward, and there is no feedback. It is difficult to predict the result from various algorithms [33]. Optimization techniques can also be used for minimizing the cost function to define the weights of the MLP. The cost function depends on numerous parameters, so the optimization is single and multi-dimensional. Optimization helps in increasing the computational speed and enhances the convergence rate. In this paper, LM technique is used as the basis to formulate optimization-based neural network MLP. This paper proposes a new GWO-trained MLP and compares the proposed model with the models designed using LM technique and an intelligent optimization technique—PSO.

2.1.1 Levenberg–Marquardt Algorithm

LM algorithm is also known as damped least-square technique. This technique is designed to enhance the training performance without computing Hessian matrix [34]. This method is especially designed for functions relating to sum of squared error. The highlight of this algorithm is training speed, but limits the usage where large dataset for training is present. It also requires large computational memory. The cost function ($V(x)$) for LM is a sum of squares of error (e_i) (typically for feedforward networks) and is given by

$$V(x)=\sum_{i=1}^{N}e_i^2(x) \tag{6}$$

$$\nabla V(x)=J^{\mathrm{T}}(x)e_i(x) \tag{7}$$

$$\nabla^2 V(x)=J^{\mathrm{T}}(x)J(x)+S(x) \tag{8}$$

where $J(x)$ is the Jacobian matrix

$$J(x)=\begin{bmatrix}\frac{\partial e_1(x)}{\partial x_1} & \frac{\partial e_1(x)}{\partial x_2} & \cdots & \frac{\partial e_1(x)}{\partial x_n}\\ \frac{\partial e_2(x)}{\partial x_1} & \frac{\partial e_2(x)}{\partial x_2} & & \frac{\partial e_2(x)}{\partial x_n}\\ \vdots & \vdots\ddots & & \vdots\\ \frac{\partial e_N(x)}{\partial x_1} & \frac{\partial e_N(x)}{\partial x_2} & \cdots & \frac{\partial e_N(x)}{\partial x_n}\end{bmatrix} \tag{9}$$

and

$$S(x)=\sum_{i=1}^{N}e_i^2(x)\nabla^2 e_i^2(x) \tag{10}$$

Hessian matrix ($H(x)$) is approximated as

$$H(x)=\nabla^2 V(x)=J^{\mathrm{T}}(x)*J(x) \tag{11}$$

and gradient (g) is calculated as

$$g=J^{\mathrm{T}}(x)*e_i(x) \tag{12}$$

where $H(x)$ is the Hessian matrix, e_i is vector of network errors and $J(x)$ is the Jacobian matrix containing the first-order derivative of the network error with respect to the weights and biases. The key step in this method is computation of Jacobian matrix. This algorithm approximates the Hessian matrix using the following equation:

$$x_{i+1}=x_i-\left[J^{\mathrm{T}}(x)*J(x)+\mu I\right]^{-1}j^{\mathrm{T}}(x)e_i(x) \tag{13}$$

If μ is zero, then the equation is just Gauss–Newton methods, when μ is large, it becomes gradient descent with a small step size. This training algorithm approximates the error and discards second-order derivatives of the error.

2.1.2 Grey Wolf Optimization

Optimization algorithm in a neural network-based MLP has two phase cycles, namely propagation and weight update. The input data are propagated from input layer to the output layer.

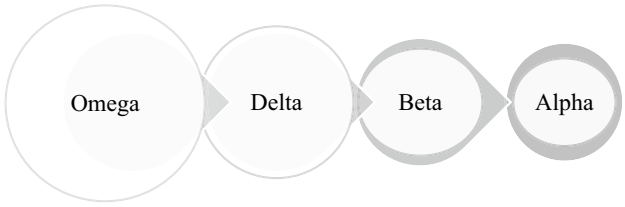


Fig. 5 Social hierarchy in grey wolves

The output from the output layer is compared with the target values, and a cost function is created in form of error. GWO algorithm mimics the properties of a grey wolf (Canis lupus). This algorithm outlines the mathematical model of the social hierarchy, tracking and hunting of grey wolves. In GWO, the fittest solution is considered as alpha (α), the second best is said to be beta (β), the third is called as delta (δ), and the rest are omega (ω). Figure 5 shows the social hierarchy among the pack of grey wolves. This complete is known as pack, in which α is the dominant factor to guide the hunting of the prey [35]. GWO is fast, requires less parameter definition and comparatively low computational memory. It focuses more on exploration, rather than exploitation, of a meta-heuristic technique. The below expressions show the behavior of grey wolves such as hunting and attacking.

$$\vec{D} = \left|\vec{C} \cdot \overrightarrow{X_{\mathrm{p}}}(t) - \vec{X}(t)\right| \tag{14}$$

$$\vec{X}(t+1) = \overrightarrow{X_{\mathrm{p}}}(t) - \vec{A} \cdot \vec{D} \tag{15}$$

where $\vec{D}$ defines the encircling behavior of the grey wolves, $\vec{X}$ indicates the position vector of grey wolf, $\overrightarrow{X_{\mathrm{p}}}$ is the position of the prey, $\vec{A}$, and $\vec{C}$ are coefficient vectors. The vectors $\vec{A}$ and $\vec{C}$ are calculated as follows:

$$\vec{A} = 2\vec{a} \cdot \vec{r_1} - \vec{a} \tag{16}$$

$$\vec{C} = 2 \cdot \vec{r_2} \tag{17}$$

where $\vec{a}$ are linearly decreased from 2 to 0 over the course of iterations, and r_1, r_2 are random vectors in [0, 1].

The following equation represents the hunting characteristics of the grey wolf.

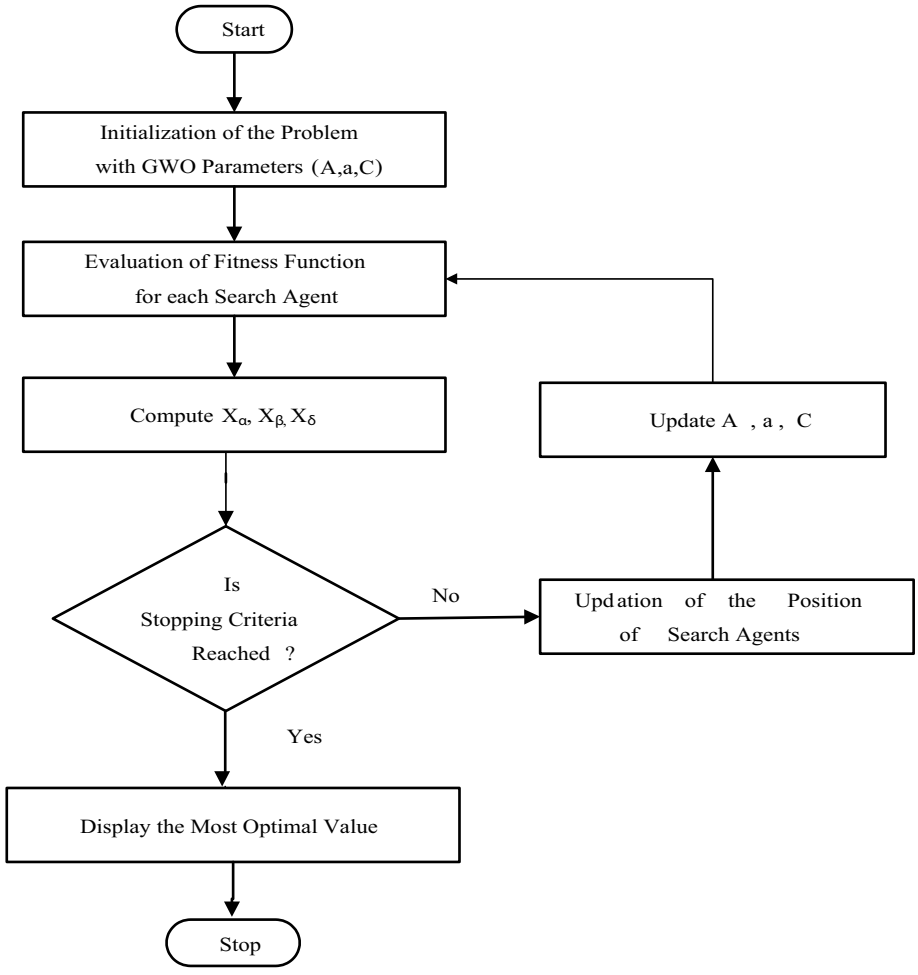


Fig. 6 Flowchart of GWO algorithm

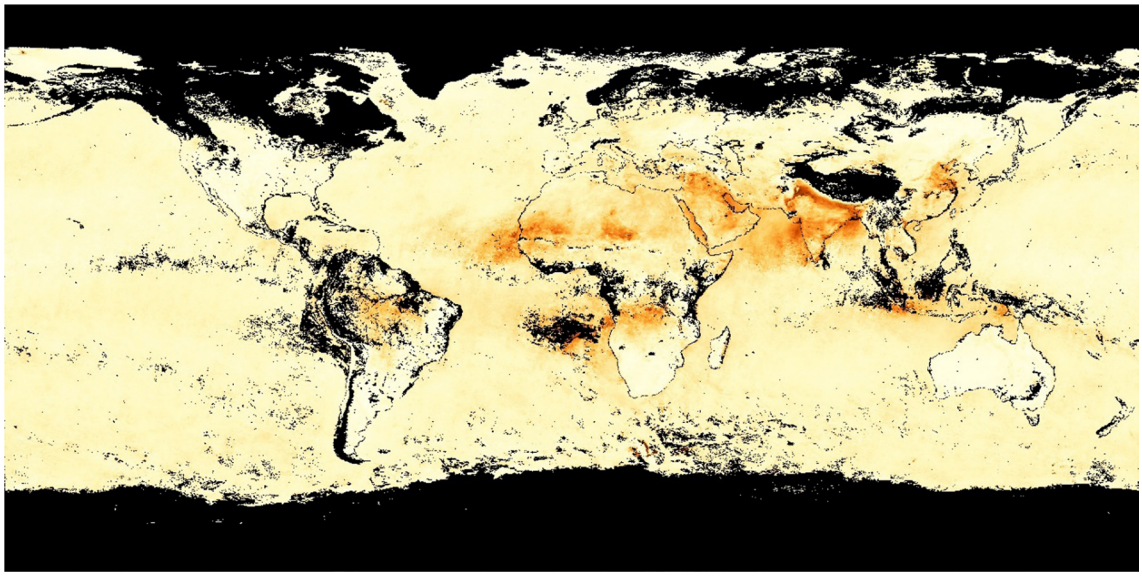


Fig. 7 Aerosol optical depth of the atmosphere for the month of October 2018 worldwide. *Source*: NASA Earth Observations (Terra/MODIS)

$$\begin{aligned}
&\overrightarrow{D_\alpha} = \left|\vec{C}\cdot\overrightarrow{X_\alpha} - \vec{X}(t)\right|; \quad \overrightarrow{D_\beta} = \left|\vec{C}\cdot\overrightarrow{X_\beta} - \vec{X}(t)\right|; \\
&\quad \overrightarrow{D_\delta} = \left|\vec{C}\cdot\overrightarrow{X_\delta} - \vec{X}(t)\right| \\
&\overrightarrow{X_1} = \overrightarrow{X_\alpha} - \vec{A}\cdot\overrightarrow{D_\alpha}; \quad \overrightarrow{X_2} = \overrightarrow{X_\beta} - \vec{A}\cdot\overrightarrow{D_\beta}; \\
&\quad \overrightarrow{X_3} = \overrightarrow{X_\delta} - \vec{A}\cdot\overrightarrow{D_\delta} \\
&\vec{X}(t+1) = \frac{\overrightarrow{X_1} + \overrightarrow{X_2} + \overrightarrow{X_3}}{3}
\end{aligned} \tag{18}$$

$\overrightarrow{X_\alpha}$, $\overrightarrow{X_\beta}$, $\overrightarrow{X_\delta}$ are the position vectors of alpha, beta and gamma, respectively. The GWO algorithm is applied to the practical problem of solar PV power forecasting using neural network MLP model. The steps involved in the development of GWO algorithm are given in flowchart and shown in Fig. 6. The stopping criterion used in the proposed technique is formulated by (i) tolerance error (tol_{err}), (ii) error is no longer decreasing or decreasing too slowly and (iii) maximum iteration is reached.

2.1.3 Particle Swarm Optimization

PSO algorithm is also a widely used technique to minimize or maximize a cost function. PSO is a population-based stochastic, meta-heuristic optimization method developed by Kennedy and Eberhart [36]. It mimics the social behavior of a flock of birds or a swarm of fishes [36]. They have discussed the use of PSO for multi-dimensional experiments such as training of weight in feedforward MLP neural network model. PSO is a technique that lies in between genetic algorithm and evolutionary programming. There are four parameters on

Table 1 τ_{L} and corresponding atmospheric conditions

Linke turbidity factor (τ_{L})	Sky conditions
$\tau_{\text{L}} = 1$	Dry and clean
$1 < \tau_{\text{L}} < 3$	Deep blue (high water vapor content)
$3 < \tau_{\text{L}} < 5$	Turbid atmosphere
$5 < \tau_{\text{L}} < 7$	Polluted atmosphere
$8 < \tau_{\text{L}}$	Hazardous

which performance of this optimization technique depends: personal best (p_{best}), global best (g_{best}), c_1 and c_2. The performance (search and exploit) of PSO is dependent on defining of parameters. The movement of the particles to the near global optimum position is described in the following equations

$$\begin{aligned}
v_{\text{id}}(t+1) = w_{\text{pso}}(t) * v_{\text{id}}(t) + c_1 * r_{\text{pso1}} * \left(p_{\text{best}} - x_{\text{id}}(t)\right) \\
+ c_2 * r_{\text{pso2}} * \left(g_{\text{best}} - x_{\text{id}}(t)\right)
\end{aligned} \tag{19}$$

$$x_{\text{id}}(t+1) = x_{\text{id}}(t) + v_{\text{id}}(t+1) \tag{20}$$

where $v_{\text{id}}(t+1)$ and $v_{\text{id}}(t)$ are the velocities of the particles at $(t+1)$ and t instances, respectively, $x_{\text{id}}(t+1)$ and $x_{\text{id}}(t)$ are the position of the particles at $(t+1)$ and t instances, respectively, c_1 and c_2 are exploitation and exploration coefficients, respectively, r_{pso1} and r_{pso2} are the random variables associated with exploitation and exploration coefficients, respectively, p_{best} and g_{best} are particle having personal best and global best values, $w_{\text{pso}}(t)$ is the weights of the particle that are given by Eq. (21)

$$w_{\mathrm{pso}}(t) = w_{\max} - \frac{\left(w_{\max} - w_{\min}\right)}{T} \times t \tag{21}$$

$w_{\max}$ is 0.9, $w_{\min}$ is 0.4, t is the current iteration, and T are the maximum number of iterations.

3 Impact of Aerosol Data on PV Power Forecasting

The air quality is measured using an air quality index (AQI). Atmospheric condition of a particular site in today's world not only considers the cloud and its movement but also the particulate matter of varying sizes. The movement of clouds and solar cycle has been considered as key factors modulating solar surface radiation. But, atmospheric aerosols and urban haze in cities are the growing problem for all the photovoltaic resources. Aerosols are the minute particles of dust, pollution, sea salts, etc., floating in the atmosphere. Aerosols have property to trap the ambient heat and increase the temperature; with an increase in aerosol content, the atmospheric condition changes and has a knock-on effect on the photovoltaics. Studies have been performed showing that 20% of the average solar dimming in China has been attributed due to increase in aerosols when wind speed is pretty low [37]. Studies have shown that that a large difference (≥ 25 W/m^2) is observed in Delhi, India, and 24.10% of aerosol optical depth from 2000 to 2011 has contributed to expansion of aerosol layer over 6 km [38]. The present estimates of the aerosols' forcing on global climate range from minus 0.6 watts per square meter to minus 4.0 watts per square meter, when combining both the direct and indirect effects of sulfate aerosols and biomass burning aerosols. Aerosol optical depth data from NASA observatory (Terra/MODIS) of the atmosphere for the month of October 2018 worldwide are shown in Fig. 7.

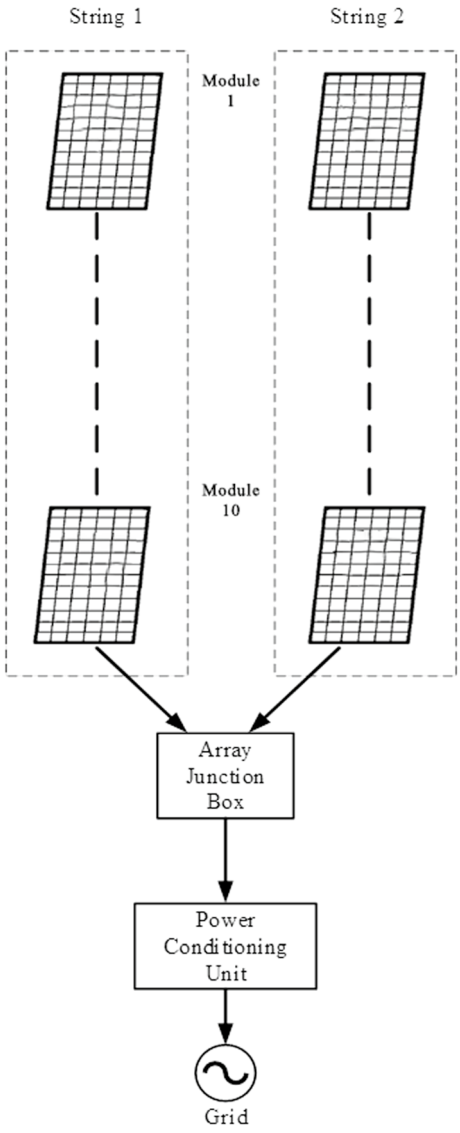


Fig. 8 Schematic single-line diagram of the PV system

Table 2 Specification of the solar PV system

Description	Value
Type	Polycrystalline
Maximum power of module	250 Wp
Open-circuit voltage (Voc) of module	37.05 V
Short-circuit current (Isc) of module	8.580 A
Maximum power voltage (Vpm) of module	30.74 V
Maximum power current (Ipm) module	8.15 A
Output tolerance of module	±2%
Total number of modules	20
Site layout	10×2
Exposure area of a module	1.640 m^2
Temperature coefficients of PV module	
α_{pv}	+0.058%/°C
β_{pv}	−0.330%/°C
γ_{pv}	−0.430%/°C

Aerosol index (AI) is the index that detects the presence of UV absorbing aerosols; this is given by TOMS and OMI AI from NASA Goddard Space Flight Centre [39], defined as

$$\mathrm{AI} = 100\left[\log_{10}\left(\frac{I_{360}}{I_{331}}\right)_{\mathrm{measured}} - \log_{10}\left(\frac{I_{360}}{I_{331}}\right)_{\mathrm{calculated}}\right] \tag{22}$$

The aerosol data can be adequately depicted by calculating the Linke turbidity (τ_{L}). The relationship between aerosols and Linke turbidity factor is discussed in the literature [40, 41]. τ_{L} describes the optical thickness of atmosphere due to various

aerosol particles and water vapor. It defines the absorption and scattering of solar radiation in different sky conditions. A high value of τ_L indicates higher number of aerosol particles in the atmosphere, and thus increasing the diffused radiation and reducing the direct radiation

$$\tau_L = \tau_{Lk}\left(\frac{\rho_{rk}}{\rho_{ra}}\right) \tag{23}$$

$$\tau_{Lk} = (0.9 + 9.4\sin(h)) \cdot \ln\left(\frac{I_0\varepsilon}{I_n}\right) \tag{24}$$

$$\rho_{rk} = 1/\left(9.4 + 0.9m_a\right) \tag{25}$$

$$\frac{1}{\rho_{ra}} = 6.5567 + 1.7513m_a - 0.1202m_a^2 + 0.0065m_a^3 - 0.00013m_a^4 \tag{26}$$

$$m_a = \frac{\text{Pressure}}{101{,}325}\left[\sin(h) + 0.15(h + 3.885)^{-1.253}\right]^{-1} \tag{27}$$

where τ_{Lk} is Linke turbidity factor under clear sky; according to Kasten, ρ_{rk} is Rayleigh integral optical thickness, and ρ_{ra} is integral optical thickness by Louche and modified by Kasten [42]. m_a is the atmospheric air mass, h is the sun elevation angle, and I_0 and I_n are sun constant and direct normal solar radiation. Table 1 shows the relation of varying sky conditions with respect to τ_L. The calculated τ_L is verified from the data along with the ranges given in Table 1 and was found to be accurately depicting the sky conditions.

4 Site Details and Data Processing

A 5 kW_p grid-connected PV system is installed at rooftop of the laboratory, and the data of the system are collected through Solar System Analyzer 9018BT during daytime for every 2 min. The inputs considered for the proposed PV forecasting model are solar irradiance (W/m^2) incident on the PV panel, cell temperature (°C), Linke turbidity (τ_L) and wind speed (m/s). A linear regression study was performed for selection of the various aforementioned inputs on the basis of 1-month historical data. The linear dependence of X and Y is studied using Pearson product moment correlation coefficient, to find the strength of linear dependence of nominal PV power generation output to the inputs. The negative correlation shows that larger values of τ_L and wind speed will result in smaller solar PV output. A large number of samples were collected for training of the neural network based on MLP model. The training dataset comprised of 80% of the time-series data, and 20% of the dataset was used for testing (10%) and validation (10%). The above selection of 80%, 10% and

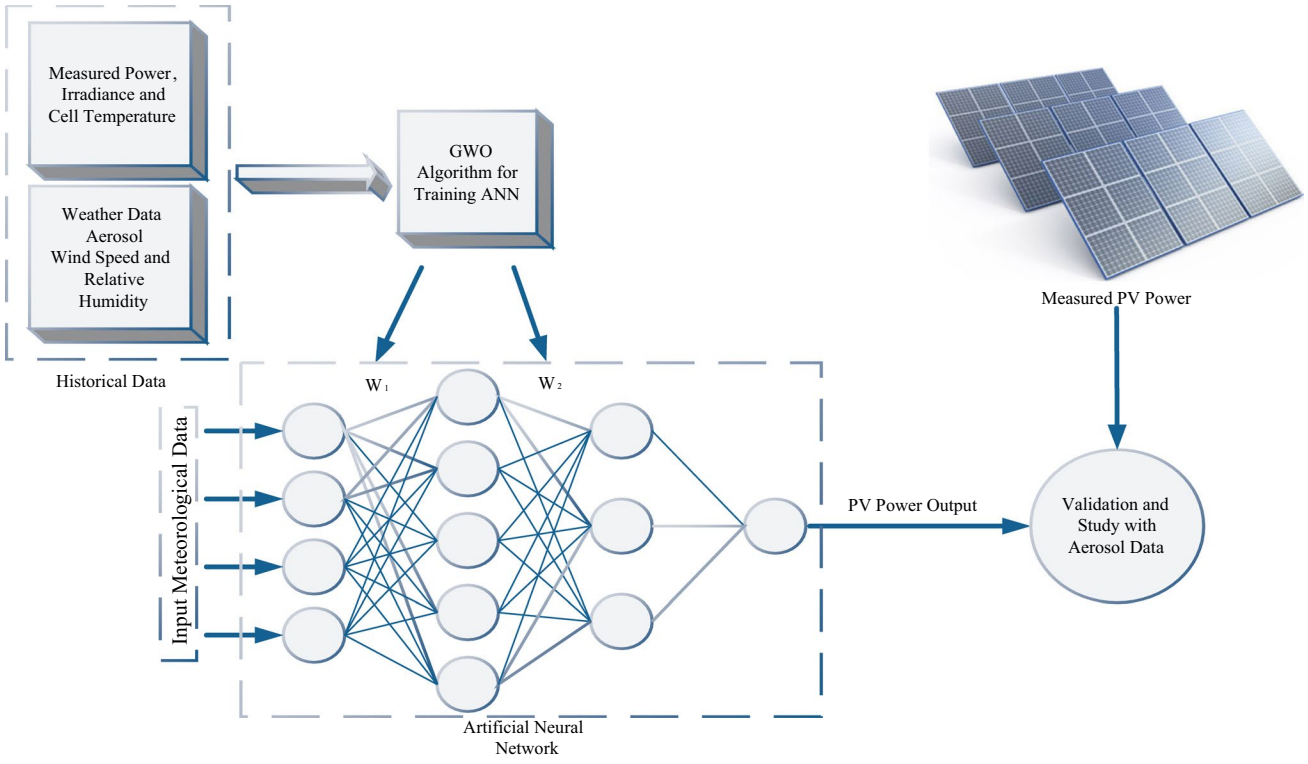


Fig. 9 Functional block diagram of the proposed hybrid intelligent model

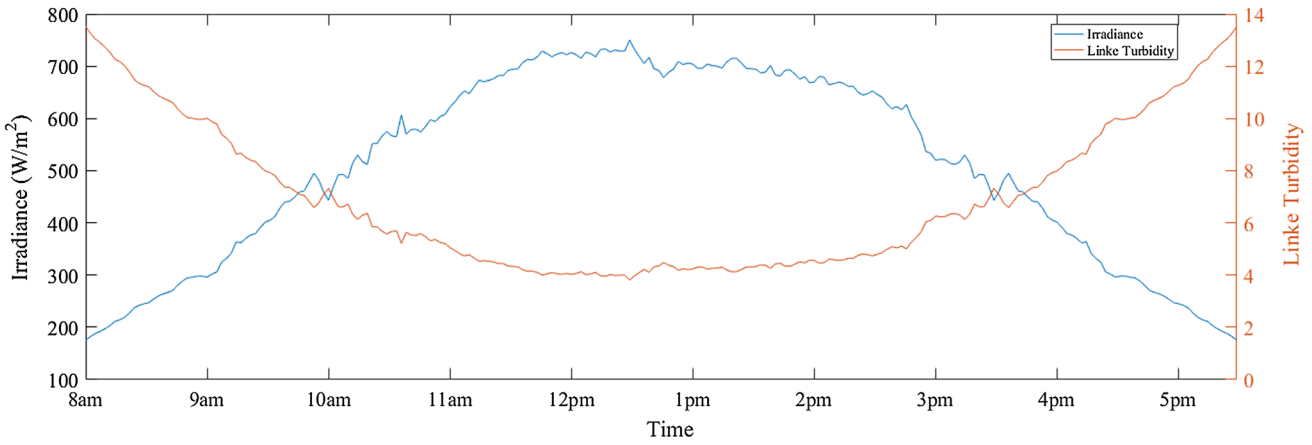


Fig. 10 Variation of irradiance and Linke turbidity for a highly polluted day

10% for different stages has been done using trial and error technique. The performance of the proposed model using above values is shown in Fig. 12. The measured data were verified with the data obtained from NREL and NOAA-GFS model for the last 10 years. The measured irradiance data were in the range of the irradiance obtained from the sources for the selected site. The schematic representation of the PV system considered is shown in Fig. 8, and the specifications of the solar PV system used in this study are provided in Table 2.

The ranges of each input vary in relation with the target values in ANN and other data mining techniques, so to prevent ill conditioning of the network normalization is required. Normalization also helps in having the same range values for different inputs. Normalization is preferred to reach a linear and robust relationship between nonlinear dataset. The inputs and the target values collected from 9018BT have been normalized using Eq. (28)

$$X^* = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \tag{28}$$

where $0 \leq X^* \leq 1$, X_i is the current value of data X, $X_{\max}$ and $X_{\min}$ are the maximum and minimum value of the dataset X.

5 Results and Discussion

5.1 Statistical Indicators

The performance of the proposed hybrid intelligent forecasting model is analyzed using the statistical indices. The performance indicators used for proposed PV power forecasting model are normalized error (NE), normalized squared error (NSE) and training error (TE) or mean square error (MSE). The accuracy of the prediction is evaluated considering these metrics

$$\mathrm{NE} = \frac{P_{\text{out}} - P_{\text{e}}}{\mathrm{Max}_{i=1}^{M}\left(P_{\text{out}}\right)} \tag{29}$$

Here, P_{out} and P_{e} are actual PV power and estimated power output at every t_{h}, and M is the total number of samples

$$\mathrm{NSE} = (\mathrm{NE})^2. \tag{30}$$

NSE is calculated for easier calculation of normalized root-mean-square error (NRMSE)

$$\mathrm{NRMSE} = \sqrt{\frac{1}{M} \times \sum_{i=1}^{M} (\mathrm{NSE}) \times 100.} \tag{31}$$

Normalized mean bias error (NMBE) (%) and normalized mean absolute error (NMAE) (%) are also considered as performance indices

$$\mathrm{NMBE} = \frac{100}{M} \times \sum_{i=1}^{M} (\mathrm{NE}) \tag{32}$$

$$\mathrm{NMAE} = \frac{100}{M} \times \sum_{i=1}^{M} |\mathrm{NE}| \tag{33}$$

$$\mathrm{TE} = \mathrm{MSE} = \frac{1}{n}\left(\sum_{i=1}^{M} \left(x_i - \hat{x}_i\right)^2\right) \tag{34}$$

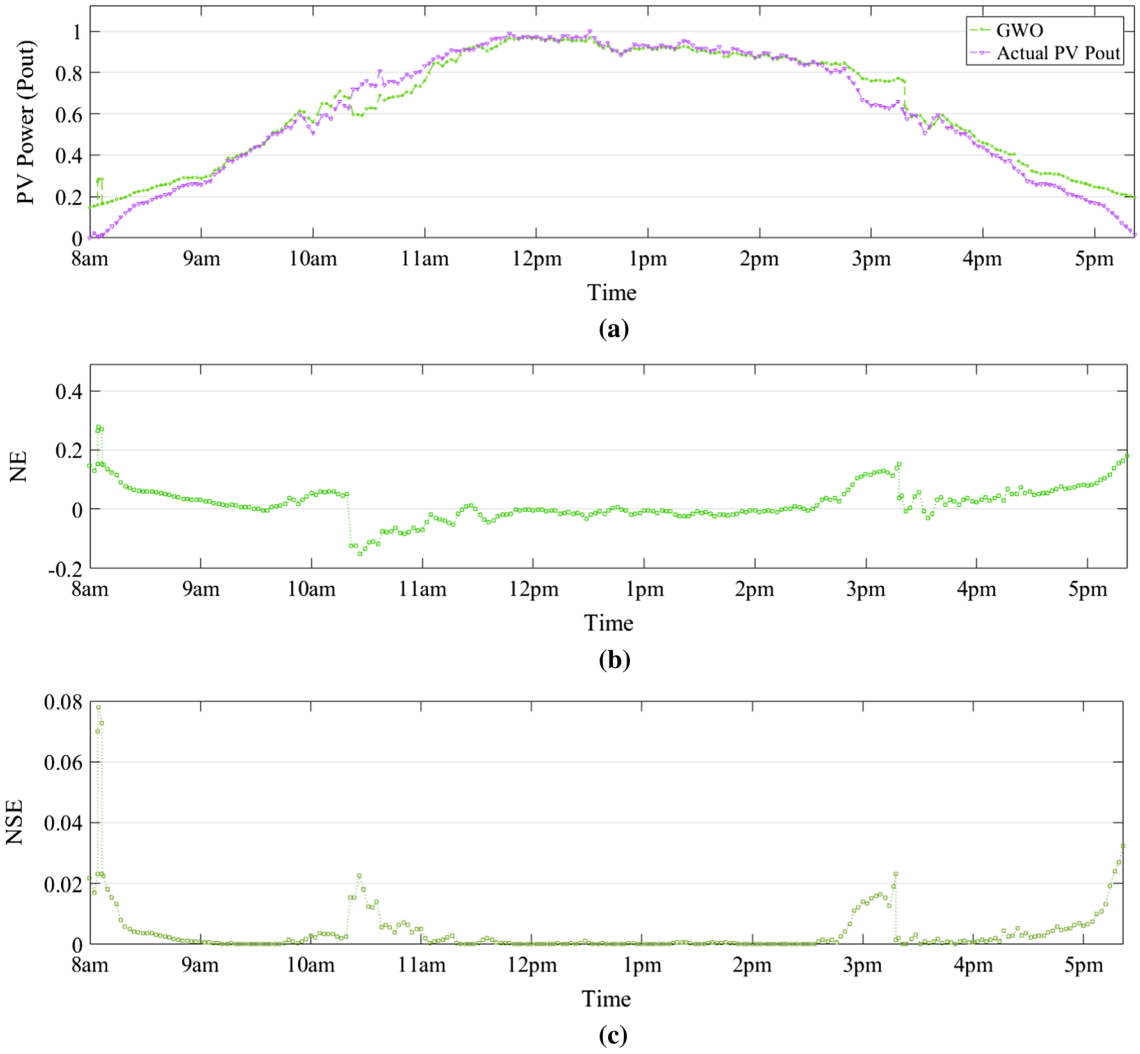


Fig. 11 **a** Performance of GWO model, **b** variation of NE for GWO model, **c** variation of NSE for GWO model

where x_i is the output of the trained model, $\hat{x}_i$ is target value, and M is the number of samples of data.

5.2 Performance of the Proposed Model

The proposed neural network model is trained using GWO algorithm considering four input parameters such as irradiance, cell temperature, wind speed and relative humidity. The training and validation stages of the proposed GWO-based ANN model, to predict the PV power, are shown in Fig. 9. The GWO algorithm acquires the previous data and assists in effective modeling of ANN network for forecasting the PV power. This meta-heuristic optimization algorithm is used to assign optimized weights to the layers of neural network by minimizing the cost function. The proposed hybrid intelligent model uses mean square error as its objective function for a maximum of 1000 iterations; the model comprises of four input nodes with five hidden layers and one output. The proposed method as defined in Eq. (23) accurately calculates the τ_L in accordance with the sky conditions and irradiance prevailing at the time, and it is given in Fig. 10. In this analysis, the effect of different parameters, such as irradiance, wind speed, cell temperature and aerosol data, is given in Table 6 of the appendix section. From Table 6, it is observed that the impact of τ_L on the PV power output is -38.2%; here, negative sign shows that as τ_L increases, the PV power generation decreases. The accuracy

Table 3 Performance indicators of different training algorithms during training state

Technique	Training error (TE)	Iterations/epochs	Regression coefficient (R)
GWO	3.23×10^{-4}	148	0.99720
PSO	4.33×10^{-4}	963	0.98575
LM	1.24×10^{-3}	89	0.98410
ANF	7.68×10^{-3}	678	NA

of the aerosol data and its effect on the PV power generation are studied in Fig. 18.

The performance of the proposed model has been assessed with changing air quality and is shown in Fig. 10. This figure shows the variation of irradiance and Linke turbidity for a highly polluted day. The mean τ_L for the day is calculated to be 8.5 (very polluted condition), and the density of aerosol particles during the morning and evening hours is very high and gradually decreases and is constant for the afternoon hours, for the selected site.

The forecasted PV power output from GWO–MLP model is also compared with the actual solar PV power measured using 9018BT Solar Analyzer. The measured PV power (P_{out}) and forecasted PV power using GWO model is shown in Fig. 11a. From Fig. 11a, it can be observed that the forecasted PV power is very close to the measured data during the afternoon (i.e., high sunshine hours). However, the performance of the proposed model is deteriorated in case of more diffuse solar irradiance. The NE for the designed model ranges from 0.27 to −0.12 as shown in Fig. 11b. GWO technique being more of an exploration-based computational

Table 4 Parameters of various training algorithms for designing of models

GWO					
iter_{max}	No. of hidden layers in neural network (n_h)		$\vec{a}$ decrease factor	$\vec{a}$ initial	tol_{err}
1000	5		$2-\frac{\text{iter}*2}{\text{iter}_{max}}$	2	10^{-6}
PSO					
c_1	c_2	w_{max}	w_{min}	n_h	tol_{err}
1.9	2.1	0.9	0.4	5	10^{-6}
LM					
No. of delays	iter_{max}	μ increase factor	μ decrease factor	n_h	tol_{err}
1	1000	10	0.1	10	10^{-6}
ANF					
Membership function type		No. of membership functions		iter_{max}	tol_{err}
Gaussian bell		[5 5 3 3]		1000	0

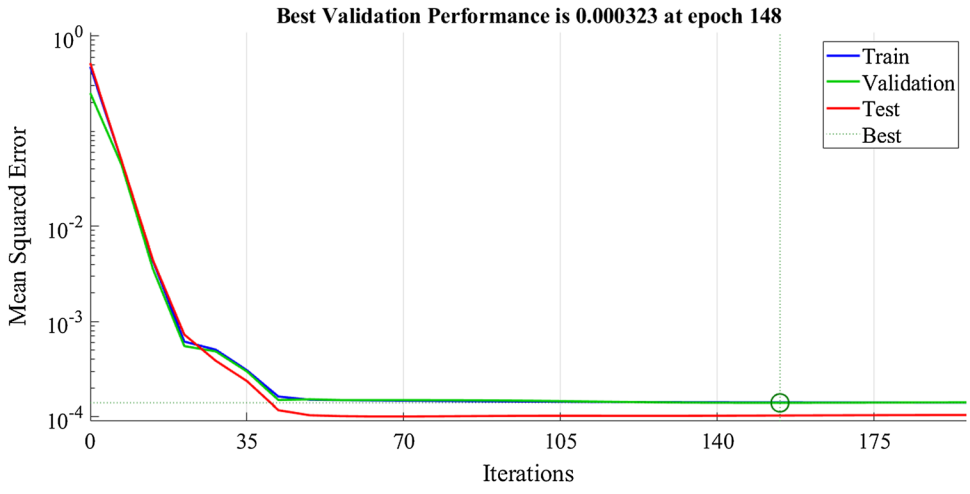


Fig. 12 MSE for testing, validation and training stages of the proposed model

technique, hence can readily forecast the variations of PV power. NRMSE, NMBE and NMAE of the proposed GWO model are 6.67%, 2.267% and 4.681%, respectively. As all the statistical indices have small values, it shows that the PV power output is accurately predicted for varying conditions that occur in a day.

A comparative analysis among the aforementioned algorithms is carried out on the basis of TE, iterations required for minimum TE and regression coefficient (R). The training error is said to be the error obtained when the trained model is again trained for target values. Regression coefficient (R) defines the associativity of the model output with its inputs. It defines the goodness of fitness function. It is a measure of how well the variation in the output is explained by the targets; the closer this coefficient is to 1, the better is the relation of inputs and outputs. Table 3 shows that the four training algorithms have been compared on the basis of TE, R and number of iterations required to reach minimum TE. The value of R using GWO-based MLP model is found to be 0.99832 with minimum number of iterations, i.e., highest computational speed training algorithm. From these results, it is revealed that the GWO-based model requires less computational time as compared to the other intelligent algorithms such as PSO. Further, the value of training error is also less for the proposed model which is 3.23×10^{-4}.

From Table 3, it can be concluded that GWO training technique is relatively better than the other training techniques. This technique has the minimum TE, and hence the associativity of inputs to output is also high, and the computational speed is faster than adaptive neuro-fuzzy (ANF) and PSO techniques. The training parameters for various algorithms are given in Table 4. The n_h value for the LM-based ANN model is 10 to reach a feasible TE of 10^{-3} order; this value of n_h is maximum in comparison with other techniques used; thus, the ANN model with LM approach is more complex than the GWO- and PSO-based models. Figure 12 clearly shows the performance of the GWO-based ANN solar PV forecasting model during different stages, i.e., training, validation and testing. The testing stage error is of the order 10^{-4}, i.e., very small, and the error for validation stage is similar to the training stage error.

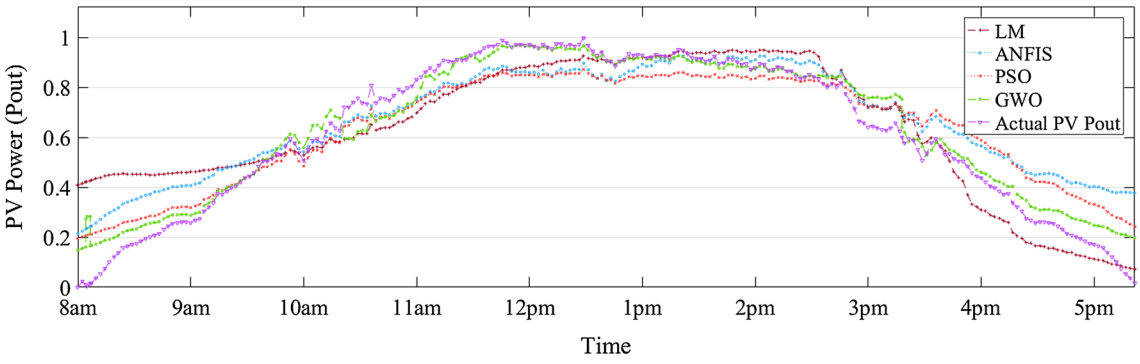


Fig. 13 PV power forecasting for high τ_L condition

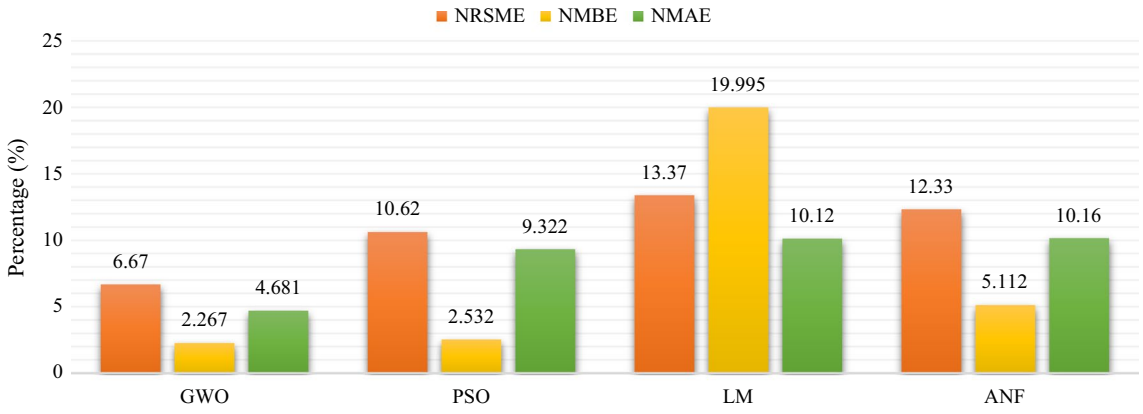


Fig. 14 Comparison of various forecasting model in terms of NRMSE, NMBE and NMAE

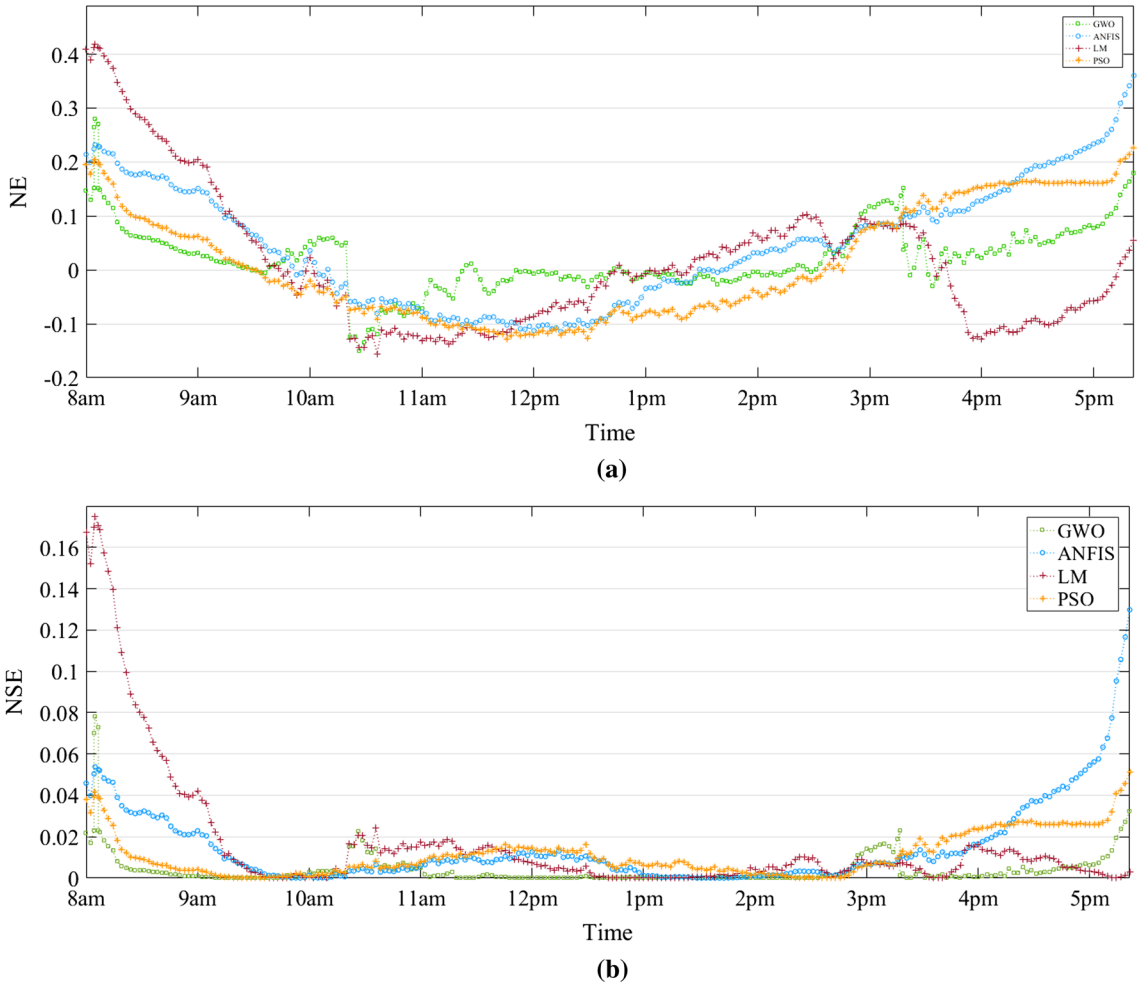


Fig. 15 **a** Variation of NE for different models, **b** variation of NSE for different models

The forecasted PV power output obtained from LM, ANFIS, PSO and GWO is compared with the measured data and shown in Fig. 13. The models forecast the PV power for every 2 min during sunshine hours. From Fig. 13, it is clearly seen that during morning hours, the PV power generation is low, and the power reaches maximum value in the noon hours, and then gradually decreases during the evening hours. However, the performance of the developed model is found satisfactory during sunshine hours. The performance indices are shown in Fig. 14 for different solar PV forecasting models. The normalized error (NE) and normalized squared error (NSE) of the predicted PV power for every sample are given in Fig. 15a, b, respectively. From Fig. 15a, b, it can be seen that NE and NSE are significantly large for the LM-based MLP model during morning and evening hours where the amount of diffuse irradiance is dominating as compared to the direct solar irradiance. When τ_L is very large, i.e., ≥ 9, diffuse irradiance increases significantly as the aerosol particle also increases the atmosphere. As τ_L is in the range of 5–7, the impact of aerosol particles do not affect the proposed PV power forecasting model as NE is small. However, the performance of the model has been improved significantly during noon hours. The summary of results obtained using Fig. 15a, b is also given in Table 5. The performance indicators such as NE, NRMSE, NMBE and NMAE for the different approaches are shown in Table 5. The value of statistical indicators used in this work is better for the proposed GWO model as compared to the other intelligent approaches.

Table 5 Performance comparison of the proposed models for prediction stage

Technique	NE	NRMSE	NMBE	NMAE
GWO	[0.27, −0.12]	6.67	2.267	4.681
PSO	[0.21, −0.10]	10.62	2.532	9.322
LM	[0.41, −0.15]	13.37	19.995	10.12
ANF	[0.38, −0.10]	12.33	5.112	10.16

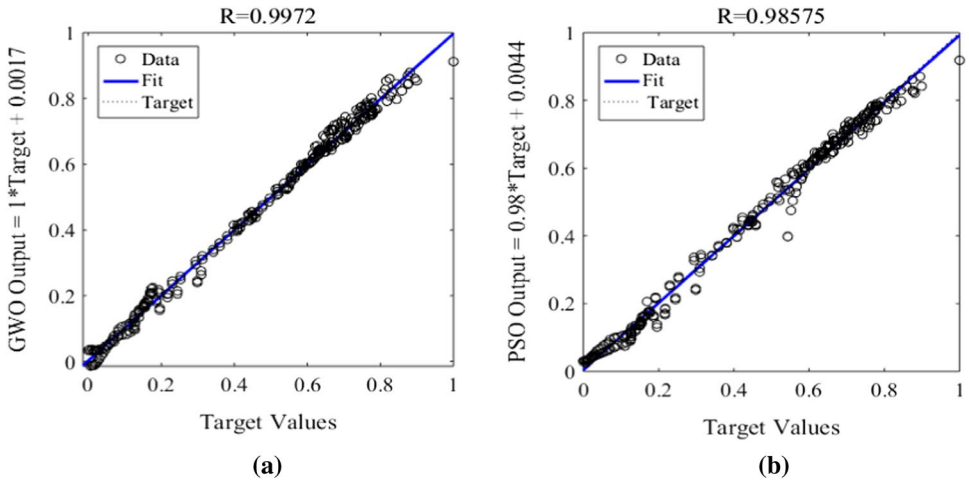


Fig. 16 **a** Regression plot of GWO in training state, **b** regression plot of PSO in training state

From the above figures and Table 5, it is concluded that the performance of the GWO-based approach is found better as compared to the other intelligent techniques. The obtained results from the proposed model may be used for various applications such as demand response, PV plant operation and control in smart grid environment. The regression coefficient plot for GWO- and PSO-based approach is also given in Fig. 16a, b, respectively.

6 Conclusion

In the proposed work, a new model based on GWO approach for solar PV power forecasting has been developed and presented for smart grid environment. The proposed model is developed on the basis of meteorological parameters to forecast the PV power output. The performance of the model is evaluated on the basis of statistical indicators and validated from the dataset obtained from 5 kWp PV system installed at the rooftop of the laboratory. The forecasted PV power model performance has also been analyzed with respect to changing air quality. This paper provides performance analysis of the proposed solar PV power forecasting model for varying air quality modeled using the Linke turbidity factor for the selected site. A comparative analysis of proposed GWO–MLP model with other intelligent techniques is also performed and was found better. The performance indicators such as TE, iterations and R were 3.23×10^{-4}, 148 and 0.99720, respectively, suggesting that the proposed hybrid intelligent model is better. The proposed model also offers relatively fast speed for training of the neurons and is comparatively less complex. The obtained results may be used for smart grid energy management and demand response applications and would be helpful in getting reliable, load smoothing and economical electrical energy to consumers.

Appendix

See Figs. 17, 18 and Table 6.

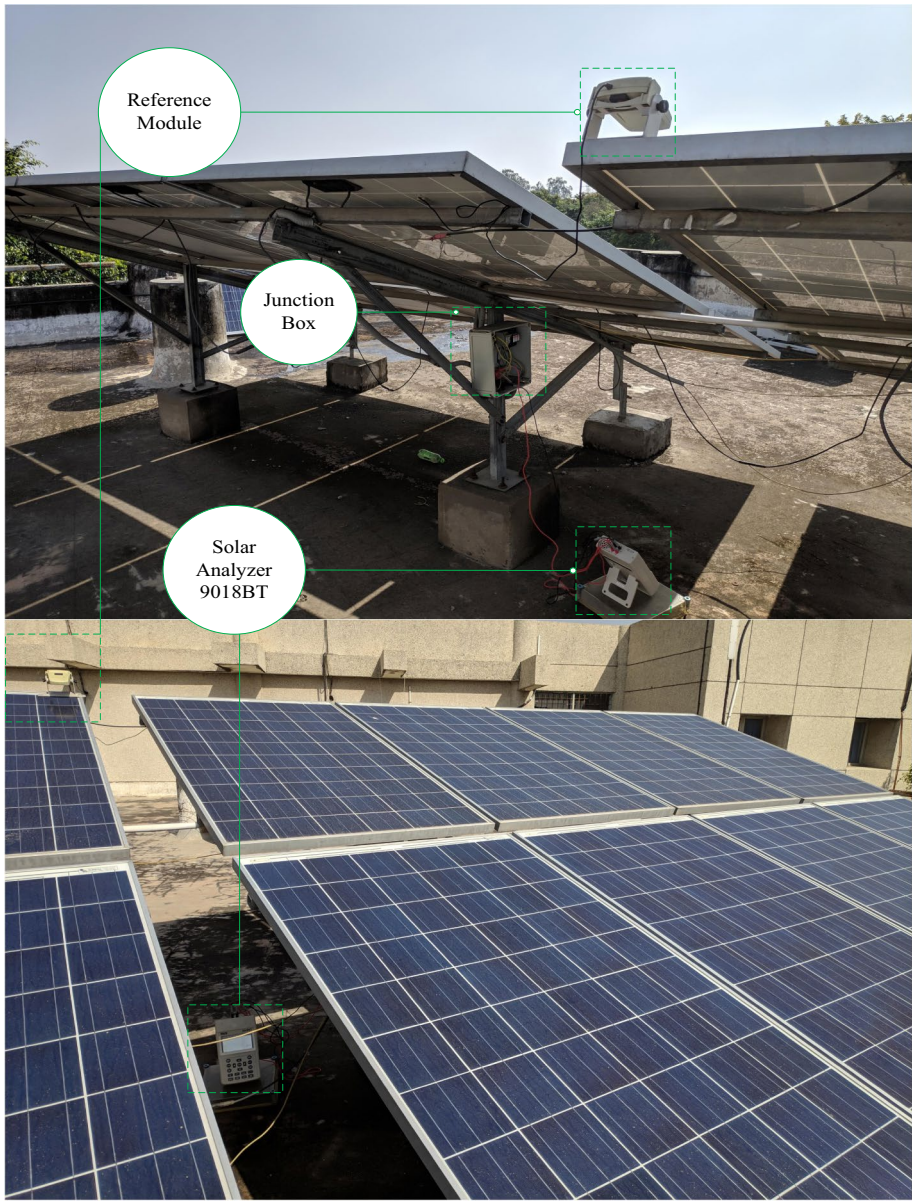


Fig. 17 Experimental setup for solar PV power forecasting

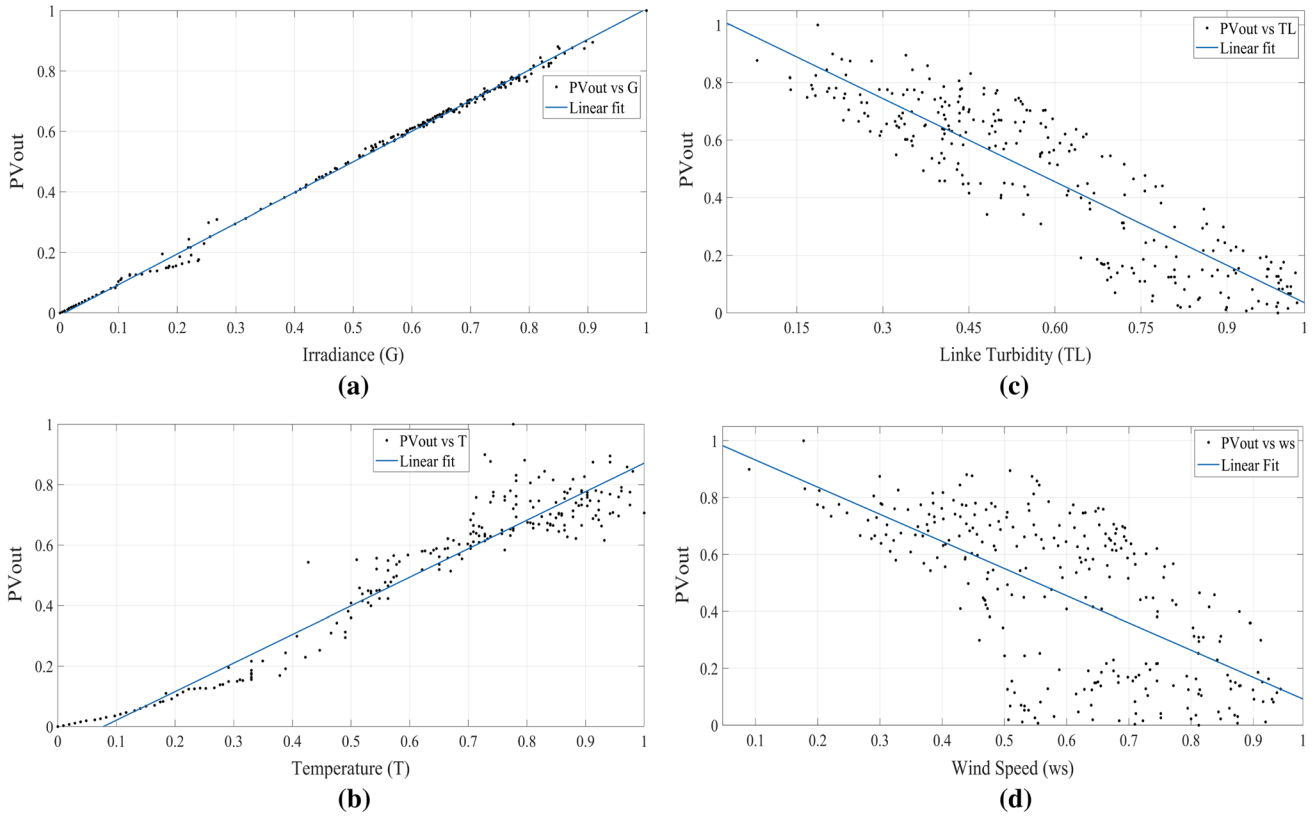


Fig. 18 Linear regression plot of the inputs and the solar photovoltaic power, **a** PV power versus irradiance, **b** PV power versus temperature, **c** PV power versus Linke turbidity, **d** PV power versus wind speed

Table 6 Daily PV output and its correlation with the inputs considered

Inputs	Correlation coefficient
Irradiance	0.993
Temperature	0.912
Linke turbidity	−0.382
Wind speed	−0.191

References

1. Hosenuzzaman, M.; Rahima, N.A.; Selvaraj, J.; Hasanuzzaman, M.; Maleka, A.B.M.A.; Nahar, A.: Global prospects, progress, policies, and environmental impact of solar photovoltaic power generation. Renew. Sustain. Energy Rev. **41**, 284–297 (2015)
2. Long, H.; Zhang, Z.; Su, Y.: Analysis of daily solar power prediction with data-driven approaches. Appl. Energy **126**, 29–37 (2014)
3. Rizwan, M.; Jamil, M.; Kirmani, S.; Kothari, D.P.: Fuzzy logic based modeling and estimation of global solar energy using meteorological parameters. Energy **70**, 685–691 (2014)
4. Gala, Y.; Fernandez, A.; Díaz, J.; Dorronsoro, J.R.: Hybrid machine learning forecasting of solar radiation values. Neurocomputing **176**, 48–59 (2016)
5. Marzband, M.; Azarinejadian, F.; Savaghebi, M.; Pouresmaeil, E.; Guerrero, J.; Lightbody, G.: Smart transactive energy framework in grid-connected multiple home microgrids under independent and coalition operations. Renew. Energy **126**, 96–106 (2018)
6. Monfared, H.; Ghasemi, A.; Loni, A.; Marzband, M.: A hybrid price-based demand response program for the residential microgrid. Energy **185**, 274–285 (2019)
7. Al-Sumaiti, A.; Salama, M.; El-Moursi, M.; Alsumaiti, T.; Marzband, M.: Enabling electricity access: revisiting load models for AC-grid operation—part I. IET Gener. Transm. Distrib. **12**, 2563–2571 (2019)
8. Aryanpur, V.; Atabaki, M.; Marzband, M.; Siano, P.; Ghayoumi, K.: An overview of energy planning in Iran and transition pathways towards sustainable electricity supply sector. Renew. Sustain. Energy Rev. **112**, 58–74 (2019)
9. Voyant, C.; Notton, G.; Kalogirou, G.; Nivet, M.L.; Paoli, C.; Motte, F.; Fouilloy, A.: Machine learning methods for solar radiation forecasting: a review. Renew. Energy **107**, 569–582 (2017)
10. Wan, C.; Zhao, J.; Song, Y.; Xu, Z.; Lin, J.; Hu, Z.: Photovoltaic and solar power forecasting for smart grid energy management. J. Power Energy Syst. **1**, 38–46 (2015)
11. Antonanzas, J.; Osorio, N.; Escobar, R.; Urraca, R.; Martinez-de-Pison, F.J.; Antonanzas-Torres, F.: Review of photovoltaic power forecasting. Sol. Energy **136**, 78–111 (2016)
12. Chaturvedi, D.K.; Isha: Solar forecasting: a review. Int. J. Comput. Appl. **145**, 28–50 (2016)
13. Martinez-Anido, C.B.; Botor, B.; Florita, A.R.; Draxl, C.; Lu, S.; Hamann, H.F.; Hodge, B.M.: The value of day-ahead solar power forecasting improvement. Sol. Energy **129**, 192–203 (2016)
14. Sanusi, Y.K.; Ojo, M.O.: Evaluation of clearness index and diffuse ratio of some locations in South Western, Nigeria using solar radiation data. J. Appl. Phys. **5**, 45–51 (2015)

15. Chu, Y.; Urquhart, B.; Gohari, S.M.I.; Pedro, H.T.C.; Kleissl, J.; Coimbra, C.F.M.: Short-term reforecasting of power output from a 48 MWe solar PV plant. Sol. Energy **112**, 68–77 (2015)
16. Lia, X.; Wagner, F.; Peng, W.; Yang, J.; Mauzeralla, D.L.: Reduction of solar photovoltaic resources due to air pollution in China. Proc. Natl. Acad. Sci. **114**, 1–6 (2017)
17. Diagne, M.; David, M.; Lauret, P.; Boland, J.; Schmutz, N.: Solar irradiation forecasting: state-of-the-art and proposition for future developments for small-scale insular grids. Renew. Sustain. Energy Rev. **27**, 65–76 (2013)
18. De Giorgi, M.G.; Congedo, P.M.; Malvoni, M.: Photovoltaic power forecasting using statistical methods: impact of weather data. IET Sci. Meas. Technol. **8**, 90–97 (2014)
19. Garba, M.B.; Muhammad, A.; Musa, M.; Mohammed, A.G.: Assessing the performance of global solar radiation empirical equations in Sokoto, Nigeria using meteorological parameters. Niger. J. Technol. **37**, 358–364 (2018)
20. Amrouche, B.; Pivert, X.L.: Artificial neural network based daily local forecasting for global solar radiation. Appl. Energy **130**, 333–341 (2015)
21. Mellit, A.; Pavan, A.M.: A 24-h forecast of solar irradiance using artificial neural network: application for performance prediction of a grid-connected PV plant at Trieste, Italy. Sol. Energy **84**, 807–821 (2010)
22. Ehsan, R.M.; Simon, S.P.; Venkateswaran, P.R.: Day-ahead forecasting of solar photovoltaic output power using multilayer perceptron. Neural Comput. Appl. **28**, 3981–3992 (2017)
23. Chen, S.X.; Gooi, H.B.; Wang, M.Q.: Solar radiation forecast based on fuzzy logic and neural networks. Renew. Energy **60**, 195–201 (2013)
24. Safi, S.; Zeroual, A.; Hassani, M.: Prediction of global daily solar radiation using higher order statistics. Renew. Energy **27**, 647–660 (2002)
25. Chupong, C.; Plangklang, B.: Forecasting power output of PV grid connected system in Thailand without using solar radiation measurement. Energy Proc. **9**, 230–237 (2011)
26. De Giorgi, M.G.; Ficarella, A.; Tarantino, M.: Assessment of the benefits of numerical weather predictions in wind power forecasting based on statistical methods. Energy **36**, 3968–3978 (2011)
27. Yang, H.T.; Huang, C.; Huang, Y.; Pa, Y.: A weather-based hybrid method for 1-day ahead hourly forecasting of PV power output. IEEE Trans. Sustain. Energy **5**, 917–925 (2014)
28. Basaran, K.; Özçift, A.; Kılınç, D.: A new approach for prediction of solar radiation with using ensemble learning algorithm. Arab. J. Sci. Eng. **4**, 1–13 (2019)
29. Douiri, M.R.: A predictive model for solar photovoltaic power based on computational intelligence technique. Arab. J. Sci. Eng. **2**, 1–18 (2019)
30. Haykin, Simon: Neural Networks: A Comprehensive Foundation. Prentice Hall PTR, Upper Saddle River (1998)
31. Zhang, Z.: Artificial neural network. In: Multivariate Time Series Analysis in Climate and Environmental Research, pp. 1–35. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-67340-0_1
32. Piliougine, M.; Elizondo, D.; Mora-Lopez, L.; Sidrach-de-Cardona, M.: Multilayer perceptron applied to the estimation of the influence of the solar spectral distribution on thin-film photovoltaic modules. Appl. Energy **112**, 610–617 (2013)
33. Sundar, C.; Chitradevi, M.; Geetharamani, G.: Classification of cardiotocogram data using neural network based machine learning technique. Int. J. Comput. Appl. **47**, 93–103 (2012)
34. Hagan, M.T.; Menhaj, M.B.: Training feedforward networks with the Marquardt algorithm. IEEE Trans. Neural Netw. **5**, 989–993 (1995)
35. Mirjalili, S.; Mirjalili, S.M.; Lewis, A.: Grey Wolf optimizer. Adv. Eng. Softw. **69**, 46–61 (2014)
36. Kennedy, J.; Eberhart, R.: Particle swam optimization. In: IEEE International Conference on Neural Networks, pp. 1942–1948 (1995)
37. Lin, C.; Yang, K.; Huang, J.; Tang, W.; Qin, J.; Niu, X.; Chen, Y.; Chen, D.; Lu, N.; Fu, R.: Impacts of wind stilling on solar radiation variability in China. Sci. Rep. **5**, 1–7 (2015)
38. Srivastava, P.; Dey, S.; Agarwal, P.; Basil, G.: Aerosol characteristics over Delhi national capital region: a satellite view. Int. J. Remote Sens. **35**, 5036–5052 (2014)
39. Liu, J.; Fang, W.; Zhang, X.; Yang, C.: An improved photovoltaic power forecasting model with the assistance of aerosol index data. IEEE Trans. Sustain. Energy **6**, 434–442 (2015)
40. Chaabane, M.; Masmoudi, M.; Medhioub, K.: Determination of Linke turbidity factor from solar radiation measurement in northern Tunisia. Renew. Energy **29**, 2065–2076 (2004)
41. Wang, L.; Chen, Y.; Niu, Y.; Salazar, G.; Gong, W.: Analysis of atmospheric turbidity in clear skies at Wuhan, Central China. J. Earth Sci. **28**, 729–738 (2017)
42. Kasten, F.: A simple parameterization of the pyrheliometric formula for determining the Linke turbidity factor. Meteorol. Rdsch. **33**, 124–127 (1980)

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/336513567

Digital Democracy and India: Study of Evolution in Citizen

Chapter · October 2019

CITATIONS
0

READS
100

2 authors:

Sharique Hassan Manazir
Jawaharlal Nehru University
12 PUBLICATIONS **3** CITATIONS
SEE PROFILE

Madhav Govind
Jawaharlal Nehru University
25 PUBLICATIONS **71** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

e-Participation platform based shared bicycle system & urban infrastructure sustenance: Critical analysis of "Greenolution" bicycle sharing ecosystem in JNU View project

All content following this page was uploaded by Sharique Hassan Manazir on 14 October 2019.

The user has requested enhancement of the downloaded file.

Emerging Trends in Information Technology

Emerging Trends in Information Technology

Contemplating Some Crucial Research Issues

Editor

Dr. Nidhi Arora

Proceedings: Selected Peer Reviewed Research Papers
of
UGC Sponsored National Conference on Emerging Trends in Information Technology
(NCETIT-2019)

held during

1st and 2nd August 2019
at
Kalindi College, University of Delhi, India

BLOOMSBURY
NEW DELHI • LONDON • OXFORD • NEW YORK • SYDNEY

BLOOMSBURY INDIA
Bloomsbury Publishing India Pvt. Ltd
Second Floor, LSC Building No. 4, DDA Complex, Pocket C – 6 & 7,
Vasant Kunj, New Delhi 110070

BLOOMSBURY, BLOOMSBURY PRIME and the Diana logo are
trademarks of Bloomsbury Publishing Plc

First published in 2019

Copyright © Nidhi Arora, 2019

Nidhi Arora have asserted their right under the
Indian Copyright Act to be identified as the Editor of this work

All rights reserved. No part of this publication may be reproduced or
transmitted in any form or by any means, electronic or mechanical,
including photocopying, recording, or any information storage
or retrieval system, without prior permission in
writing from the publishers

Bloomsbury Publishing Plc does not have any control over,
or responsibility for, any third-party websites referred to or in this book.
All internet addresses given in this book were correct at the time of going to
press. The author and publisher regret any inconvenience caused if
addresses have changed or sites have ceased to exist, but can accept
no responsibility for any such changes

ISBN: 978-93-89165-99-9

2 4 6 8 10 9 7 5 3 1

Typeset by Shinu, Moti Nagar, New Delhi Printed
and bound in India by Replika Press Pvt Ltd

To find out more about our authors and books, visit
www.bloomsbury.com and sign up for our newsletters

Acknowledgements

The conference was possible with the financial support of UGC. I, on behalf of Kalindi College and entire staff of our Department of Computer Science, extend my gratitude to UGC for all their extended support and help. I would like to express my deepest gratitude to the Patron and Advisor of this conference, Principal Kalindi College, Dr. (Ms.) Anula Maurya, for providing all kind of administrative support while planning and execution of this conference. I thank all the invited keynote speakers for their enlightening talks. I sincerely owe my gratitude to all the Technical Programme Committee members for their timely review and comments on all the submissions. I sincerely thank all the teaching staff of Department of Computer Science, Kalindi College: Dr. Vandana Gupta, Ms. Shalini Sharma, Dr. Reena Jain, Ms. Anshula, Dr. Yogendra, Dr. Susheel, Ms. Ramya, Ms. Rajani, Mr. Harshvardhan; Computer lab staff Ms. Mamta Sachdeva, Ms. Asha Khantwal and Mr. Rajiv; and all the student team members for their support in successfully organizing this conference.

I am also thankful to Bloomsbury Publishing India Pvt. Ltd. for taking interest to publish this book. We are thankful to all the chapter contributors for providing their revised papers and supporting us whenever any information was required for completing the manuscript of the book. I wish all the readers a knowledge enhancing experience while reading this book.

Best wishes!

Nidhi Arora

Preface

Twenty first century has been defined by advancement in information technology and communication. It has become an integral part of our daily life and, has been the fastest growing area. The world has now become a global arena, with no virtual boundaries. Research is being conducted at a very fast pace, and practically every day there is addition to the existing sphere of knowledge with primary focal areas being Artificial Intelligence, Cloud Computing, Internet of Things, E-learning, Big Data Analytics and many more.

This book "Emerging Trends in Information Technology" is the outcome of the National Conference on Emerging Trends in Information Technology organised by Department of Computer Science, Kalindi College during 1st and 2nd August 2019. The conference was fully funded by University Grants Commission. The book consolidates some of the high quality original papers from the set of papers presented in the National Conference on Emerging Trends in Information Technology (NCETIT 2019). The main aim of the conference was to bring together experts from academic institutions, industries, leading engineers, students, scientists & scholars to facilitate close interaction for exchange of views & dissemination of knowledge with prime focus on the practical challenges encountered and the solutions adopted in all the fields related to Computers and Information Technology. Parallel invited talks on both days by experts from renowned Indian Universities attracted large number of audience from various colleges. Due to wider applicability of the conference theme, large number of abstracts and full papers were submitted and presented, out of which only sixteen high quality papers were selected to be published as chapters in this book after passing through peer review and plagiarism check though abstracts of all the presented papers were published in a souvenir booklet released during the conference.

Introduction

Nidhi Arora[1]

In today's fast growing highly connected world, where technology becomes obsolete even before reaching masses and data becomes big and diverse even before information could be fetched, there lays a tremendous pressure to keep technological growth pace in line with the issues of information and data handling. Each day brings new challenges in the field of emerging trends related to technological growth for efficient information handling.

There has become a direct relationship between a country's economy and its technological state of growth, since the technology has sweeped in almost every facet of life not even in urban brackets but also in rural setups. Hence, this is the prime research area in present world which makes it inevitable for researchers from all disciplines, ranging from humanities and pure sciences, to work in isolation leaving technology behind. Hence, technology and its applications is an interdisciplinary area where diverse research fields are centrally connected giving this field of research many dimensions and directions. Some hot topics one can see and identify are Big Data, Cyber Security, Social Media Analytics, Cloud Computing, Optimization, Artificial Intelligence, Machine Learning, Social Networks, Digital Image recognition and E-Services including E-Governance, E-Learning, E-Commerce and so on. The list mentioned is not consolidated but has many more sub fields which are independent and fascinate attention among masses.

To be truly able to apply and use any of these or all of these emerging technologies, one not only needs to fine tune its understanding regarding advantages but also the disadvantages of these technologies. Even though lot of knowledge is present on web portals, common physical platform of discussion, where mixed genres are facilitated to share their knowledge remains the necessity. This book provides a common medium of resource, where in each chapter presents a crucial discussion at length on a particular area of research in information technology field. The chapters

[1]Assistant Professor, Department of Computer Science, Kalindi College, University of Delhi
E-mail : nidhiarora@kalindi.du.ac.in

are a collection of papers presented in a National Conference on Emerging Trends in Information Technology wherein researchers and industry experts from all over India discussed these topics and facilitated peer community to enhance their knowledge about the state of the art in these fields.

Chapter 1 presents the current state of the art on classification methods of supervised machine learning for cyber security intrusion detection. The chapter enlisted approximately 40 related studies on cyber security for IDS that focusing on the architecture (single, hybrid or ensemble classifier), classifier algorithms using well-known public datasets for experiment, the classification algorithms based on supervised learning were evaluated using metrics (F1score, precision, false positive rate, false alarm rate, accuracy) and feature selection for supervised machine learning algorithms.

Chapter 2 focuses on in-depth analysis of two cryptographic techniques namely RSA and ECC in digital signature applications. It also compares the performance and security aspects of these techniques.

The epicenter of the study presented in **Chapter 3** deals with the role and impact of social media on Indian students. Parameters such as political mobilization, political influence, opinion, changing behaviour in youth are used to evaluate the role of social media. The methodology used in the research is mainly conducting (most of the data are collected using a questionnaire as a research tool).

Chapter 4 discusses sentiment analysis, a key research area among IT researchers. It discusses application of two different approaches are presented to extract the sentiment of the reviews at the sentence level. In the first approach, TextBlob library was used to calculate the polarity of the sentence whereas the second method follows a rule-based approach in which two Lexicons of polarised words (negative and positive) were used.

Chapter 5 presents various innovative frameworks to understand the transition of citizens' direct participation in Democracy to their e-participation in Digital Democracy. This study provides new insight into how different tools are used to engage citizen in governance through various e-participation models and what are challenges to enable the citizen to participate in the democratic process in India.

In-depth literature survey on the state of the art of Big Data paving its way to E-Business is discussed in **Chapter 6.** The survey focuses on the

different sorts of data held and its inverse usage for e-business along with different ways of providing security and safety for data corresponding to anonymity. It also highlights the issues in big data with respect to e-businesses and tries to trigger the maximum usage of big data applications effectively.

Chapter 7 revolves around the discussion on formulating an unconstrained nonlinear programming problem and determines the optimal solution using gradient search technique. Also the authors discuss the post optimality analysis by varying the coefficient of the objective function and the obtained optimal solution values respectively.

Chapter 8 authors beautifully discussed postmodern techno culture and the need for feminist intervention by use of mobile applications for women safety in India. The presented research explore the author's previous case study of "My Safetipin" mobile application from a fresh perspective of analysing female participation not as decision makers to look at possible reasons for failure of this e-participation platform.

Chapter 9 gives an overview on some significant works in the field of Recommender System on the basis of its three techniques namely collaborative filtering, content based filtering and hybrid filtering. The Chapter also includes various issues and application domains of Recommender System.

Chapter 10 investigates synchronization of two identical fractional order (FO) Chua's chaotic system with disturbance using the technique of Adaptive Sliding Mode. The introduced techniques has many applications in secure communication and neural network.

Chapter 11 presents a model developed for single server finite capacity Markovian queuing incorporating feedback of customers, encouraged arrivals and the concept of reverse reneging. In today's highly volatile uncertain environment, understanding customer behavior in advance is of utmost importance for any firm in order to stay ahead of competitors. Firms need to design effective operational strategies for overall functioning of the system.

Chapter 12 revolves around the discussion to absorb the ideas of Cyber Security alongside the unsafe impacts of Malicious Software and finding proper ways of protecting the system and simultaneously highlighting the cyber security on a global scale and especially in India. Malicious software, or Malware, is a Cyber Security threat that affects the user's computer system by exploiting the system's vulnerabilities.

Chapter 13 aims at assessing and benchmarking the operational efficiency of 12 scheduled (domestic and international) passenger airlines in India during the year 2016–17. Projections are discussed and improvements are mentioned which will help the decision makers to improve their companies performance.

Chapter 14 presents an in-depth survey on the applications of big data in real world. Use of Big data in all spheres of personal & professional lives be it agriculture, banking, travelling, healthcare, financial institutions, food industry, power sector, managing smart cities and homes is discussed at length.

Chapter 15 authors discussed hybrid projective combination synchronization (HPCS) between chaotic generalized three species Lotka-Volterra (GLV) biological system using adaptive control method. Lotka and Volterra introduced the well-known examples of basic biological models of two species predator and prey model. numerical simulations are performed to show the efficiency and accuracy of the suggested scheme using MATLAB.

Chapter 16 focusses on E-learning trends among students. One case study that provides a direct comparison of direct classroom teaching and teaching through E-Learning environment is presented to know whether there is any major difference in students learning and their performance when learning is through E-learning based approach and through classroom lecture based learning approach. In addition to this a survey was conducted in this study which allows the students to give their feedback and suggestions on both the learning approaches.

Contents

Acknowledgements *v*
Preface *vii*
Introduction *ix*

1. A Literature Survey on Cyber Security Intrusion Detection Based on Classification Methods of Supervised Machine Learning 1
   Anu Yadav and Ela Kumar
2. In-depth Analysis of the Performance of RSA and ECC in Digital Signature Application 15
   Arokia Ramya Terrance, Shalini Sharma, Anu Sharma, Kanika Tyagi, Nikita Shokeen and Nitya Saini
3. Social Media—A Platform for Political Discourse: An Evaluation 28
   Abhishek K. Singh
4. Feedback Ratings are Better Indicators of User Sentiments than Whole Review 38
   Geeta Arneja and Harmeet Kaur
5. Digital Democracy and India: Study of Evolution in Citizen e-Participation Frameworks 48
   Sharique Hassan Manazir and Madhav Govind
6. A Literature Analysis on Big Data Paving Its Way to E-Business 64
   Suman Madan, Kirti Bhardwaj and Shubhangi Gupta
7. A MATLAB Programming Applied in Gradient Search Method to Solve Quadratic Optimization Problem 78
   L.N. Das and Sanyam Gupta
8. Postmodern Techno Culture and the Need for Feminist Intervention: Study of Mobile Applications for Women Safety in India 90
   Rubina and Madhav Govind
9. Recommender System: Principles, Techniques and Application 98
   Sushma Malik, Yukta Pandey and Amanpreet Kaur

10. Adaptive Sliding Mode Disturbance Observer Control Base Synchronization in a Class of Fractional Order Chua's Chaotic System 107
Ayub Khan and Uzma Nigar
11. A Single Server Finite Capacity Markovian Feedback Queuing System with Encouraged Arrivals and Reverse Reneging 119
Bhupender Kumar Som and Sunny Seth
12. Study on Cyber Security and Malware Protection 127
Tripti Lamba and Shivani Kandwal
13. Efficiency Evaluation and Benchmarking of Air Carriers in India for the Year 2016–17 Using Data Envelopment Analysis (DEA) 141
Bharti Seth, Punita Saxena and Shalini Arora
14. Big Data Applications in Real-World 158
Shalini Sharma, Naresh Kumar and Kuldeep Singh Kaswan
15. Adaptive Control and Hybrid Projective Combination Synchronization of Chaos Generated by Generalized Lotka-Volterra Biological Systems 174
Ayub Khan and Harindri Chaudhary
16. E-Leaning: Emerging Trend in Learning 185
Diksha Grover and Shweta Paradkar

Contributors 197

Author Index 199

A Literature Survey on Cyber Security Intrusion Detection Based on Classification Methods of Supervised Machine Learning

Anu Yadav[1] and Ela Kumar[2]

ABSTRACT: *In the era of big data, with the rapid change in communication technology and the internet, the number of data features increasing and decreasing intrusion detection system performance in terms of accuracy. The growth of the Internet and its usage has risen a problem of network attacks; hence cyber security has become an important issue. The intrusion detection system (IDS) is used to classifying the patterns for attacks on the network and to identify unauthorized access due to cyber-attacks. In this literature survey paper enlisted approximately 40 related studies on cyber security for IDS that focusing on the architecture (single, hybrid or ensemble classifier), classifier algorithms using well-known public datasets for experiment, the classification algorithms based on supervised learning were evaluated using metrics (F1score, precision, false positive rate, false alarm rate, accuracy) and feature selection for supervised machine learning algorithms. This paper also included short descriptions of IDS and public dataset.*

Keywords: Machine Learning, KNN, SVM, LDA, Naïve Bayes.

1. INTRODUCTION

Today imagine our life without the Internet can't credible. There are lots of internet users who are regular user in cyberspace for various purposes such as educational, professional, commercial, and socialization. Internet increases users Tendencies to share their secret data and details, pictures and Opinions about the world. Hence it is an important aspect of

[1] PhD CSE Research Scholar, Department of Computer Science and Engineering IGDTUW, Delhi, India. E-mail: anuyadavcse@gmail.com

[2] Professor and HOD, Department of Computer Science and Engineering IGDTUW, Delhi, India. E-mail: ela_kumar@rediffmail.com

machine learning approaches to make a reliable model sooner that does not affect by cyber-attacks. This paper presents the outcome of a literature review of supervised machine learning (SML) for classification methods for cyber security. This paper enlisted more than 50 related studies which are published in the last few years and highly cited papers.

The ML (Machine learning) is a subfield of AI (Artificial Intelligence), which allows the machine to learn without programming. The ML is divided into four main categories based on the learning methods: supervised, unsupervised, semi-supervised and reinforcement.

The IDS (intrusion detection system) is used to discovered attacks on the cyber world by identifying unauthorized access, and modification of information for a particular machine learning system. The intrusion is divided into three main categories based on the detection for cyber analytics of IDS: First, the Misuse-based or Signature-based method, which detects the well-known type of attacks in the cyber world by using signatures of that particular attack by using frequent manual updates of the DB (database). Second, an anomaly-based method, which builds a network and system behaviour, and identifies anomalies. The last one is the Hybrid method which is a combination of misuse based and anomaly-based detection, they increased DR (detection rates) of known intrusions and decrease the FPR (false positive rate) for unknown attacks. The IDS is divided into two main categories based on the detection of intrusive behaviour: the host-based and network-based.

This literature paper focuses on SML Classification techniques for cyber security, the primary focus of our paper is to meet certain criteria for SML. The queries performed on IEEE, Springer, Google Scholar and Science Direct using machine learning, cyber security, KNN, SVM, LDA, Naïve Bayes and intrusion detection. The paper selected based on popular technique, the number of citations, UGC approved journal and publication year.

2. IDS PUBLIC DATASETS

The raw data is a basic component of research, which can be obtained directly or from existing public datasets. The dataset chosen should be perfect for machine learning approaches. The DARPA 1998/1999 intrusion detection dataset [1] which is used by most of the experiments conducted on IDS, it is collected and published by the CSTG (Cyber Systems and Technology Groups) of the MITLL (Massachusetts Institute of Technology Lincoln Laboratory).

The KDD cup99 dataset is more derived and contains 41 features/ attribute. Dataset aims to be freely available for survey/research and which is more realistic for IDS, a different version of this dataset provided by contest. The dataset has a connection which is labelled as normal or attack (DOS (Denial Of Service), R2L (Root 2 Local), U2R () and Probe) data.

The newly refined version of KDD Cup99 dataset is called as NSL-KDD, which introduced in KDDMT (Knowledge Discovery & Data Mining) Contest [39] to improve KDD Cup99 dataset. It identifies attribute is from intrusive or normal connections and contains 22 various attack types.

3. CLASSIFICATION METHOD FOR SML

3.1 *SVM*

SVM (Support Vector Machine) is most accurate and robust supervised learning model and construct hyperplane in the high dimensional feature space between two classes with maximum margin. As shown in figure line between support vector points is hyperplane and the closest vector points near separation hyperplane are support vector points [2].

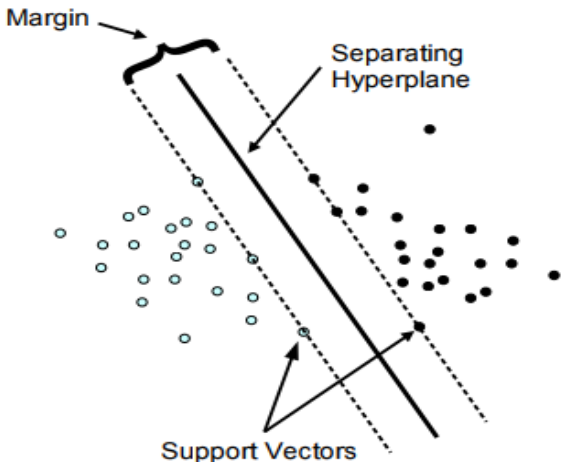


Fig. 1: Support Vectors, Hyperplane and Margin

[3] shows a combination of DBN (Deep belief network) as a DRA (dimensionality reduction algorithm), one-class SVM and anomaly detection technique for not labelled set of data.

[4] they used multi-class SVM and feature selection by chi-square for IDS model to decrease the FAR (false alarm rate) and increase the DR (detection rate) and classification accuracy (ACC) of network attacks on NSL-KDD dataset. In SVM use two parameters! (gamma) and C (over fitting constant) used in parameter tuning technique for RBF kernel (Radial Basis) optimization Function. They used 3 performance metrics and found that the PCA-GA-SVM, KPCA-GA-SVM, and SVM (single) methods have less time for training than N-KPCA-GA-SVM model.

[5] they explained a hyper-heuristic SVM configuration process as a bi-objective optimization problem for accuracy and complexity. The hyper-heuristic framework consists of high-level strategy (the search performance to control the selection) and low-level heuristics (to generate a new SVM configuration) with the Pareto set of SVM approach. Their framework evaluated Microsoft malware big data classification and anomaly intrusion detection cyber security problems.

[6] they use ELM and SVM in multi-level hybrid intrusion detection method for known and unknown attacks and obtained improved detecting efficiency. They used K-means to build a high-quality training dataset using KDD Cup 1999 for improving the performance of the classifiers and reduce the training time of classifiers. Their results show an accuracy of 95.75% with an improved detection rate of R2L, DoS, and U2R attacks.

[7] FWP-SVM-GA (Feature selection Weight Parameter optimization of SVM based on Genetic Algorithm) for human-centred IDS on KDD Cup 99 datasets. First, it optimizes mutation and crossover probability of GA based on the population evolution algebra search strategy and fitness function value that increases true positive rate, decreases the classification time and decreases the SVM error rate. The proposed model increases intrusion detection rate (96.61 and 100), true positive rate (0.9066%), and the accuracy rate (0.9975%) but it decreases FPR (3.39 and 0) and FNR (0.07%) for 10000 training set.

The K-means and multi-level SVM algorithm [8] for network IDS on NSL-KDD dataset which divides the data into different clusters by using Kmeans and mark the abnormal cluster using multi-level SVM for classification. The results show that increases the detection rate and reduce the false alarm rate by comparing the proposed model with SVM+CSOACN and C4.5+SVM. Experimental results showed that TPR and FPR for different attack types are DoS 98.64% (True Positive) 4.52%

(False Positive), Probe 97.81% (TP) 3.85% (FP), R2L 89.42% (TP) 20.21% (FP) and U2R 95.54% (TP) 6.97% (FP).

The problems due to a virus attacking the application layer network protocol, TCP communication protocol is analysed by IDS based on clustering and SVM [9]. Combination of unsupervised FCM and supervised SVM to calculate the distance between communication data and cluster centre. Experimental results show that improved accuracy and detection precision with the reduction in detection time and training time. The outcome of SVM detection accuracy of the test set is 94.8571% and time is 0.76647 sec where the outcome of FCM 96.8571% (accuracy), time is 0.56155 sec for test set and training set the accuracy is 97.3333% and the time is 0.49516 sec.

[10] used ADFALD dataset for host-based anomaly detection system with the one-class SVM algorithm. Their experiments show that model removed duplicate entries as well as achieve satisfactory performance and reduce the computational cost but the recognition rate will remain low.

[11] described the separation of control plane and data plane by using SDN (Software Defined Networking) controller for DDOS (Distributed Denial of Service) attacks. Among all algorithms, SVM classifier gives higher accuracy and less false positive rate.

[12] explained the SVM classifier method by combining feature selection and multiple class NSL-KDD CUP 99 intrusion detection dataset. The method best performance with an F1 score of 0.99 and worst performance F1 score of 0.77 which unable to detect intrusion in network. Their results show classification accuracy 91% using 3 features, 99% using 36 features, and 99% when 41 training features selected.

3.2 *KNN*

The KNN (K nearest neighbour) algorithm is simple, depends on accuracy as well as prediction and do not contain any learning phase hence called as a lazy learner but memorize the training data [13]. The KNN classifies the new data from the available dataset based on similarity measures using distance metrics (Euclidean distance, Mahalanobis or Hamming distance) to predict new unseen data point.

[14] The unbalanced normal data and intrusion data are obstacles which are solved by FCM (fuzzy C-means) and KNN algorithm to construct feature vectors in FCM based on clustering central point and fuzzy

matrix wherein KNN based on classifier on KDD-Cup 99 dataset. The FCM intrusion detection model is superior to the TANN (Triangle Area Based Nearest Neighbours) and CANN (combining cluster centres and nearest neighbours) models the accuracy of FCM+KNN is 98.73% where TANN is 98.47% and CANN is 96.07%, the detection rate of FCM+KNN is 96.23% where TANN is 94.5% and CANN is 86.05% and false alarm rate for FCM+KNN is 0.28%, TANN is 0.4% and CANN 0.75%.

[15] The fusion approach to detect intrusion that hybridises different classifiers of SVM, k-NN, and Primal-Dual PSO (Particle Swarm Optimisation) on the KDD99 dataset. There result shows that the highest accuracy of SVM, k-NN and pdAPSO is about 97%, 98% and 99.8% respectively, for R2L attack. This fusion approach generating accuracy of 98.55%.

[16] The intrusion detection using KNN-ACO (ant colony optimization) and SVM method performed on KDD1999 Cup dataset for improving classification accuracy and generates low false alarm rate. The overall accuracy rate for KNN-ACO is 94.17%, Back propagation neural network (BPNN) is 93% and SVM is 83%. The overall false alarm rate for KNN-ACO is 5.82%, BPNN is 6.90% and for SVM is 16.90%.

[17] Explained the CANN and KNN classifier performance by combining them with K-MEAN clustering on NSL-KDD dataset. To classify their data, they used the nearest neighbour (KNN) and the farthest neighbour (k-FN). The ten times comparison of CANN with k-FN CANN results shows improved accuracy and detection rate with reduces false alarm rate.

[18] To detect unknown attacks with insufficient training data by using Fuzzy C-Means algorithm, distance-weighted k-Nearest Neighbours classifier and Dempster Shafer Theory to assess the belief function and pignistic probabilities on KDD'99 datasets. The experimental result ofthe Fuzzy and DST recorded for DOS class 5.22% others 6.57% to 11.18%, Probe 3.90% others 6.21% to 12.07%, U2R 8.12% others 13.13% to 35.06%, R2L 9.98% others 10.29% to 24.60% and UA 11.25% others 40.68% to 47.94%.

[19] They proposed two-layer hybrid classification and detection process for network attacks detection on KDD99 dataset where first layer have GBDT classifier for identifying DoS attacks and second layer of KNN and stacking ensemble classifier which optimized by FOA (Fly Optimization Algorithm) to divide the non- DoS data from Normal, R2L,

Probe and U2L. Evaluation implemented by accuracy, precision rate and recall rate. The classification of DoS data with KNN on attacks like a smurf, neptune pod, back, and teardrop has high recall rate and precision rate of smurf and neptune is higher than 99% but back and the pod is higher than 60%. The higher accuracy SMOTE+GBDT+SEM method compared with SVM+ANN and SMOTE+GBDT methods.

[20] proposed KNN-TSA intrusion detection model based on TSA (tree-seed algorithm) reduces feature redundancy and extract feature for input and KNN is used for classification on KDD CUP 99 datasets to improve the accuracy and efficiency of network IDS. The experimental outcome shows Accuracy of KNN-PSO 83.65%, KNN-GA 78.78% and TSA-KNN 87.34%.

A hybrid approach of k-means clustering and KNN classification algorithm is presented in [21] to detect intrusion on the cloud system using NSL-KDD data set for evaluation of experiments. The experimental outcome proves the ability of the proposed model for detection of intrusion attacks.

3.3 *LDA*

The LDA (Linear Discriminant Analysis) is a simple, complex and robust classification method of SML proposed by RA fisher (1936), which provides accuracy based on linear combination of predictors/ variables for separating 2 classes/targets using Mahalanobis distance, coefficient vector, class probability, covariance matrix and score function.

[22] LDA- LR (Linear Discriminant Analysis and Logistic Regression) anomaly-based IDS model which performs well for binary class and multiclass classification problem. LDA – LR experiment on NSL-KDD dataset and comparison analysis between performances of other IDS based on Naive Bayes, C4.5 and SVM and LDA-LR. The Comparison of intrusion detection models on the binary class for 23 Features and measured accuracy (ACC) and DR in terms of percentage where SVM have 99.63% (ACC) 99.16% (DR), Naive Bayes 98.16% (ACC) 89.97% (DR), C4.5 have 99.85% (ACC) 90.95% (DR), LDA have 99.83% (ACC) 90.89% (DR), And LR have 98.27% (ACC) 91.59% (DR). The ACC and DR for LDA and LR models on the multi-class is 95.05% (ACC) &95.05% (DR) for LDA and 95.31% (ACC) & 98.74% (DR) for LR.

LDA feature extraction [23] method used to detect cyber-attacks and remove the useless information from original data by using truncated

mean to estimate the class mean vector. The experiments consist the 1000 training samples with the composition of 100 normal, 100 DOS, 100 R2L, 100 PROBE and 50 U2R for evaluation of truncated LDA (with different *P)* on KDD cup99 and NSLKDD dataset. Detection Rate and FPR of truncated LDA, LDA and median LDA compared and found that FP lower than 2.3%.

R1-PCA (Rotational Invariant L1-norm Principal Component Analysis) and median LDA (median Linear Discriminant Analysis) algorithm for Anomaly IDS [24]. The reduction dimension algorithms using L1-norm instead of the PCA Euclidean distance and LDA uses the class median vector to estimate the CMV (class mean vector) on KDDcup99 dataset. The experiments consist the 1000 training samples with the composition of 100 normal, 100 DOS, 100 R2L, 100 PROBE and 50 U2R for evaluation FPR and DR.

The ensemble of PCA-LDA (Principle Component Analysis -Linear Discriminant Analysis) feature extraction algorithms [25]. The experiment outcome shows that feature reduction, increased detection rate and accuracy. They compared PCA-SVM and LDA-SVM in terms of ACC, FP and FN. In PCA-SVM Overall-accuracy is 0.8902, False-positive is 0.0089 and False-negative is 0.1704 where In LDA-SVM ACC is 0.8993, FP is 0.0203, and FN is 0.1133.

The OPMCA (optimal mean PCA) with LDA [26] on KDDcup99 and NSL-KDD provide accuracy to improve defence from cyber-attacks and increase detection rate and decrease FPR. Dataset consists of 100 normal, 100 DOS, 100 R2L, 100 PROBE and 50 U2R with training data at 1000 samples. The outcome is compared with OMPCA+LDA to LDA and PCA datasets.

The classification algorithms LDA-PCA to identify Intrusion and anomaly on NSL-KDD dataset [27]. The PCA selecting eigenvectors and correlation matrix which obtained correlation between two variables count (94.4%).

D-LDA (Direct Linear Discriminant analysis) algorithm [28] proposed using work of PCA and two-stage PCA+LDA algorithms which eliminates singularity problem to enhance defence against cyber-attacks. The experiment performed on 10% KDD Cup 99 dataset and outcome is accuracy DR is 97.75%, false alarm rate 0.83% and DOS attack has the best classification which is 82%.

[29] GLDA-SVM (Genetic Linear Discriminant Analysis and Support Vector Machines) IDS is a hybrid feature selection approach of LDA, Genetic Algorithm and SVM Kernels as classifiers on NSL-KDD dataset. The experiment outcome shows that this efficient soft computing approach provides high DR and low false alarms, maintaining low cost and low detection time for the intrusion.

3.4 Naïve Bayes

The NB (Naïve Bayes) classification technique is a powerful predictive model but simply based on Bayes theorem and probabilistic classifiers with assumption between features. The NB used in SPAM filtering, diagnosis, new categories and weather prediction.

If irrelevant data is processed than computational power and resources are wasted hence to improve IDS more efficient the false positive rate must be reduced by combining SVM, Naïve Bayes, and decision trees [30]. The overall accuracy of this model is 99.62% and the false-positive rate of 1.57%.

[31] Explained IDS based on Bayesian network using the wrapper approach to eliminate irrelevant features and reduce time complexity and high accuracy. Used NSL-KDD datasets with WEKA and they obtained overall performance 98.2653 in terms of accuracy, 98.9 precision, 1.73 error rate, and 0.007 low false-positive rates.

The IDS based on Bayesian network and feature selection for network traffic data which is evaluated using Kyoto dataset [32]. The Experimental results show that approach has a high level of detection rate of 99.9%, high efficiency and low FAR (0.015%) in detecting network attack.

The two algorithms hybrid approach which consist PCA for feature extraction as well as reduces high dimensionality and Naive Bayes for classification on the KDD99 dataset [33]. The outcome indicates a high accuracy rate and lower FPR. The performance (accuracy) of the Naive Bayes for Attack types DOS is 87.47%, U2R is 95.76%, PROBE is 97.12%, and R2L is 72.16%. According to result as the number of features decreases accuracy, DR and FPR increases, i.e. for 41 number of feature DR is 79.461%, FPR 15.362% and ACC is 80.143%; for 35 number of features, DR is 82.131%, FPR 4.680% and ACC is 82. 312%; for 10 number of feature DR is 96. 52%, FPR 1.16% and ACC is 96.240%.

[34] IDS detect security problems, improves the detection rate and reduced the false alarm rate of TLMD (Two Layers of Multi-class

Detection) that performed on KDDcup99. TLMD method consists of the C5.0 and the Naive Bayes algorithm for adaptive network IDS, which handles imbalanced datasets, deal with a continuous attribute, and reduce noise in the dataset. The comparison of Average Accuracy of SVM is 83.97%, C5.0 is 92.77% and TLMD Method is 93.32%, which proved that TLMD has the highest accuracy among them.

[35] The k-means clustering and naïve Bayes classification hybrid method used to overcome issues of Misuse detection and anomaly detection mechanisms for IDS. The distance sum-based features (the distances from each sample to several centroids found) are used for classifier training and detection by using a clustering algorithm. The experiment result provides improved accuracy for Probe, R2L and U2R attacks on NSL-KDD dataset, but poor results for Normal and DoS attacks, the outcome is Normal 74.11%, DoS 86.05%, Probe 92.48%, R2L 32.02% and U2R 19.0%.

The Naive Bayesian network algorithm for IDS based on the PCA [36] which calculate the CV (characteristic value) of the data by network attack then extract the properties and principal components through the PCA. The experiment performed on KDD CUP 99 dataset which including 4 types of attacks Probe, DoS, U2L and R2L whose 41 attributes taken as extraction, the outcome showed that the algorithm reduces the data dimension and improve the detection rate.

[37] Developed an IDS using two techniques which are based on neural network, first PCA method for feature extraction and second technique is combination of Bayesian (for noise reduction) and KPCA (kernel principal component analysis for feature extraction). The experiment is implemented using java technology, learning rate value=0.7, Hidden unit=4 and performed on KDD Cup 99 dataset and its result gives high accuracy and low error rate. The experiment outcome produced based on two methods IDS based on neural network using PCA (ACC 90.13%, ERROR 9.87%, TIME 18874 ms, MEMORY 21735120 bytes) and IDS based on neural network using Bayesian and KPCA (ACC 92.07%, ERROR 7.93%, TIME 9010ms, MEMORY 38000136 bytes) results show that the accuracy%, error % and time are better.

The online Naïve Bayes (NB) classifier [38] detect normal or bad intrusion continuously and keep updating mean and standard deviations for features on KDD CUP 99 dataset. The approach is able to perform incremental update. The performance of different class problems (23-class problem, 5-class problem, and 2-class problem) are compared based on

F1 score and type of class. Then comparison between online NB and online KNN for *Normal* 0.2440% (NB) 0.2917% (KNN), *DoS* 0.8951% (NB) 0.8997% (KNN), *Probe* 0.0168% (NB) 0.0161% (KNN), *R2L* 0.0041% (NB) 0.0043% (KNN) and *U2R* 0.0422% (NB) 0.0506% (KNN). The *Accuracy* comparison of machine learning methods on 2class problem is0.93% for Online NB, Perceptron Learning *is* 0.87% and for Online k-NN 0.96%.

4. CONCLUSION

As cyber-attacks increasing due to easy internet usage and access, it is important to study a systematic survey on improved IDS and machine learning algorithms for cyber security concerns. This paper presents a literature review of supervised machine learning for classification methods security on the cyber world using different public Datasets. The paper focused on recent year's improvement in the performance of supervised machine learning in the area of IDS. It is not easy to choose the best effective approach because each method of classification for an IDS has its pros and cons, most of the paper provide comparisons between there proposed method with another approach.

REFERENCES

[1] Available: http://kdd.ics.uci.edu/databases/kddcup99.

[2] Corinnacortes and Vladimir, Vapnik, "Support-Vector Networks", Machine Learning, 20, pp. 273–297, 1995.

[3] Erfani, M.; Rajasegarar, Sutharshan; Karunasekera, Shanika and Leckie, Christopher, "High-dimensional and large-scale anomaly detection using a linear one-class SVM with deep learning Sarah", *Pattern Recognition,* Vol. 58, pp. 121–134, 2016.

[4] Ikram, Sumaiya Thaseen and Cherukuri, Aswani Kumar, "Intrusion detection model using fusion of chi-square feature selection and multi class SVM", *Journal of King Saud University Computer and Information Sciences,* Vol. 29, pp. 462–472, 2017.

[5] Sabar, Nasser R., Yi, Xun and Song, Andy, "A Bi-objective Hyper-Heuristic Support Vector Machines for Big Data Cyber-Security", *IEEE Translations and content mining,* Vol. 6, pp. 2169–3536, 2018.

[6] Al-Yaseen, Wathiq Laftah; Othman, Zulaiha Ali and Nazri, Mohd Zakree Ahmad, "Multi-level hybrid support vector machine and extreme learning machine based on modified K-means for intrusion detection system", *Expert Systems with Applications,* Vol. 67, pp. 296–303, 2017.

[7] Tao, Peiying; Sun, Zhe and Sun, Zhixin, "An Improved Intrusion Detection Algorithm Based on GA and SVM", Special Section on Human-Centered Smart Systems and Technologies, Vol. 6, pp. 13624–13631, 2018.

[8] Shang, Wenli; Cui, Junrong; Song, Chunhe; Zhao, Jianming and Zeng, Peng: "Research on Industrial Control Anomaly Detection Based on FCM & SVM", *IEEE International Conference on Trust, Security & Privacy In Computing & Communications*, pp. 218–222, 2018.

[9] Xiaofeng, Zhang and Xiaohong, Hao, "Research on Intrusion Detection Based on Improved Combination of K-means andMulti-level SVM", *17th IEEE International Conference on Communication Technology*, pp. 2042–2045, 2017.

[10] Xie, Miao; Hu, Jiankun and Slay, Jill, "Evaluating host-based anomaly detection systems: Application of the one-class SVM algorithm to ADFA-LD", *11th International Conference on Fuzzy Systems and Knowledge Discovery,* 2014.

[11] Kokila, R.T., Selvi, S.T. and Govindarajan, K., "DDoS detection and analysis in SDN-based environment using support vector machine classifier," *in Proc. 6th International Conference Adv. Computation,* pp. 205–210, 2015.

[12] Pervez, M.S., Farid: "Feature selection & intrusion classification in NSL-KDD CUP 99 dataset employing SVMs," in *Proc. 8th International Conference Software, Knowledge, Inf. Manage. Appl. (SKIMA)*, 2014.

[13] Cover, T. and Hart, P., Nearest neighbour classification*, IEEE Trans. Information Theory*, 13, pp. 21–27, 1967.

[14] Duan, Liyu and Xiaa, Yauan, "An Intrusion Detection Model Based on Fuzzy Cmeans Algorithm", *8th International Conference on Electronics Information and Emergency Communication, IEEE,* 2018.

[15] Dada, E.G., Faculty of Engineering, University of Maiduguri "A Hybridized SVM-kNN-*pdAPSO* Approach to Intrusion Detection System" University of Maiduguri Faculty of Engineering Seminar Series, Seminar Series, Vol. 8, 2017.

[16] Vishwakarma, S., Sharma, V. and Tiwari, A., "An intrusion detection system using KNN-ACO algorithm'', *International Journal of Computer Appl.*, Vol. 171, No. 10, pp. 18–23, 2017.

[17] Shapoorifard, Hossein and Shamsinejad, Pirooz, "Intrusion Detection using a Novel Hybrid Method Incorporating an Improved KNN", *International Journal of Computer Appl.,* Vol. 173, No. 1, pp. 0975–8887, 2017.

[18] Jing, Xueyan, Bi, Yingtao and Deng, Hai, "An Innovative Two-Stage Fuzzy kNN-DSTClassifier for Unknown Intrusion Detection", *The International Arab Journal of Information Technology,* Vol. 13, No. 4, pp. 359–366, 2016.

[19] Sun, Chong; Lv, Kun; Hu, Changzhen and Xie, Hui, "A double-layer detection and classification approach for network attacks", *27th International Conference on Computer Communication and Networks, IEEE*, 2018.

[20] Chen, Feng; Ye, Zhiwei; Wang, Chunzhi; Yan, Lingyu and Wang, Ruoxi, "A Feature Selection Approach for Network Intrusion Detection Based on

Tree-Seed Algorithm and K-Nearest Neighbor", *4th IEEE International Symposium on Wireless Systems within the International Conferences on Intelligent Data Acquisition and Advanced Computing System,* 2018.

[21] Sharifi, Aboosaleh M., Amirgholipour, Saeed K. and Pourebrahimi, Alireza, "Intrusion Detection Based on Joint of K-Means and KNN", *Journal of Convergence Information Technology (JCIT),* Vol. 10, No. 5, September 2015.

[22] Subba, Basant; Biswas, Santosh and Karmakar, Sushanta, "Intrusion Detection Systems using Linear Discriminant Analysis and Logistic Regression", *IEEE INDICON,* 2015.

[23] Zyad, Elkhadir; Khalid, Chougdali and Mohammed, Benattou, "An effective network intrusion detection based on truncated mean LDA", *3rd International Conference on Electrical and Information Technologies ICEIT*, 2017.

[24] Zyad, Elkhadir; Khalid, Chougdali and Mohammed, Benattou, "Combination of R1-PCA and median LDA for anomaly network detection", *3rd International Conference on Electrical and Information Technologies ICEIT,* 2017.

[25] Aburomman, Abdulla Amin and Reaz, Mamun Bin Ibne, "Ensemble of binary SVM classifiers based on PCA and LDA feature extraction for intrusion detection", *International Conference on Advances in Electrical, Electronic and System Engineering,* pp. 636–640, 2016.

[26] Hadri, Amal; Chougdali, Khalid and Touahni, Raja, "Combination of OMPCA and LDA for anomaly network detection", *International Symposium on Advanced Electrical and Communication Technologies (ISAECT),* 2018.

[27] Ibrahimi, Khalil and Ouaddane, Mostafa, "Management of Intrusion Detection Systems based-KDD99: Analysis with LDA and PCA", *International Conference on Wireless Networks and Mobile Communications (WINCOM), IEEE,* 2017.

[28] Alaoui-AdibSaad, Chougdali Khalid and Mohamed, Jedra, "Network Intrusion Detection System based on Direct LDA", *Third World Conference on Complex Systems, IEEE* 2015.

[29] Abdullah, Azween and Zheng, Cai Long, "Improving Intrusion Detection using Genetic Linear Discriminant Analysis", *International Journal of Intelligent Systems and Applications in Engineering, IJISAE*, 2015, Vol. 31, pp. 34–39.

[30] Goeschel, Kathleen, "Reducing false positives in intrusion detection system using data mining techniques utilizing SVM, Decision Trees and NAÏVE Bayes for offline analysis", *SoutheastCon, IEEE*, 2016.

[31] Kabir, Md Reazul; Onik, Abdur Rahman and Samad, Tanvir, "A Network Intrusion Detection Framework based on Bayesian Network using Wrapper Approach", *International Journal of Computer Appl.,* pp. 0975–8887, Vol. 166, 2017.

[32] Jabbar, M.A., Rajanikanthaluvalu S. Sai Satyanarayana Reddy, "Intrusion Detection System using Bayesian Network and Feature Subset Selection", IEEE *International Conference on Computational Intelligence and Computing Research 2017*.

[33] Almansob, Saqr Mohammed and Lomte, Santosh Shivajirao, "Addressing Challenges for Intrusion Detection System using Naive Bayes and PCA Algorithm", *2nd International Conference for Convergence in Technology (I2CT) IEEE 2017*.

[34] Yuan, Yali, LiuweiHuo and Hogrefe, Dieter, "Two Layers Multi-class Detection Method for Network Intrusion Detection System", *IEEE Symposium on Computers and Communications (ISCC),* 2017.

[35] Varuna, S. and Natesan, P., "An Integration of K-Means Clustering and Naïve Bayes Classifier for Intrusion Detection", *3rd International Conference on Signal Processing, Communication and Networking (ICSCN),* 2015 IEEE.

[36] Xiaoyan, HAN; Liancheng, XU; Min, REN and Weiping, GU1, "A Naive Bayesian network intrusion detection algorithm based on Principal Component Analysis"*, 7th International Conference on Information Technology in Medicine and Education,* pp. 325– 328, 2015.

[37] Sonawane, Harshal A. and Pattewar, Tareek M., "Neural Network based Intrusion Detection using Bayesian with PCA and KPCA Feature Extraction", *IEEE International Conference on Computer Graphics, Vision and Information Security (CGVIS),* pp. 83–88, 2015.

[38] FatmaGumus, C. OkanSakar, ZekiErdem and OlcayKursun, "Online Naive Bayes Classification for Network Intrusion Detection", *IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM),* 2014, pp. 670–674.

[39] "KDD-97: A Preview", The Third International Conference on Knowledge Discovery and Data Mining, Newport Beach, California, Sponsored by the American Association for Artificial Intelligence, 1997.

In-depth Analysis of the Performance of RSA and ECC in Digital Signature Application

Arokia Ramya Terrance[1], Shalini Sharma[2], Anu Sharma[3], Kanika Tyagi[4], Nikita Shokeen[5] and Nitya Saini[6]

ABSTRACT: *After passing the Information Technology Act 2000, the digitally signed documents become legally valid in India. The digital signature is used to digitally sign the electronic documents. It is used in many applications like eTender, e-filing the Income-tax returns, etc. Around 20 million users are benefited by the eMudhra Digital Signature Certificate Company in India. The online transactions and e-authentication have witnessed tremendous growth in the digital world, hence cyber security has become a critical element for safe and secure operations. While transferring data files, it becomes ineluctable not only to authenticate the users but also to make sure that the original data has not tampered. The Digital signatures are used to uncover unauthorized alteration to data and provide cyber security to authenticate the identity. The Digital signature application makes use of cryptographic techniques in order to provide safe and secure information to the users. The paper presents an in-depth analysis of two cryptographic techniques namely RSA and ECC in the Digital Signature applications and compares the performance and security aspects of these techniques.*

Keywords: Cyber Security, Digital Signature, Cryptography, ECC, RSA.

1. INTRODUCTION

The digital signature is an electronic signature used to authenticate the digital file. The digital signature is necessary for safe and secure

[1,2] Assistant Professor, Department of Computer Science, Kalindi College, University of Delhi, India. E-mail: [1]arokia.ramyat@gmail.com; [2]Shalinisharma@kalindi.du.ac.in

[3–6] Student, Department of Computer Science, Kalindi College, University of Delhi, India. E-mail: [3]anu08111999@gmail.com; [4]tyagikanika99@gmail.com [5]nikitashokeen1@gmail.com; [6]nitya41099@gmail.com

communication in the digital world. It gives the recipient a very strong reason to believe that the message was sent by the claimed user and that message was not altered in transit. Similar to the hand-written signature perform the role in the traditional system to construct the authentication of the paper document, the digital signature performs the role of authenticating the digital document. The four main characteristics of the digital signature are as follows:

- ***Integrity:*** Integrity means that a digital signature ensures the originality of data sent that is the message or sent data is not modified or tampered while transferring.
- ***Confidentiality:*** Confidentiality means that digital signature confirms the privacy of the data sent, that is it attempts to make the content of the message secret.
- ***Authentication:*** Digital signature ensures proof of identity. It identifies the claimed sender and the receiver of the message correctly.
- ***Non-repudiation:*** Digital signature confirms that neither the sender nor the receiver can oppose to the fact that the message exists.

The Digital signatures are constructed with the help of public-key/ asymmetric cryptography. Asymmetric cryptography makes use of two distinct keys, named as the private key and public key, for the process of creation of the digital signature and verification of the digital signature. As shown in Figure 1, the sender's private key is used to generate a digital signature at the sender side, whereas his/her public key is used for digital signature verification at the receiver side.

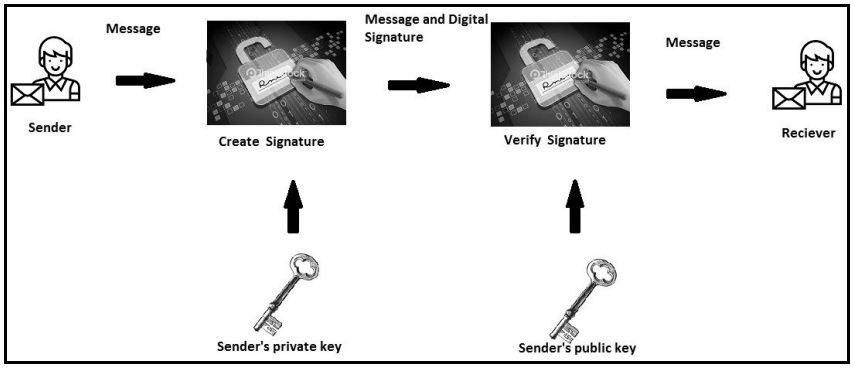


Fig. 1: Digital Signature

The applications of the digital signature include sending secure email, web-based monetary transactions, and code signing of any software developed

by any developer, e-Filing and e-tender. The digital signature is generated uniquely for every signer that is, no two different signers can have the same signature code. In India, the Digital Signature Certificate (DSC) authorized by the Controller of Certifying Agencies (CCA) helps different organizations or firms to have a smoother transaction of the Digital Signature application related to the Registrar of companies or Ministry of Corporate Affairs (MCA).

2. DIGITAL SIGNATURE PROCESS

The digital signature process is divided into the following three phases.

2.1 Key Generation

The process of key generation practices to produce the keys that are to be used in signing and verification process of digital signature. Whenever it executes, algorithm produces a key pair comprise of private (signing) key and an analogous public (verification) key. This process comprises of a random number generator and that will generate a unique key pair whenever it is used. SK stands for secret key because in the procedure the private key (signing key) is secret and PK stands for public key as the key is distributed to all users willing to verify the signature [1].

2.2 Digital Signature Creation

The creation of the Digital Signature involves the generation of the hash value for the given message, encrypting it, and transmitting it to the receiver, along with the original message and the process flow diagram is shown in Figure 2.

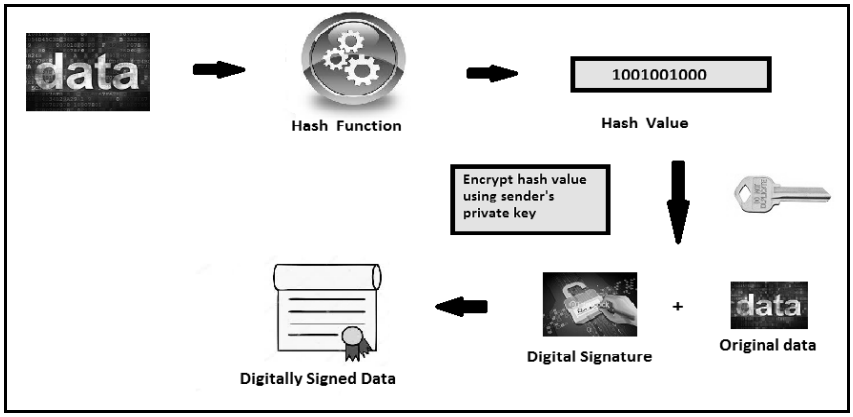


Fig. 2: Process Flow Diagram for the Creation of Digital Signature

Step 1: *Generating the Hash value*—The hash value of the original message is created, using one of the hash algorithms namely MD5, SHA-1, SHA-512.

Step 2: *Computing digital signature*—The created hash value should now be encrypted by the sender's private key, using any mathematical cryptographic encryption algorithm. The encrypted hash value is called as the digital signature

Step 3: *Transferring digital signature and original message*—The digital signature is obtained in the second step and the original message are transported to the recipient, who wants to verify the digital signature.

The above steps are illustrated in Figure 3, the Finite State Machine for signing the document by the sender.

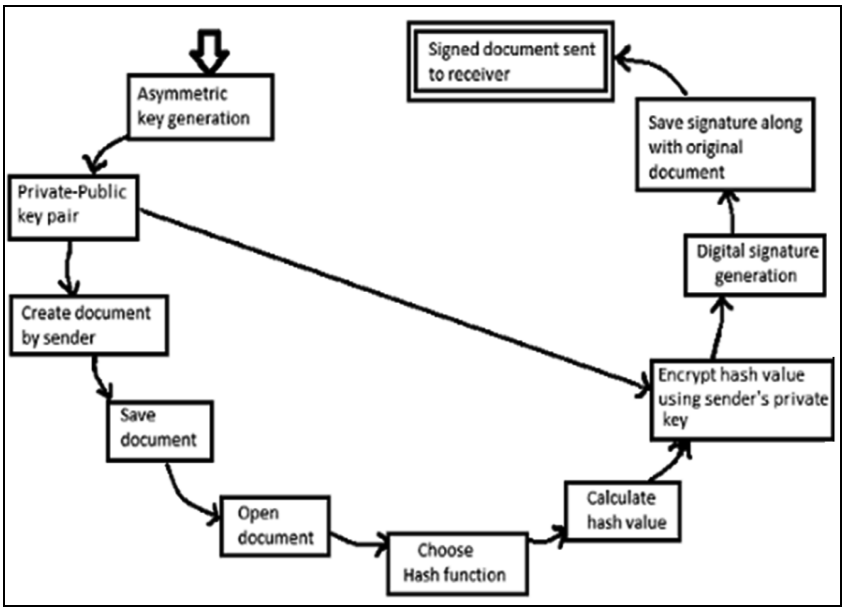


Fig. 3: Finite State Machine for Signing the Document by the Sender

2.3 Digital Signature Verification

The verification of the Digital Signature involves the computation of the hash value of the message and comparing it with the hash value obtained by decrypting the received digital signature and the process flow is shown in Figure 4.

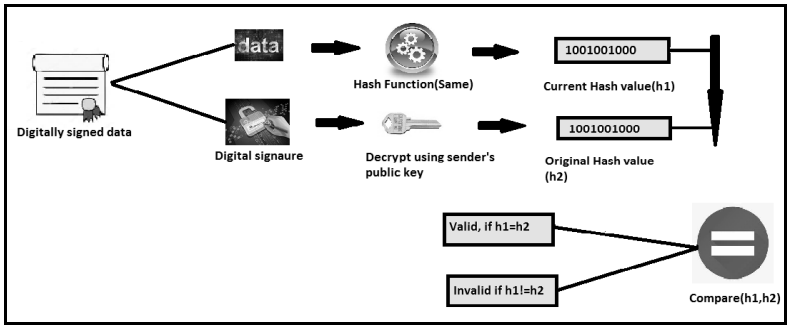


Fig. 4: Process Flow of the Verification of Digital Signature

Step 1: *Computing current hash value*—A Hash value for the message received by the recipient is computed using the same hashing algorithm used during the signing procedure. Let the resultant hash value be called as h1.

Step 2: *Obtaining hash value*—The digital signature received by the recipient is decrypted, using the same decryption algorithm corresponding to that which was used during the signing procedure. The decryption is done by the public key of the sender. Let the resultant hash value be called as h2.

Step 3: *Comparing hash values*—The hash value (h1) computed in the first step is compared with the hash value (h2) computed in the second step. If the values h1 and h2 are the same, then the verification process is fruitful and this proves that digital signature is valid. If the value of h1 and h2 is different, then the digital signature is invalid.

The above steps are illustrated in Figure 5, the Finite State Machine for verifying the document by the receiver.

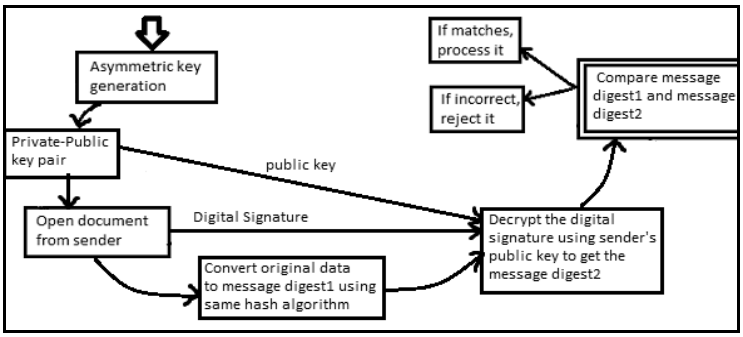


Fig. 5: Finite State Machine for Verifying the Document by a Receiver

The public key cryptosystem used to produce digital signatures can cause a sizeable difference in the performance of the digital signature process. The two most widely used algorithms in the digital signature are RSA and ECC. A brief analysis of these algorithms is done in this paper.

3. RSA

RSA is named after its devisers, Ron Rivest, Adi Shamir and Leonard Adleman [6]. RSA is a public-key cryptography that uses asymmetric key algorithms, where the key meant for encrypting a message is different from the key used for decrypting a message. RSA is the first algorithm that can be used both for encrypting the data and for generating digital signatures. RSA is most commonly used to establish truthfulness of data [7]. In the algorithm, two large prime numbers are used for creating the public key and the private key. RSA uses numbers e, d and n where (d, n) is known as the private key that must be kept secret and (e, n) is known as the public key that can be shared [3]. The number e is known as the public exponent, the number d is known as the private exponent, and n is known as the modulus. The key length of RSA is equal to the length of the modulus.

Key Generation

1. Choose two prime numbers p and q, where p and q should be distinct.
2. $n = pq$.
3. Let $\varphi(pq) = (p - 1)(q - 1)$ where φ is totient function.
4. Pick an integer e such that $1 < e < \varphi(pq)$, where e is an encryption key and $\gcd(e, \varphi(pq))=1$.
5. Calculate decryption key d such that, $d*e \bmod \varphi(pq) = 1$, where d is a secret key exponent.

Signature Creation

Let us consider the message being signed, m is in the range (1, 2, 3, ..., n–1). The Signature s is computed by encrypting the value of the hash function of the message h(m) where h is the hash function and m is the message sent from the sender end using the private key e.

$$s = h(m)^e \bmod n$$

Signature Verification

Signature is verified with the help of public-key d.

1. $h' = s^d \bmod n$, is computed.

2. If h' = h(m') mod n, then the signature is valid, and invalid otherwise where h is the hash function and m' is the message received at the receiver.

4. ECC

Elliptic Curve Cryptography, discovered by Miller and Kobiltz [5] is a cryptographic technique that makes use of public keys on elliptic curves over finite fields. ECC has been around since 1985. This type of system is most suitable for memory constraint devices such as Palmtop, Smartphone, Smartcards, [8]. For cryptographic use, the curve is considered over finite fields, not over real numbers. Here, we use an elliptic curve over prime fields, Z_p. An elliptic curve for all points (x, y) over Z_p is defined as follows:

$$y^2 = x^3 + ax + b \bmod p,$$

Where $p > 3$ and $a, b \in Z_p$, and the condition $4a^3 + 27b^2 \neq 0 \bmod p$ is satisfied.

The basic idea behind that the addition of two points P and Q, on an elliptic curve result in a third point, R on the curve, that is, P + Q = R. The addition is well defined only if there exists an identity element O, such that P + O = P. The inverse -P is defined as P + (-P) = O.

4.1 Elliptic Curve Digital Signature Algorithm

ECDSA is a revised version of Digital Signature Algorithm that uses elliptic curve cryptography. ECDSA was accepted as an American National Standards Institute (ANSI) in 1998 [4]. Here, we define our ECDSA standard for curves over prime fields Z_p.

Key Generation

1. Let us consider an elliptic curve E, of the general form described above, on which is a point A that generates a cyclic group whose order is q.
2. Choose an arbitrary integer t, and compute B = tA.
3. We now have the keys where,
   (a) k1 = (p, a, b, q, A, B), is the public key, and
   (b) k2 = (t), is the private key.

Signature Generation

The signature consists of a pair of integers (r, s). Both r and s are of the same bit length as q. Signature is generated as follows:

1. Choose an integer k_e with $0 < k_e < q$. Here, k_e is a key lasting for a short period of time.
2. Calculate $R = k_eA$.
3. Let r = x-coordinate of R.
4. Compute the value of the hash function of the message h(m) where h is the hash function and m is the message.
5. Compute signature, $s \equiv (h(m) + t.r)\ ke^{-1} \bmod q$,

Signature Verification

1. Compute value $v \equiv s \bmod q$.
2. Compute value $w_1 \equiv v.h(m) \bmod q$.
3. Compute value $w_2 \equiv v.r \bmod q$.
4. Compute $P = w_1A + w_2B$.
5. If $X_p = r \bmod q$, where X_p is the x-coordinate of point P, then the signature is valid and invalid otherwise.

5. COMPARISON BETWEEN ECDSA AND RSA

In this paper, we have implemented a digital signature application using a cryptographic technique named as RSA and ECC. We have executed the digital signature application on the original message of word count 268 using RSA with sample data inputs of 512, 768, 1024 and 2048 bits and ECDSA with sample data inputs of 192, 239 and 256 bits in CrypTool respectively and shown its impact in all the major phases namely Key Generation, Creation of the Signature and Verification of the Signature of the Digital Signature application. The hash function used is SHA-1 hash function. All the computations are performed using CrypTool 1.4.41 on a Microsoft Windows platform. CrypTool is an open-source project. The bit sizes are chosen as per its usage of these cryptographic techniques in most of the real-world devices.

5.1 Key Generation Phase

Table 1: Execution Time of Key Generation under Different Scenario

RSA		*ECC*	
RSA Modulus (Bits)	*Execution Time (Seconds)*	*Curve Parameter (Bits)*	*Execution Time (Seconds)*
512	0.1	192	0.027
768	0.13	239	0.018
1024	3.314	256	0.02
2048	4.086	Nil	Nil

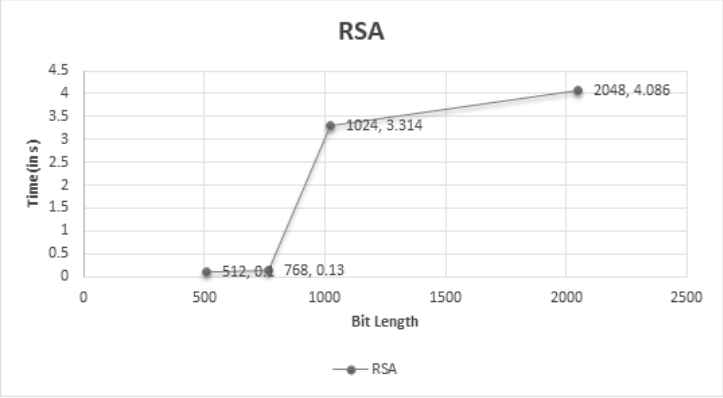


Fig. 6: Graph of the Execution Time for the RSA Key Generation Step for Distinct RSA Modulus

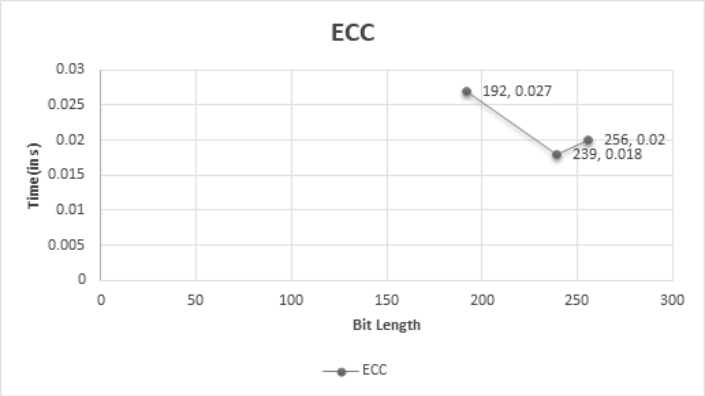


Fig. 7: Graph of the Execution Time for the ECC Key Generation Step for Distinct Curve Parameters

5.2 Signature Generation Phase

Table 2: Execution Time of Signature Generation under Different Scenario

RSA		*ECC*	
RSA Modulus (Bits)	*Execution Time (Seconds)*	*Curve Parameter (Bits)*	*Execution Time (Seconds)*
512	0	192	0.028
768	0.002	239	0.002
1024	0.004	256	0.002
2048	0.016	Nil	Nil

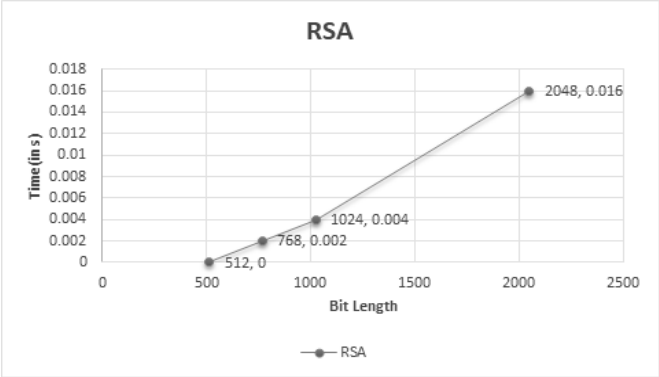


Fig. 8: Graph of the Execution Time of the RSA Signature Generation Phase for Different RSA Modulus

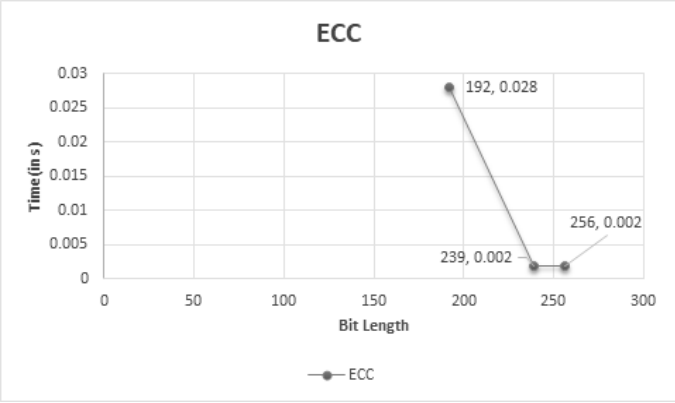


Fig. 9: Graph of the Execution Time of the ECC Signature Generation Phase for Different Curve Parameters

5.3 Signature Verification

Table 3: Execution Time of Signature Verification under Different Scenario

RSA		*ECC*	
RSA Modulus (Bits)	*Execution Time (Seconds)*	*Curve Parameter (Bits)*	*Execution Time (Seconds)*
512	0	192	0.002
768	0	239	0.006
1024	0	256	0.006
2048	0	Nil	Nil

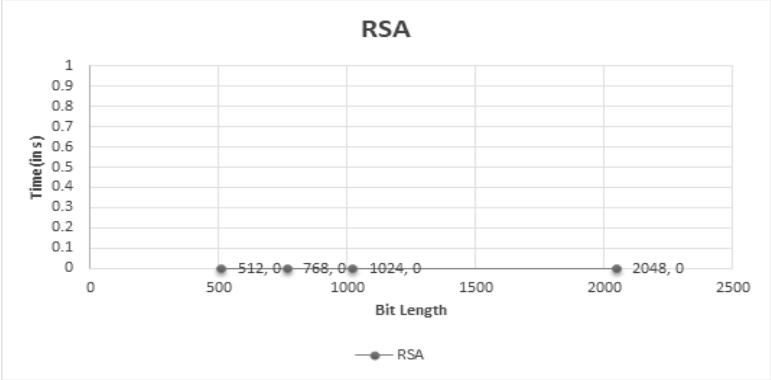


Fig. 10: Graph of the Execution Time of the RSA Signature Verification Phase for Different RSA Modulus

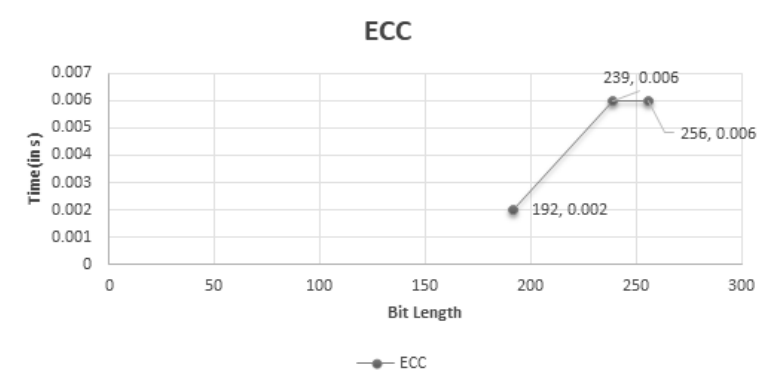


Fig. 11: Graph of the Execution Time of the ECC Signature Verification Phase for Different Curve Parameters

The above results show that the step for key generation of ECDSA is a lot swifter than that for RSA, whereas signature generation takes similar time and signature verification is faster for RSA, but not to a larger degree. Also, it is well-founded that Elliptic Curve point of a multiplication operation is found to be computationally more efficient than RSA exponentiation. [9]

Table 4: Comparison between ECDSA and RSA

Comparison Parameter	*RSA*	*ECDSA*
Encryption Time	Increases to a larger degree for all key sizes	Takes very less time to encrypt even very large keys
Decryption Time	Takes a lot more time than ECC to decrypt large messages	Takes more time than RSA to decrypt small messages but very less time to decrypt large messages
Key Generation Time	Slower, but quite comparable for smaller key sizes (e.g., 128 bits)	Outperforms RSA and is more efficient for bigger key sizes (e.g., 512 bits, 1024 bits, etc.)
Signature Generation Time	Slightly slower for smaller key sizes	Outperforms RSA and is evident for key sizes as large as 1024 bits
Signature Verification Time	Outperforms ECC even for very large key sizes	Takes more time than RSA

6. CONCLUSION

RSA cryptographic technique is comparatively easier to understand and implement than ECC. The main advantage of RSA over ECC, established from the research is that the operation like signature verification is faster in RSA, because of the smaller length of its public keys. In most of the practical scenarios, a message is generally signed only once and gets verified more than one time. It is a well-established fact that the computational power of the devices shows significant improvement with time. So, it will become relatively easier to detect the private key in the RSA signature scheme, as it requires factorizing the key into two primes. ECC, on the other hand, can be abused or we can say attacked if someone knows the ephemeral key, which is a slight advantage over RSA as it lasts

for a very short period of time. It is clear from the experiments conducted that ECC is able to provide almost a similar security level for digital signature operations as RSA, but for smaller bit lengths. In conclusion, ECC is an extremely appealing algorithm to be used in memory-constrained devices such as smartphones, tablets, etc.

REFERENCES

[1] Ali, A.I. (2015, June). Comparison and Evaluation of Digital Signature Schemes Employed in Ndn Network. *International Journal of Embedded Systems and Applications*, 5(2), 15–29.

[2] Srivastava, S., Trehan, V. and Yadav, P. (2012). Digital Signature. *International Journal of Engineering and Management Sciences*, 3(2), 115–118.

[3] Preetha, M. and Nithya, M. (2013). A Study and Performance Analysis of RSA Algorithm. *Journal of Computer Science and Information Technology*, 2(6), 126–139.

[4] Christof Paar and Jan Pelzl. Understanding Cryptography (2010). *Springer Nature.*

[5] Mahto, Dindayal; Khan, Danish Ali and Yadav, Dilip Kumar (2016). Security Analysis of Elliptic Curve Cryptography and RSA. *Proceedings of the World Congress on Engineering Vol. I, WCE, June 29 – July 1, 2016.*

[6] Gupta, S., Sinha, R. and Srivastava, H.K. (2013, May). Performance-Based Comparison Study of RSA and Elliptic Curve Cryptography. *International Journal of Scientific & Engineering Research*, 4(5), 720–725.

[7] Bhujade, V. and Chaudhary, D. (2015). Analysis and Implementation of Combined Approach of RSA and ECC Algorithm for Enhanced Data Security. *International Journal of Computer Applications*, 129(15), 37–42.

[8] Mahto, D. and Yadav, D.K. (2017). RSA and ECC: A Comparative Analysis. *International Journal of Applied Engineering Research*, 12(19), 9053–9061.

[9] Prabu, M. and Shanmugalakshmi, R. (2009, June). Research Issues on Elliptic Curve Cryptography and Its applications. *International Journal of Computer Science and Network Security*, 9(6), 19–22.

[10] Alese, B.K., Falaki, S.O. and Philemon, E.D. (2012). Comparative analysis of public-key encryption schemes. *International Journal of Engineering and Technology*, 2(9), 1552–1568.

Social Media—A Platform for Political Discourse: An Evaluation

Abhishek K. Singh[1]

ABSTRACT: *With the advent of media technologies, social media has gained eminence. Social media is becoming a very significant medium that communicate information through multimedia. Todays, social media is very powerful in disseminating messages and have become a very crucial part of political campaigns. The present study thus, evaluate how social media used to communicate and manage political discourse. In this study, researcher deals with the relation between politics and modern media tools that promotes the interaction between the politically influenced public and alternative sources of information. The focus of the research is around the issues of social networks which act both as a vehicle for social networking as well as a platform for political debate by incorporating various cases that suggests a polarization in the sources of information indicating the role of social media as an agent for polarization of opinion in social discourse in special contest to Indian youth. The epicenter of the study deals with the role and impact of social media on Indian students. Parameters such as political mobilization, political influence, opinion, changing behavior in youth are used to evaluate the role of social media. The methodology used in the research is mainly conducting (most of the data are collected using a questionnaire as a research tool)*

Keywords: Social Networking, Facebook, Political Discourse, New Media, Agenda Setting.

1. INTRODUCTION

The beginning of new media technologies has given a new platform for socio-political communication, including social media where common people can act react and interact with open, courageous and bold comments. Consequently, use of social media is very crucial as a public

[1]Academic Expert and Media Researcher, Ex-Faculty, Ministry of Higher Education, Sultanate of Oman. E-mail: asingh8319@gmail.com

sphere. The social media is not only a boon for social awareness, but it gives new astuteness to the political discourse as political mechanism. It is generally accepted that social media play a central role in information diffusion. As a result, many socio-political communities are taking it as the digital platform for political discussions (Chadwick, 2013). It provides important communication channel through which the political community or citizens, as ordinary contributor, can share information about its activities, publish its position on various topics, allocate information from multiple sources, as well as report on the issues surrounding them at a given time.

Social media is really a big demand because it gives new meanings to communication and information flow which greatly influences the way people think. Graham (1999) notes that present media technology certainly has an ability to create and support sociopolitical activities in the society as it directly influence the identity of individuals and shape their mindset towards specific outlines. Social media like Facebook, Twitter, YouTube and blogs have appeared, changing the feature and mechanism of modern communication. Social media enhances the link between social structure and public interactions for better peoples understanding towards the decision making process.

Media researchers have speculated that the use of social media influences citizen's political behavior (Gil de Zuniga, *et al.,* 2012). Many studies have shown that digital stands like social media have a deep penetration for social reform and if it is used suitably in an organized way, it has an ability to change public orientation. The social media are not the only source of information, identity and social relations, but it also used in the generation of a certain ideology and motivated people for political participation (Katz and Gurevich, 1974). As a result, social media is a very powerful tool that will spread, motivate, speed up information and give new meaning to political expressions.

"Political commitment and social media" has now become very smart and innovative approach in the field of contemporary social research. Specifically, in terms of Indian society, the social media is very popular and have a high range of reachability in its domain. In this research, the researcher has followed the social track to monitor the reactions related to political activity on social media. Social media are certainly contained very influential messages and have high impact on the sentiments of common users. In this paper, researcher has investigated the social media and its various provocative posts to understand the following questions :-

1. In what ways the young people are getting political information for political discourse on social media.
2. How do the general people show political sentiment in social media?
3. What forms of messages commonly used for political action in social media?
4. How do youngsters start involving themselves as a part of political discourse using political posts on social media.

2. LITERATURE REVIEW

Political discourses are mainly characterized by the different levels of issues and propagation of political, ideological based images and comments through digital media for communication and coordination. Social media like Facebook normally supports self expression and encourage social relations. Social media has a deep impact on the emotion of common users and it affects the diffusion of emotion. The social media accelerate the broader horizon of sentiments. Therefore, it can be pointed out here that SNS environment has wide converge than the traditional media networks (Brady, Wills, Jost, Tucker and van Bavel, 2017). Growing media use has been become a center of attraction of many researchers, especially, for political campaigns. Such as how social media is being used a tool for political agenda framing and how social media is being used to facilitate social sentiments politically (Zhang & Lee, 2018).

The Internet is growing and expanding like mushroom both vertically and horizontally due to the cheapening of technology. Especially among young people, internet is as essential as education. One research indicates that use of Twitter, Facebook, YouTube, Myspace is increasingly very common among younger age groups within the Arab world, especially the 20 to 30 year old age group, which uses the net more avidly compared to the rest of the population (Abdulla, 2006: p. 80).

Arab spring was the great example of social media used to unite social sentiments. In this respect, Egypt, the 90% (approx) population were active users of social media and very active in the political movements. In fact social media was the real driving force behind the Egypial revolution (Sahar Khamis & Katherine Vaughan, 2011). Some more studies say, new media revolution erupted in the Arab world after 1990, inspired by the introduction of both satellite television channels and the Internet (Khamis & Sisler, 2010), this situation is rapidly changing, since many Arab countries are currently striving to increase Internet penetration rates.

However, research is very limited in the field of social media as a platform to create emotional impact on political operation. Today's journalistic practices are used to create powerful and interesting news. Looking for something sensational, surprising or shocking to catch the public's attention is always been a motive of today's news makers to influence the public sentiments. Previous research has shown that the season of political operation strongly influences the quantity, style and (political) news about the actors (van Aelst & Swert, 2009). Different media deal with the same news in different ways. Moderators and serious newspapers and television news programs have most often used frameworks of responsibility and conflict in the presentation of news, while sensationalist media have most often used the framework of human interest (Semetko and Valkenburg, 2006).

3. OBJECTIVE OF THE STUDY

1. To analyze the social media as a catalyst to activate sentiments for political discourse.
2. To evaluate the usefulness of social media for news consumption, opinion, expression and provide a virtual space for political discourse.

As mentioned above, the research is analytical in nature as it sharply shows the power of social media in the promotion of political activity and how social media is used by political associations to motivate social elements for protest participation by influencing their online behavior and expressions. Furthermore, the issues highlighted in social media are certainly results, changes in expression and development of new political behavioral pattern in the social capitals.

4. RESEARCH METHODOLOGY

This study used a random sampling method to interpret the collected data. This study also takes into account some secondary data sources. Survey-based research conducted online and as emails. This was used with about 50 university students in Delhi-NCR. Information was also collected mainly by e-mail. The respondents were graduate students.

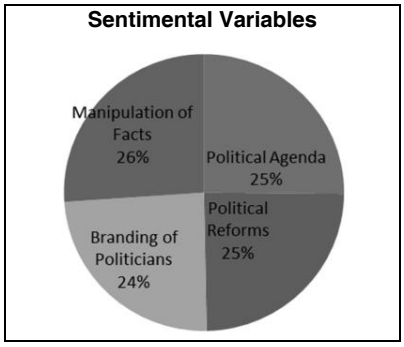
Respondents are mainly academic intellectuals of age group 20 to 30 years old. The data were collected on the basis of questionnaires based on the variables related to social media, political discussion and public opinion or sentiments, which is a main theme of the present research topic. The samples were consulted for their experience on social media and its role as a tool for political discourse.

5. ANALYSIS AND DISCUSSION

In this research the main focus area is the applications of social media and its related variables which play a very important role in the formation of the new culture of political discourse. Presently, online engagements have become a part of social civilization. Social media positively plays a very significant role in the political discourse and undoubtedly, it gives a new direction to the social perception and distress the political opinion. The present research pointed out how the social media affects the style of new communication systems to promote a specific agenda (mainly political). The research analysis was mainly focused on the political criteria of discussion on the social media which may have ability to change the entire political beliefs. The samples were asked questions keeping in mind the fact that social media surely have an powerful role in manipulating the political commitment to the society and it helps largely to polarize the users.

The present analysis was based on the following parameters:

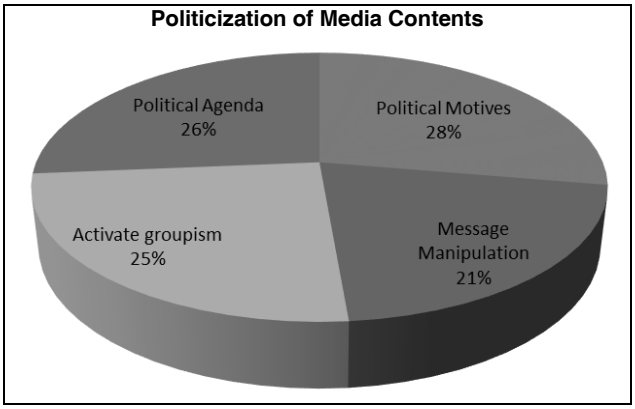
1. ***Rationality and sentimental judgement:*** The respondents pointed out that the social media are very powerful and have intense ability to change the sentiment of users. Social media certainly emphasize the role of impression formation and allow political judgement. The rationality and its related factors play a very crucial role in the formation of agenda based mental image in the mind of users.



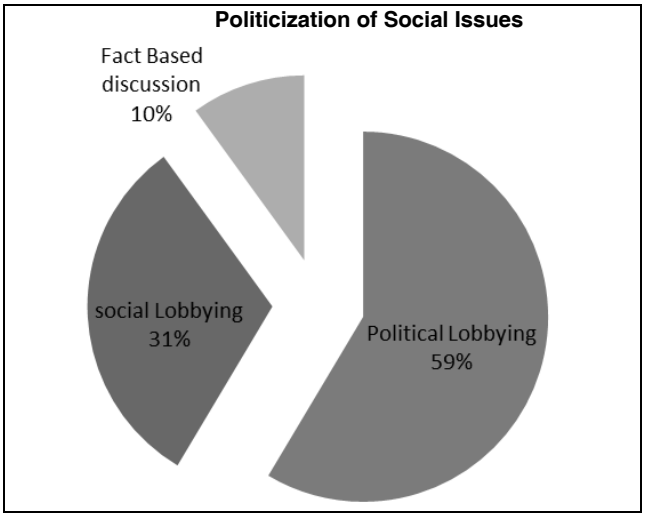
The present analysis indicated that the sentiment and the judgement for the comment shows the dialog for the news and updates are closely related to the emotional representation of online users. The user comments act as a catalyst, which shows the formation of the specific identity of participants. Thus, it may result, online activism

for political discourse. Here, it is also noteworthy that such online activism is very fundamentally growing as a political campaign as its power for sentimental penetration is very deep and highly influential in present social-technical public sphere.

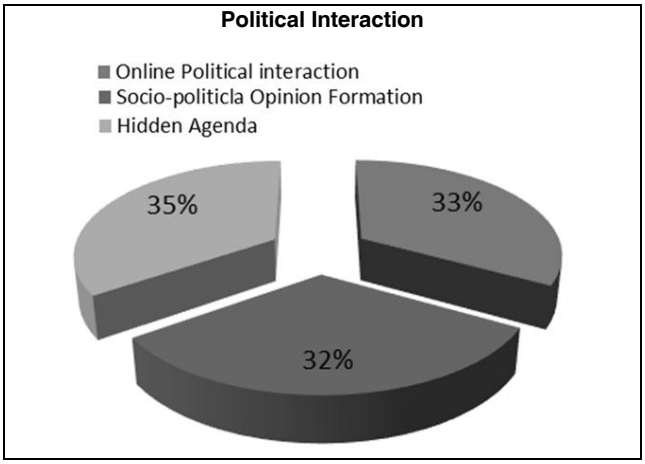
2. ***Social Media as a political Public Sphere for political activism:*** The public sphere is a space of communication where common people discuss and analyze the information. Certainly, it creates a kind of public opinion and gives an issue based discourse on the social platform. Keeping in mind, this study mainly focuses on the three dimensions which usually conceptualize the online discussion constructively. These dimensions are:
   - ***Democratic rights, i.e. freedom of expression are used for Political Mobilization:*** The concepts of freedom of expression are used for the manipulation and diversion of the message. Thus, the response from the people are mentioned as:



   - ***Examine politicization of social issues for political lobbying:*** The general public have generally very limited reach of the content of political affairs. It's been noticed that the general people are using social media as a main source of information as its very easy and attractive medium where people easily participate and discuss with free space and opinion. It can be concluded here, the social issues are very much glamorized and politicized for political lobbying on social media. Political parties, activists and common groups are using social media as a political tool to reach maximum people of the society and to increase their members.



- ***Social understanding and interactive representation of specific ideology***: Social media as public spheres are mostly used for the social unification for presentation and representation of specific ideology. As per the responses for the respondents. The specific ideaology are most of the time have some hidden agenda and theme, which is usually political in nature. These kind of messages and compose in expression is used to spread some specific opinion and create a group of similar voice on online media. The responses are mentioned below:



This analysis based on all the considered factors pointed out the political utility of social networking sites, which indicate that social media such as Facebook, YouTube and Twitter is a legitimate place to discuss political issues and share political ideology. Under investigation in a conversation consists of participants with opposing views. Similarly, the research has apparently succeeded in overcoming the polarization of online discussion, which has penetrated online political debate in the social community. However, the representation of views within the various social media group was strongly biased in favor of discussions of like-minded participants with around 70% of all social media users. This shows that the unique construction of social networks, social networking sites can play a small role in facilitating the exposure of participants to different parties and the heterogeneity of online political discussions.

6. CONCLUSION

The results of this study are considerably pointed out the capacity of social media and social networking, which enable people from the dissimilar standpoint to incorporate and contribute in political discourse. In addition, people seek entertainment and use various platforms of social media and develop the capabilities of technologies to engage in political discussions and express their views on issues of interest to them. Even if the online political discussion existed long before the advent of social networks, the richness and the popularity of these technologies, and the reduction in the cost of Internet access, it may be that more people and new participants are attempting online political discourses. Besides, as this study pointed out, these sites can promote civility between different social groups in discussions and interactions. In these circumstances, we can conclude that growing citizens are dominantly using social media as a political tool. Social media regularly adding new creative users, who is encouraging such discussions at higher levels to reach maximum contributors. The study clearly stated all the research questions mentioned above. The social media, mainly points the youth and through various platform social media tried to reach and influence the youth for political discourse. Secondly, the social media, especially allows people to share political sentiments and get attached to the specific group as member for political discourse. The study startlingly stated that the social media is a kind of convenient media as it available in all possible languages with interactive text, picture, video and suitable use of multimedia, these new media features creates visual customization effect to attract and interact common users.

This study also gives rise to some optimism in the online political discourse. Social media, in the present era of information technology, certainly have a dominantly role in opinion formation by penetrating the social sentiments and criteria of political judgment. Social media is a major public sphere, where ordinary people gather to politely and rationally discuss perspectives and arrive at the best possible conclusion. Lastly, in my view point, social media were developed with the aim to connect and understand our society in a much better way, but how people are using it, is totally different as society is getting divided into different groups based on the specific political frame but isolated on humanitarian ground.

REFERENCES

[1] Abida, E. (2013). Impact of New Media on Dynamics of Pakistan Politics. *Journal of Political Studies,* 21(1), 113–130.

[2] Aday, S., Farrel, H., Freelon, D., Lynch, M., Sides, J. and Dewar, D. (2013). Watching From Afar: Media Consumption Patterns around the Arab Spring American Behavioral Scientist, 57(7), 899–919.

[3] Barnes, S.H. and Kaase, M. (Eds.). (1979). Political action: Mass participation in five Western democracies. Beverly Hills, CA: Sage.

[4] Bekkers, V., Beunders, H., Edwards, A. and Moody, R. (2011). New media, micromobilization, and political agenda setting: Crossover effects in political mobilization and media usage. Information Society, 27, 209–219. doi:10.1080/01972243.2011.583812.

[5] Bennett, W.L. and Segerberg, A. (2011). Digital media and the personalization of collective action: Social technology and the organization of protests against the global economic crisis. Information, Communication and Society, 14, 770–799. doi:10.1080/1369118X.2011.579141.

[6] Ender Yunus, The Role of Social Media in Creating Political Awareness and Mobilizing Political Protests; Royal Institute of Technology School of Computer Science and Communication, Sweden, 2013.

[7] Qasim Mahmood, Amira Sariyati Binti Firduas and Hamedi Bin Mohd Adnan, PUTAJ – Humanities and Social Sciences; Social Media and Protest Behavior, Vol. 23, No. 2, 2016.

[8] Sebastián Valenzuela, Unpacking the Use of Social Media for Protest Behavior: The Roles of Information, Opinion Expression, and Activism, American Behavioral Scientist, SAGE Publications, 2013; DOI: 10.1177/0002764213479375.

[9] Shah, D.V. and Kwak, N. Information and expression in a digital age modeling Internet effects on civic participation. Communication Research, pp-. 531–565, 2005 doi:10.1177/0093650205279209.

[10] Dalton, R.J., Sickle, A.V. and Weldon, S. The individual-institutional nexus of protest behaviour. *British Journal of Political Science,* 40, 51–73, 2009, doi:10.1017/S000712340999038X.

[11] Lovejoy and Saxton, 2012 Lovejoy, K. and Saxton, G.D. Information, community, and action: How nonprofit organizations use social media. *Journal of Computer-Mediated Communication,* 2012, pp. 17.

[12] Katz and Gurevitche's, Utilization of mass communication by the individual. In J.G. Blumler and E. Katz (Eds.), The uses of mass communication (pp. 19–32). London, UK: Faber, 1974.

[13] Kroh and Neiss. On the causal nature of the relationship between Internet access and political engagement: Cambridge University Press, 2012.

Feedback Ratings are Better Indicators of User Sentiments than Whole Review

Geeta Arneja[1] and Harmeet Kaur[2]

ABSTRACT: *Sentiment analysis is a field of NLP (Natural Language Processing) that is used to discover opinions from the textual data. These opinions have a great impact on customers to make decisions regarding online shopping and choosing products. With the advancement of digital technologies, a large amount of textual data is available in the form of customer reviews. Sentiment analysis systems are used to classify textual data on the basis of their polarity which can be positive, negative or neutral. In this paper, two different approaches were applied to extract the sentiment of the reviews at the sentence level. In the first approach, TextBlob library was used to calculate the polarity of the sentence whereas the second method follows a rule-based approach in which two Lexicons of polarised words (negative and positive) were used. In each of the approach, the value of sentiment was calculated both for the review and the summary provided in the dataset and ranked from 1-5. By performing statistical analysis on the relationship between reviews, summary and overall rating specified in the dataset, it becomes evident that the summary and ratings provide a much more accurate opinion of the reviewer. Since it is more time consuming to process textual data, ratings can be safely used for analysis of user sentiments*

Keywords: Sentiment Analysis, Rule-based Approach, TextBlob, Polarity, Lexicon.

1. INTRODUCTION

With the advancement in the field of opinion-rich resources like online review sites, it has become possible to use this information for

[1] Guest Faculty, Department of Computer Science, Rajdhani College, University of Delhi, Delhi, India. E-mail: arnejageeta@gmail.com

[2] Department of Computer Science, Hansraj College, University of Delhi, Delhi, India. E-mail: hkaur@hrc.du.ac.in

better understanding of people's opinions. Individuals rely on these reviews to make decisions related to online shopping, choosing events, products and entities. However, it would not be easy for an average human reader to detect relevant sites and mining the sentiment from the huge amount of unstructured text.

Sentiment analysis is the computational approach used to extract opinions, attitudes and assessments towards issues, products, services and events etc. An objective sentiment analysis system is needed that can automatically calculate the sentiment concealed within the unstructured text available on the Web. Sentiment analysis systems classify textual data on the basis of their polarity which can be positive, negative or neutral. It can be done at different levels like document level, sentence level and sub-sentence level. Different methods and algorithms have been devised to implement sentiment analysis. They can be classified as rule-based approaches, that are based on manually defined rules; automatic systems, that rely on machine learning techniques and hybrid systems which combine both of them.

In this paper, the dataset retrieved from Amazon beauty products was analyzed. It contains reviews of the customers about the beauty products sold by Amazon. Two different approaches were applied to extract the sentiment of the reviews at the sentence level:

1. TextBlob library was used to calculate the polarity of the sentence.
2. A rule-based approach in which two Lexicons of polarised words (negative and positive) were used.

In each of the approach, the value of sentiment was calculated both for the review and the summary provided in the dataset and ranked from 1–5. Correlation, a statistical measure, was then used to study the relationship between reviews, summaries and ratings for both the approaches. A formula based on the number of occurrences of positive and negative words, to compute the polarity of a review was devised and applied on the dataset. The polarity was modified if any positive or negative word was preceded either by negations (not, never etc.) or by emphasizers (highly, very etc.) accordingly.

The result sets derived from both the approaches were statistically analyzed against the overall ratings and against each other. Assuming that an average human expresses the precise opinion in summary, the results showed that the sentiments derived from summary data are much closer to the overall rating as compare to that of the whole review. Thus, in

order to make decisions, an individual can rely on feedback ratings instead of going through the whole reviews.

The structure of paper is as follows. Section 2 presents literature review. In section 3, we describe the research objective. Methodology is presented in section 4, which describes the approaches used and section 5 concludes the obtained results. The final section outlines the expansion possibilities in future.

2. LITERATURE SURVEY

Several researches have been made regarding sentiment analysis of data available on web in the form of reviews with focus on categorizing them as positive, negative and neutral. A huge amount of data is available on the web in the form of customers' reviews. Sentiment analysis of online customer reviews is under high demand from the perspective of customers and vendors. In [1] research problems based on the level of granularities have been discussed. In general, sentiment analysis has been investigated mainly at three levels: document level, sentence level and entity and aspect level.

Several researchers have worked on feature extraction and opinion-oriented words [2]. The task is performed by identifying the features of the product that customers have expressed opinions on (called opinion features) and ranking the features according to their frequencies that they appear in the reviews. Then, for each feature, the number of positive and negative customer reviews have been identified. The specific reviews are then attached to their corresponding features.

The machine learning approaches have also been used for sentiment analysis [3]. They use bag-of-words (BoW), Part-Of-Speech (POS) information and sentence position as features. They represent reviews as feature vectors that are sent as input to a learning device usually Naïve Bayes and SVM.

In [4], the sentiment analysis of unstructured reviews has been carried out. The authors identified the features present in the review and then classified them using Naïve Bayes classifier and calculated their positive, negative and neutral polarity distribution.

The problem of assigning a numerical score (one to five stars ranking) to a review is presented [5]. They described it as a multi-label classification problem, and presented it using two supervised learning approaches; Naïve Bayes (NB) and Support Vector Machines. Classification of movie

reviews is presented in [6] that used Support Vector Machine (SVM) algorithm with 'bag of words' (BoW), in which some of the special features are chosen. However, the constraint here is its main focus on adjectives and modifiers that express assessment. The researchers used sentiment analysis of user reviews, and proposed an effective method to classify the reviews into positive, negative, and neutral classes by including slangs using different types of lexicons [7]. The analysis of word dependency and structure of contextual information was incorporated with Rule based approaches for sentiment analysis [8, 9].

The lexicon-based approach proposed by [10] uses rule-based technique to classify the reviews. They classified the reviews using three steps, namely, (i) filtering, (ii) subjectivity classification, and (iii) sentiment scoring.

The foremost motivation of this work is presented by a lexicon-based method to sentiment classification [11]. The author used dictionaries of positive or negative words to for classifying reviews. A semantic calculator was proposed by incorporating intensifiers and negation words.

3. METHODOLOGY

In this paper, review data, summary and ratings provided in the dataset have been analysed and it was observed that feedback ratings are better indicators of customers' opinion in comparison to that of the whole review. Reviews, summary and overall rating from the dataset were extracted. Reviews and summaries, were then processed to calculate their sentiment value by using two different approaches; one is lexicon-based method [11] with a rule-based approach and second using TextBlob library.

The resultant sentiment values were then ranked from 1 to 5 based on proposed rule and finally a statistical comparison was done to show that feedback ratings are better indicator of sentiments.

3.1 Dataset

The dataset used in the current research has been downloaded from "Amazon Product Data" [12] that contains 1,98,502 records of review data related to beauty products from 2009 to 2014. With rapid growth of internet, the number of reviews available online is increasing every year. The review data of 2014 that contained maximum number of records (a total of 58,309) against previous years was extracted. So, 58,309 records were analysed for the experiment. The structure of each record is given in Table 1.

Table 1: Structure of Record

Field Name	*Description*
reviewerID	ID of the reviewer
asin	ID of the product
reviewerName	name of the reviewer
helpful	helpfulness rating of the review
reviewText	text of the review
overall	rating of the product
summary	summary of the review
unixReviewTime	time of the review (unix time)
reviewTime	time of the review (raw)

Out of the above mentioned 9 fields in Table 1, we extracted reviewText, overall and summary fields for our research.

3.2 Lexicon-Based Approach

Two different dictionaries have been used here in order to check the polarity of words [13]. One contains a list of 4784 negative words and other one contains a list of 2006 positive words. We also created two files; one for negative words (like not, never etc.) which when precedes some positive or negative word, flips its polarity and one for modifiers (like very, highly etc.) that adds onto addition weightage to the polarity when precedes any word. Python 3.2 is a versatile language and supports extremely stable libraries. It is used to implement the algorithms. For human language preprocessing, Natural language Toolkit (NLTK) was used. It provides an effective platform for building Python programs.

3.2.1 *Preprocessing*

A typical review contains word variations, negative words, modifiers (that adds on to the weightage of positive or negative words) etc. For implementing lexicon-based method, the reviews need to be preprocessed in order to normalize the text into a suitable form to extract the sentiments.

The preprocessing steps used are:

- Tokenization: Tokenization is a step that splits up longer strings of text into smaller pieces, called tokens, that retain all the essential data. Reviews were present in the form of sentences in the dataset which were tokenized using python's NLTK (Natural Language Toolkit) into a list of words. Further processing was performed after a review has been appropriately tokenized.
- Stop Words: A stop word is a commonly used word (such as "the", "a", "an", "in") that must be ignored as they are not supposed to take up unnecessary space and valuable processing time. For this, we removed them by updating the list of stop words provided by python's NLTK (Natural Language Toolkit) with more words that we consider to be included in stop words (such as ";", "!", "-").
- Lemmatization: It is the process of grouping together the varied inflected forms of a word so that they can be analysed as a single unit. It is basically the morphological analysis of words as it links words with similar meaning to one word. NLTK (Natural Language Toolkit) Word Netlemmatizer has been used for this.

3.2.2 *Sentiment Calculation*

Sentiment calculation is the total sum of the sentiment bearing positive or negative words of a review. The sentiment calculation is derived from established heuristics that are based on sentiment orientation of the words. Each word from the review is compared against positive and negative dictionary and its occurrence is recorded as count. If a match is found in negative dictionary the value of count is decremented and if it is found in positive dictionary the value is incremented. Whenever a match is found, the preceding word is checked against the list of negations (like "not", "never" etc.) and the value of count is accordingly flipped. If the preceding word is a modifier (like "very", "highly" etc.) then count is incremented/decremented by 50%. Finally, each review is assigned a rank from 1 to 5 using a set of heuristics.

The sentiment of summary data provided in the dataset is also processed in the same way. Both set of ranks, one derived from reviews and other from the summary data were then statistically compared with each other and against the overall rating provided in the dataset using coefficient correlation to show how reviews and summary data fare against the overall rating.

The sentiment calculation algorithm is shown in Algorithm 1.

Algorithm 1: Sentiment Calculation Using Lexicon-Based Method

```
Data: Pre-processed Amazon Review data
Result: Output: Reviews Ranked from 1 to 5
count as Sentiment value
rank as Rank value
For each Review in Review_set:
    For each word in the Review:
        If word in negative_dictionary then
                Decrement count
                If preceding word in negation then
                        Increment count by 2
                If preceding word in modifiers then
                        Subtract 0.5 from count
        If word in positive_dictionary then
                Increment count
                If preceding word in negation then
                        Decrement count by 2
                If preceding word in modifiers then
                        Add 0.5 to count
    If count <=-1 then
        rank=1
    Else if count >-1 And count<=-0.25 then
        rank=2
    Else if count > -0.25 And count <= 0.25 then
        rank=3
    Else if count > 0.25 And count <=1 then
        rank=4
    Else
        rank=5
    count=0
```

3.3 Sentiment Calculation Using TextBlob

In the second approach, TextBlob library of Python has been used to calculate sentiment value of reviews and summary data available in the dataset. *TextBlob* is a Python library for processing textual data. It provides an API for common natural language processing (NLP) tasks. Result sets were then ranked from 1 to 5 using a similar set of heuristics. The results are again statistically compared with each other and against the overall rating provided in the dataset.

4. RESULT AND DISCUSSION

The results obtained after statistical comparison are shown in Table 2. The value of Correlation Coefficient indicates a weak relationship

between review vs summary whereas a moderate relationship between review vs rating and summary *vs.* rating [14]. It shows a maximum value of 0.44 among summary and rating in Lexicon-based approach and 0.40 when TextBlob library is used.

Table 2: Correlation Coefficient among Review *vs.* Summary *vs.* Rating

Method Used	*Lexicon-Based Approach*	*Using TextBlob Library*
Review/Rating	0.35	0.37
Review/Summary	0.26	0.24
Summary/Rating	0.44	0.40

An average human reviewer provides a concise and accurate statement about a product in summary whereas a review may contain detailed information and a computational sentiment calculator may yield wrong results due to the overuse of positive words as compare to negative words in a negative review and vice versa. Such reviews belong to the category of thwarted expectations [15]. An example from dataset to illustrate this concept is hereunder:

> *"I am not knocking it as a product. The smell was nice and the exfoliation was gently, but it did nothing for my combination skin."*

This review has positive polarity throughout but flips in the end. In such cases, the sentiment calculated computationally may wrongly misinterpret a positive polarity whereas the summary of the review gives more precise result that matches the overall rating. The rating in the dataset corresponding to this review is 1 that represents a negative review and the summary is "this facial scrub did nothing for me ".

Thus, by analysing the relationship of reviews, summary and overall rating specified in the dataset, it becomes evident that the summary and ratings provides a much more accurate opinion of the reviewer. Since it is more time consuming to process textual data, ratings can be safely used for analysis of user sentiments.

5. CONCLUSION AND FUTURE SCOPE

We have shown the statistical comparison between summary data and review data against overall feedback rating data present in the dataset in order to validate that online customers can make use of overall rating to make decisions about any product as they are better indicators of

customers' opinion regarding a product as compare to whole review. We calculated the sentiment values of summary and review data using (i) Lexicon-based approach and (ii) TextBlob library and then sent them as inputs to the statistical comparison. The value of correlation coefficient indicates that summary data is comparatively closer to overall rating as compare to that of review data due to the presence of reviews that belong to the category of thwarted expectations. Both the approaches used above show similar results.

The proposed system can be extended to incorporate sentiment analysis at aspect level instead of sentence level. Other approaches based on machine learning or hybrid techniques can be used to calculate the sentiment that can be further analysed statistically against the results derived in the current work.

REFERENCES

[1] Liu, B., "Sentiment Analysis and Opinion Mining," Morgan and Claypool Publishers, pp. 18–19, 27–28, 44–45, 47, 90–101, May 2012.

[2] M. a. B.L. Hu, "Mining opinion features in customer reviews," in *Proceedings of the 19th National Conference on Artificial Intelligence, California*, 2004.

[3] S. A. E. a. F. S. Baccianella, "Multi-facet rating of product reviews.," in *Proceedings of the 31th European Conference on IR Research on Advances in Information Retrieval, Heidelberg,* 2009.

[4] R. M. D. Nithya, "Sentiment Analysis on Unstructured Review," *in International Conference in Intelligent Computing Applications (ICICA),* March 2014.

[5] K. P. P. a. V. V. Sarvabhotla, "Supervised learning approaches for rating customer reviews," *J. Intell. Syst*., pp. 19: 79–94., 2010.

[6] C. N. G. a. S. A. Whitelaw, " Using appraisal groups for sentiment analysis," in *Proceedings of the 14th ACM International Conference on Information and Knowledge Management,* October , 31–November 5, 2005, 2005.

[7] A. S. K. A. A. Kundi F, "Detection and Scoring of Internet Slangs for Sentiment Analysis Using SentiWordNet," *Life Science Journal*, p. 11(9):66–72., 2014.

[8] K. a. S. P. Moilanen, "Sentiment composition," in *Proceedings of the Recent Advances in Natural Language Processing International Conference,* Sept. 27–29, Bulgaria, 2007.

[9] L. a. S. H. Dey, "Opinion mining from noisy text data," *Int. J. Document Anal. Recognition*, pp. 12: 205–226, 2009.

[10] K. A. A. S. a. A. M. Kundi FM, "Lexicon-Based Sentiment Analysis in the Social Web," *Journal of Basic and Applied Scientific,* p. 4(6): 238–248, 2014.

[11] J. B. M. T. K. V. a. M. S. M. Taboada, "Lexicon-based methods for sentiment analysis," *Computational. Linguistics,* Vol. 37, pp. pp. 267–307, 2011.

[12] J. M. R He, "Ups and downs: Modeling the visual evolution of fashion trends with one-class collaborative filtering," WWW, 2016.

[13] B. L. Minqing Hu, "Mining and Summarizing Customer Reviews," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD-2004)*, Aug 22–25, Seattle, Washington, USA, 2004.

[14] B. J. T. M. A. M. Ratner, "The correlation coefficient: Its values range between +1/−1, or do they?," *Journal of Targeting, Measurement and Analysis for Marketing*, Vol. 17, No. 2, p. pp 139–142, June 2009.

[15] A. P. B. a. J. S. N. Ramteke, "Detecting Turnarounds in sentiment analysis: Thwarting," in *Conference for association of computational linguistics,* 2013.

Digital Democracy and India: Study of Evolution in Citizen e-Participation Frameworks

Sharique Hassan Manazir[1] and Madhav Govind[2]

ABSTRACT: *Just like any other social system, democracy too needs regular innovation to respond to changing societal necessities and norms. The Internet and the corresponding rise of social media have provided scope for various innovative ways of engaging citizen and the democratic electorate on issues that matter for both as one nation. Many scholars have underlined the importance of citizen's participation in democracy and how ICT can play an important role in it. For instance, Gundogan (2010) believed that political participation is an essential component of democracy and consists of transferring citizens' ideas, opinions, feelings, desires, and critics to governing process at the stage of preparing policies, developing a solution and implementing the decisions. Hence, with the passage of time and post-advent of ICT, citizens' political participation has taken diverse forms both offline as well as online which gave rise to the concept of e-participation which acts as the backbone of Digital Democracy. The range and availability of citizen e-participation platforms provided by the government, non-government organizations as well as private players are increasing day by day as the idea of governance is shifting more and more from democracy to digital democracy. Various democratic governments are trying to define the future path of citizen participation in democratic decision through ICT based platform and thus with time various academic work with an innovative approach have come up. Although there are many frameworks to understand the pattern and motives of e- participation of people, yet none of them is adequate to capture the holistic view due to variations in the social context, digital inclusion, technological advancement, demographic changes etc.*

[1] Research Scholar, Centre for Studies in Science Policy, JNU, New Delhi, India. E-mail: shmanazir@gmail.com

[2] Professor, Centre for Studies in Science Policy, JNU, New Delhi, India. E-mail: m_govind@mail.jnu.ac.in

In this study, we have analyzed the various innovative frameworks to understand the transition of citizens' direct participation in Democracy to their e-participation in Digital Democracy. This study provides new insight into how different tools are used to engage citizen in governance through various e-participation models and what are challenges to enable the citizen to participate in the democratic process in India.

Keywords: Digital Democracy, e-Participation, e-Governance, ICT, Digital Inclusion.

1. INTRODUCTION

Citizen Participation has always been an important aspect of any form of governance and becomes more important when we talk about democratically run governments. As per the available literature under democratic theories, there exists varying opinion upon citizen political participation. While scholars like Schumpeter limited idea of citizen political participation to voting only. (Shumpeter, 1946) Dahl accepted wide range of citizen action besides voting as political participation. (Bengtsson, 2008)

A well-planned citizen participation framework in governance can lead to informed citizen who would understand the technically difficult situations and would suggest a holistic, communitywide solutions. Conventionally in governance, there were different modes for ensuring citizen participation at different levels using different procedures, techniques and methods. They included electoral voting's, citizens juries, citizens panels, committees, consensus conferences, scenario workshops, deliberative polling, focus groups, multi-criteria mapping, public meetings, rapid and participatory rural appraisal etc. (Pimbert and Wakeford, 2001) Definition for political participation by (Vissers & Stolle, 2013) i.e. "political participation refers to all forms of involvement in which citizens express their political opinion and/or convey that opinion to political decision-makers," fits well into the online realm as well because it does not specifically mention whether these activities take place online or offline.

Lately, there has been various e-participation frameworks by scholars worldwide to understand and improve citizen e-participation. Though the definition of e-Participation is well understood, the study approach of "e-participation" has varied been worldwide which could be divided into five categories:

1. E-Governance Platform Analysis
2. Government Ministerial Website Content Analysis

3. Social media Platform Analysis
4. Citizen Enabled e-participation Platform Analysis
5. Government enabled e-participation Platform Analysis.

There has been enough research from the perspective of E-Governance projects evaluation. Most of the research have been done by the government department as part of its overall e-assessment strategy in collaboration with independent third parties such as IIT's, IIMs etc. (MCIT 2008; IIT and IIM-A, 2010). These studies have focused on broad parameters such as 'impact on outreach', 'cost of accessing services', 'quality of services' and "overall governance". The well-known scholars like Norris Pippa (2001) studied the e-participation in collaboration with United Nations, from the perspective of content analysis of 196 countries government websites and tried to understand how information provided on website is enabling citizen to interact with government. Scholars like Chun & Cho (2012) in their paper titled "E-Participation and transparent policy decision making" studied both the contents and the features of government websites as well as the social media platforms in Korea. Most of the research on e-participation has relied on "citizen enabled e-participation platforms". Macintosh et al are well-known name who have studied e-participation through that approach. Further similar approach and similar framework has been used by European Union too. (Lironi, 2016)

Table 1: Major form of Academic Work on E-participation

Sr. No	*Forms*	*Tools*	*Examples*	*Research*
1.	E-Governanc e Platforms	Services like feedback through email or discussion forum over dedicated e-governance scheme website	MCA21 Passport Income Tax	(MCIT, 2008); (DIT & IIM-A, 2010) etc.
2.	Ministry Website	Provision for policy feedback over specific policy draft; no formal feedback on acceptance or discussion over policy proposal or amendment.	TRAI	(Norris P., 2004)[UN World Public Sector Report Deepening Democracy via E-Governance]
3.	Social Media	Citizen share their issues and discuss with social media account of politicians, ministries etc.	Facebook, Twitter, Instagram	(Magro, 2012); (Effing, Hillegersberg, & Huibers, 2011);

Sr. No	*Forms*	*Tools*	*Examples*	*Research*
				(Bonsóna, Torresb, Royob, & Floresc, 2012); (Salamat & Hassan 2011) etc
4.	Government Enabled e-Participation Platforms	Citizen can participate in formal policy-making process through different tools.	My-Gov; Common Service Center etc	(Lamba, Yadav, & Lela, 2016) (Limited Conference Proceedings Only)
5.	Citizen Enabled e-Participation Platforms	Online Petitions, discussion forums	Change.org; gov2u.org etc	(Alathur, Ilavarasan, & M.P.Gupta, 2007); (Maciintosh, 2004)

Currently, there are various organizations engaged in the task of mapping citizen e-participation on the basis of different forms of parameters. (Heeks, 2006) Majority of research reports only focus upon e-government ranking, but UN E-Government Survey report comes up with e-participation index apart from E-Government index to evaluate yearly status of countries and rank them in order. UN derives e-Participation index by focusing on the government's use of online services for providing information to its citizens or "e-information sharing", interacting with stakeholders or "e-consultation" and engaging in decision-making processes or "e-decision-making" (UN, 2018) Below chart shows 2018 ranking of top 15 nations worldwide based upon their e-participation indices.

Table 2: e-Participation Ranking

E–Participation Index		
Sn No	*Country*	*Index*
1.	Republic of Korea	1
2.	Denmark	1
3.	Finland	1
4.	Netherlands	0.9888
5.	Japan	0.9831

E–Participation Index		
Sn No	*Country*	*Index*
6.	New Zealand	0.9831
7.	Australia	0.9831
8.	Spain	0.9831
9.	United Kingdom	0.9831
10.	United States of America	0.9831
11.	Norway	0.9775
12.	Brazil	0.9719
13.	France	0.9663
14.	Singapore	0.9663
15.	India	0.9551

2. EVOLUTION OF E-PARTICIPATION FRAMEWORK

Citizen political participation in governance decision making as well as public policy formulation through top-down approach (government or private entity initiated) or bottom-up approach (citizen-initiated) has been studied worldwide from various perspectives even before advent of Information Communication Technology.

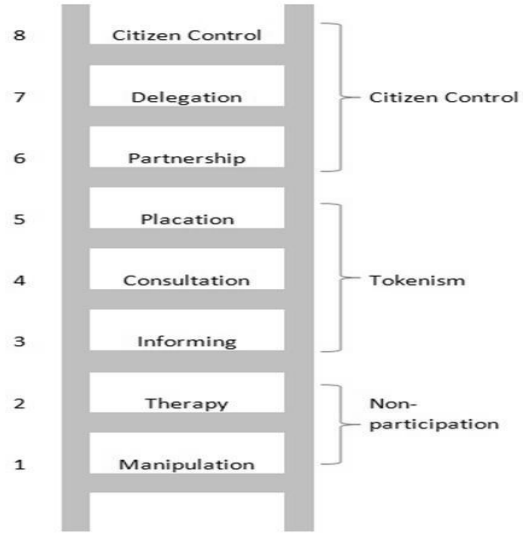


Fig. 1: Arnstein's Ladder of Citizen Participation (Arnstein S.R., 1969)

Before the advent of ICT way back in 1969, Arnstein's ladder of participation was the pioneer in study of citizen participation in governance and policymaking. She differentiated top-down approach into three layers namely Non-participation, tokenism and on top Citizen Power. (Cornwall, 2008). She characterized citizen participation as citizen power and defined it as "… the redistribution of power that enables the have-not citizens, presently excluded from the political and economic processes, to be deliberately included in the future". (Arnstein S. R., 1969) Later, Glass proposed citizen participation framework whose objectives included information exchange, education and support building, decision-making supplement and input probing. (Glass, 1979) After that, there were various theoretical model of citizen participation from various theoretical and academic perspectives which did not involved ICT or citizen participation in governance necessarily. Jule Pretty's "Typology of Participation" (1995), Sarah White's "Typology of Participation" (1996), Phil Tre Seder "Degree of Participation" (1998),

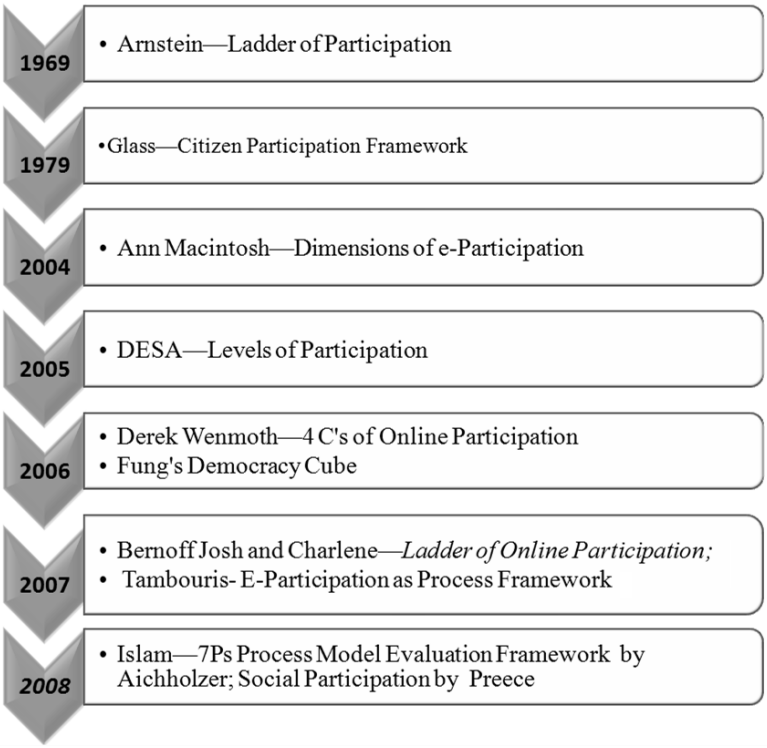


Fig. 2: Timeline for Evolution of e-Participation framework

Scott Davidson's "Wheel of Participation" (1998), OECD Active Participation Framework (2001), Herry Shier's "Pathways to Participation", UNICEF's "Strategic Approach to Participation" (2001), Jans & De Backer's "Triangle of Youth Participation (2002), IAP2's "Spectrum of Public Participation" (2007) and Diane Warbuton's "Engagement in the Policy Cycle" (2007) were worth mentioning such theoretical framework. It is an established fact that e-Participation from the perspective of e-governance and e-Democracy is still an evolving research area and thus there is very limited universally accepted framework catering all its aspects. (Kangere, 2002) First theoretical framework for citizen e-participation in public policy formulation was proposed by Ann Macintosh in the 2004. She identified e-enabling, e-engaging and e-empowering as major three dimensions of e-participation of citizen in a top-down approach for policymaking by government in consultation with citizen.

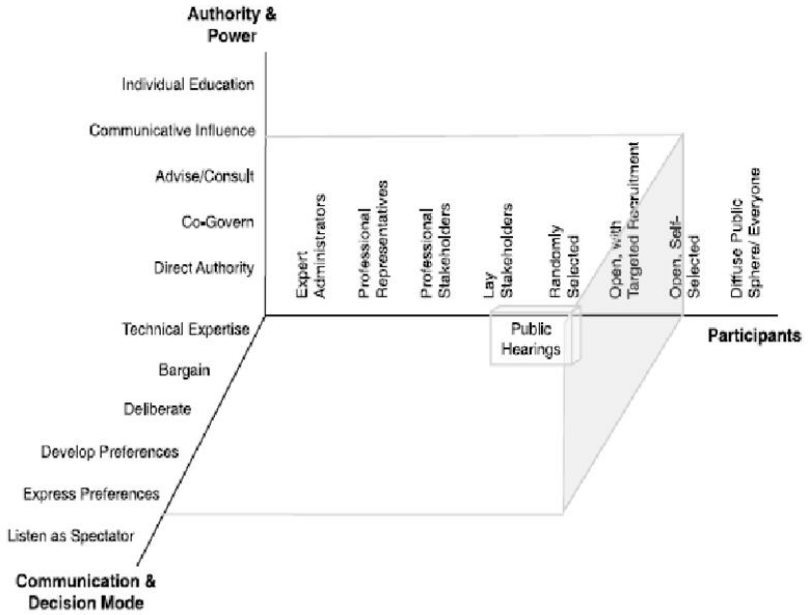


Fig. 3: Fung's Democracy Cube (Fung, 2006)

E-information, e-consultation and e-decision making are three-stage framework for explaining citizen participation in public policy formulation according to DESA approach. (DESA, 2005) Derek Wenmoth is the year 2006 developed 4C's of Citizen participation to

show how people participate in online communities in four different roles which were consumer, commentator, contributor and lastly commentator with all the four levels of participation with different set of motivation, behavior and end result of participation.. (Wenmoth, 2011) In 2006 Fung introduced "Democracy Cube" which has three dimensions: "the scope of participation", "the mode of communication and decision" among participants, and "the extent of the participants' authority". Participants could be inclusively or exclusively chosen to participate, under communication axis, six types of communication in citizen participation were identified: listening as a spectator, expressing preferences, developing preferences, aggregating and bargaining, deliberating and negotiating, and deploying expertise. Lastly, the authority axis was meant to measure the impact of public participation. (Fung, 2006)

On similar lines, Li and Bernoff came up with Ladder's of Online Participation in the year 2007 which basically dealt with user participation on different web portals with mainly six levels of user participation namely Inactive, Spectators, Joiners, Collectors, Critics and lastly Creators who create the online content. This theoretical framework too was limited for citizen participation on web portals and was not specifically dealing with public policy formulation. (Li & Bernoff, 2007)

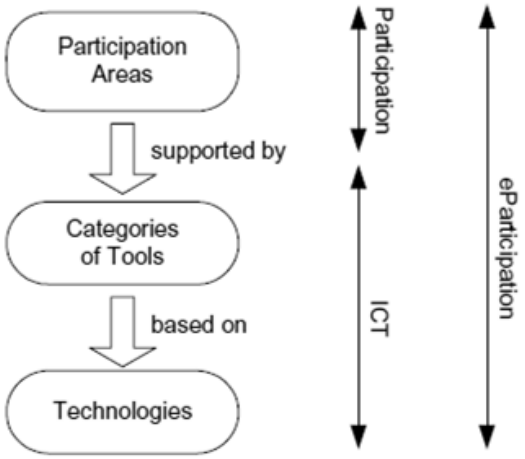


Fig. 4: Tambouris Framework of E-Participation (Efthimios TAMBOURIS, 2007)

In year 2007 Tabouris proposed framework for e-participation where he differentiated the whole process into five major levels namely Democratic processes, Participation areas, Participatory techniques, Categories of tools,

Technologies which could flow into two major directions i.e. the democratic processes which is top-down to the technologies or bottom-up technology-driven towards democratic processes. (Tambouris, 2007)

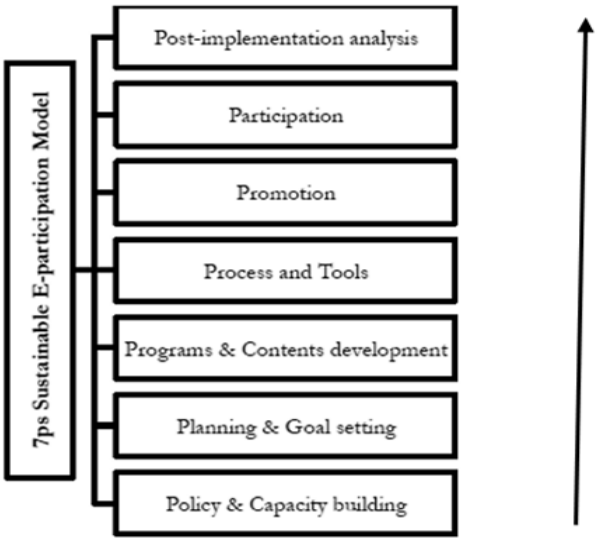


Fig. 5: Islam's 7P Sustainable E-Participation Model

Islam discussed in detail about universal applicability of e-Participation model since socio-economic, technological and digital inclusion issues vary from country to country. (Islam M. S., Towards a sustainable e-Participation implementation Process, 2008) Thus, he proposed 7Ps Sustainable e-Participation framework in year 2008 which has seven phases namely Post-implementation analysis, Participation, Promotion, Process and Tools, Programs and Contents development, Planning and Goal setting and lastly Policy and Capacity building in top-down manner.

3. INDIA AND E-PARTICIPATION FRAMEWORK POLICY

Recently Ministry of Communication and Information Technology, Government of India produced "Framework for Citizen Engagement in E-Governance" in 2012. For the first time, it defined Citizen participation, needs for ways i.e Information Sharing, Consultation, Joint Assessment and Shared Decision Making to share information and make citizens a partner in decision making. According to this framework citizen engagement in decision making would result in a better-informed decision. It defined Citizen Engagement as "an interactive two-way

process that encourages participation, exchange of ideas and flow of conversation."It conceptualized "information sharing, Consultation, Joint assessment, Shared decision-making and collaboration" as stages of an engaging citizen for participating in the governance. This framework was specifically meant for e-governance initiatives and not for e-participation platform meant for shared decision making. In the proposed framework citizen e-participation framework the question of outreach is dealt mostly from the perspective of Digital Divide i.e. accessibility while digital inclusion is given very less emphasis. While studies have found non-uniform access to ICT based on socioeconomic status, gender, age, race, geography, caste, language barrier etc scholars suggested that digital divide should must be studied as a phenomenon from the perspective of global, social as well as democratic divide. (Räsänen, 2006) According to IAMAI report based on census 2011, in Rural India only 20.25% population has access to Internet while in Urban India 65% has access to Internet. (UN, United Nations E-Government Survey 2018, 2018) In India use of internet is majorly male dominated in both rural as well as urban demography as per the latest reports. While 41% internet users in Urban India are female, only 36% of the total internet users in Rural India are females. (IAMAI, 2018) According to recent research reports, in India about 60 million people are disabled and 42.5% of them are women while 75% of people with disabilities come from rural areas. Thus, in India barriers to accessibility are all confluence when we talk about citizen political participation with governance using internet. India has total 780 official and nonofficial languages which is world's second-highest number of languages by any country.

4. FUTURE CHALLENGES

When we talk about citizen e-participation using Internet as a public sphere the question of outreach is dealt mostly from the perspective of Digital Divide while digital inclusion is given very less emphasis. (Helbig, Garcia, & Ferro, 2008) While some scholars still argue that issues pertaining to digital inclusion is always on a trajectory of being resolved with time as the new technology comes in and thus is resolved with time. (Compaine, 2001) Even according to UN E-Government Survey of 2018 in India finding content online in the local language is still a big challenge and despite progress, providing local, relevant and useful content, in addition to raising awareness about it, requires significant effort. (UN, 2018) Citizen E-Participation rights violation in form of abrupt internet access shutdown is another concern. Internet shutdowns is

not a new phenomenon but is indeed now a regular one in countries like, Syria, Nepal, Democratic Republic of Congo, Burundi, Ethiopia, Iraq, Kazakhstan, Pakistan, Sudan, Uzbekistan, Yemen, China, Zimbabwe, North Korea, Uganda, Bahrain, Bangladesh, Egypt, Turkey, Cameroon, Gambia, Myanmar, Egypt and India too. According to United Nations report shutting down internet with intention to disconnecting people even to quell political unrest is human rights violation. (Rue, 2011) Report states that "while internet blocking and filtering of internet denies users access to desired content, states even try to cut off access to the Internet completely. The Special Rapporteur considers cutting off users from internet access, regardless of the justification provided in violation of article 19, paragraph 3, of the International Covenant on Civil and Political Rights." According to recent research, India leads the world in the number of internet shutdowns, with over 100 reported incidents in 2018 alone. (Shahbaz, 2018) According to research reports by Boston Consulting Group's projection for 2016, internet contributes around 5.6% to India's total yearly GDP. In 2016 only India lost around US$968 million as total economic impact of internet disruptions, both in terms of number of days and total GDP lost. (Darrell West, 2016) Thus internet shutdowns not only are oppressive in nature but affects nations economy too since now entire economy is based on internet connectivity. Thus, we are not yet close to a robust citizen e-participation framework which caters to all the social, economic, political, demographical as well as technological issues of nations through uniform standard and hence there is immense scope of qualitative as well as quantitative research.

REFERENCES

[1] Alathur, S., Ilavarasan, P.V. and M.P. Gupta. (2007). Citizen participation and effectiveness of e-petition: Sutharyakeralam – India. *emerald insight*.

[2] Amna, E. and Ekman, J. (2014). Standby citizens: Diverse faces of political passivity. *European Political Science Review,* 261–281.

[3] Ann.Macintosh, Davenport, E., Malina, A. and Whyte, A. (2002). Technology to Support Participatory Democracy. *IGI Global*.

[4] Arnstein, S.R. (1969). A Ladder Of Citizen Participation. *Journal of the American Institute of Planners*, 216–224.

[5] Astrom, J. (2001). Digital Democracy: Ideas,Intentions and Initiatives in Swedish Local Governance. *ECPR Joint Session.* Grenoble: Department of Social Science, Orebo University.

[6] Bakker, T.P. and Vreese, C.H. (2011). Good News for the Future? Young People, Internet Use, and Political Participation. *Sage Publication*, 452.

[7] Bannerjee, I. (2003). rhetoric and reality: The Internet Challenge for Democracy in Asia. Singapore: Eastern University Press.

[8] Barboni, G., Field, E., Pande, R., Rigol, N., Schaner, S. and Moore, C.T. (2018). *A Tough Call: Understanding barriers to and impacts of women's mobile phone adoption in India.* Harvard Kennedy School–Evidence for Policy Design.

[9] Bebbinton, A. (1996). Organizations and Intensifications: Campesino Federations, Rural Livelihoods and Agricultural Technology in the Andes and Amazonia. World Development Report.

[10] Bengtsson, A. (2008). *Politisk deltagande.* Lund: Studentlitteratur.

[11] Berry, R. (2011). Older people and the internet. *ILC*.

[12] Bonsóna, E., Torresb, L., Royob, S. and Floresc, F. (2012). Local e-government 2.0: Social media and corporate transparency in municipalities. *ELSEVIER*, 123–132.

[13] Chopra, R. (2006). Global primordialities: Virtual identity politics in online Hindutva and online Dalit discourse. *New media & Society*, 187–206.

[14] Christensen, H.S. (2015). Process or outcome? How the Citizens' Initiative to Ban Fur Farming Affected Political Trust among Users of Avoin Ministeriö. *Research On Finnish Society*.

[15] Chun, S.A. and Cho, J.S. (2012). E-Participation and transparent policy decision making. *Information Polity*, 129–145.

[16] Compaine, M.B. (2001). The Digital Divide: Facing a crisis or creating a myth? Cambridge: MIT Press.

[17] Connor, D.M. (1988). A New Ladder of Citizen Participation. *National Civic Review*, 248–257.

[18] Cornwall, A. (2008). Unpacking 'Participation':models, meanings and practices. *Oxford University Press and Community Development Journal*.

[19] Cunningham, F. (2010). *Theories of Democracy: A critical introduction.* London: Routledge Publications.

[20] Dalton, R.J. (2014). Citizen politics: Public opinion and political parties in advanced industrial democracies. Los Angeles: CQ-Press, Sage.

[21] DESA, U. (2005). UN Global E-government Readiness Report 2005: From E-government to E-inclusion. New York: United Nations Publications.

[22] DIT and IIM-A. (2010). *Impact Assessment of Commercial Taxes and Computerization.* New Delhi: Nutech Photolithographer.

[23] E.Schwartzenberg. (1978). *A Historical Atlas of South Asia.* Chicago & London: University of Chicago Press.

[24] Effing, R., Hillegersberg, J.V., and Huibers, T. (2011). Social Media and Political Participation: Are Facebook,Twitter and YouTube Democratizing Our Political Systems? *IFIP International Federation for Information Processing*, 25–35.

[25] Efthimios Tambouris, N.L. (2007). A Framework for Assessing eParticipation Projects and Tools. *Hawaii International Conference on System Sciences*. Hawaii: IEEE.

[26] Eisenhart, M.A. (1991). Conceptual frameworks for research circa 1991: Ideas from a cultural anthropologist; implications for mathematics education researchers. *Proceedings of the 13th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 202–219). Blacksburg: VA.

[27] Feenberg, A. (1991). *Critical Theory of Technology*. Oxford University Press.

[28] Fraser, N. (1990). "Rethinking the Public Sphere: A Contribution to the Critique of Actually Existing Democracy". Duke University Press.

[29] Freeman, J. and Quirke, S. (2013). Understanding E-Democracy. *Journal of E-Democracy*, 141–153.

[30] Fung, A. (2006). Putting the Public Back into Governance: The Challenges of Citizen Participation and Its Future. *Public Administration Review.*

[31] Gaventa, J. and Valderrama, C. (1999). Participation, Citizenship and Local Governance. *'Strengthening participation in local governance'*. Valencia: Institute of Development Studies.

[32] Glass, J.J. (1979). Citizen Participation in Planning: The Relationship Between Objectives and Techniques. *Journal of the American Planning Association*, 180–189.

[33] GOI, M. (2017 йил 24–3). *Ministry of Electronics and Information Technology*. Retrieved 2017 йил 27–3 from meity.gov.in: http:// meity.gov.in/content/assessment-e-governance-projects

[34] Govind, M. and Basu, A. (2017). Community Participation or Manufactured Consent? Strategies for Implementation of Drinking Water Project "JalNidhi" in Kerala (India). *SAGE Publications*, 1–19.

[35] Habermas, J. (2006). Political Communication in Media Society:Does democracy still enjoy an epistemic dimension? The impact of normative theory on empirical research Communication THeory. 411–426.

[36] Habermas, J. (1989). The Structural Transformation of the Public Sphere. Cambridge: MIT Press.

[37] Hauser, G.A. (1999). *Vernacular Voices: The Rhetoric of Publics and Public Spheres.* Columbia: University of South Carolina Press.

[38] Heeks, R. (2006). Understanding and Measuring eGovernment: International Benchmarking Studies. In *E-Participation and E-Government: Understanding the Present and Creating the Future.* New York: Department of Economic and Social Affairs, Division for Public Administration and Development Management.

[39] Helbig, N., Garcia, R.G. and Ferro, E. (2008). Understanding the complexity of electronic government: Implications from the digital divide literature. *Government Information Quarterly: Elsevier*, 89–97.

[40] Hill, K.A. (1998). Cyberpolitics: Citizen Activism in the Age of the Internet. *Rowman and Littlefield.*

[41] Holzer, M. and Kim, S.-T. (2006). *Digital Governance in Municipalities Worldwide.* New York: National Center for Public Productivity.

[42] IAMAI (2018). *Internet in India 2017.* New Delhi: Kantar IMRB.

[43] Islam, M.S. (2008). Towards a sustainable e-Participation implementation model. *European journal of ePractice*.

[44] Jayaswal, K. (1943). Hindu Polity: A Constitutional History of India in Hindu Times. Bangalore: Indowestern Publishers.

[45] Jayaswal, K.P. (1924). *Hindu Polity.* Varanasi: Chaukhambha Indowestern Publication.

[46] Kalita, T. and Qasim, A. (2016). *The Future of Internet in India.* New Delhi: NASSCOM.

[47] Kangere, S.R. (2002). *What is Participation? CBR: A participatory strategy in Africa.* London: Centre for International Child Health (CICH).

[48] Keane, J. (2008, March 13–14). Monitory Democracy? *ESRC Seminar Series*.

[49] Kedzie, C.R. (1997). A Brave New World or a New World Order?" in Culture of the Internet. *Lawrence Erlbaum Associates, Inc*.

[50] Kumar, C.S. (2016). Dalit Camera: Through Untouchable Eyes. *IRA International Journal of Education and Multidisciplinary Studies*, 59–68.

[51] Kumar, C.S. and Subramani, R. (2014). Internet as an alternative media for Dalits in India: Prospects and challenges. *Journal Of Humanities And Social Science*, 125–129.

[52] Lamba, A., Yadav, D. and Lela, A. (2016). Citizen Pulse: A text analytics framework for proactive e-governance—A case study of mygov.in. *Proceedings of the 3rd ACM IKDD Conference on Data Sciences.* CODS.

[53] Li, C. and Bernoff, J. (2007). *Social Technographics.* Forrester.

[54] Lironi, E. (2016). Potential and Challenges of E-Participation in European Union. Brussels: European Parliament.

[55] Livingstone, S. and Helsper, E. (2007). Gradations in digital inclusion: children,young people and the digital divide. *SAGE Publications.*

[56] Macintosh, A. (2004). Characterizing e-participation in policy-making. *37th Hawaii International Conference on System Sciences.* Hawaii: IEEE.

[57] Mackie, G. (2004). Schumpeter's Leadership Democracy. *American Political Science Association, Hilton Chicago and the Palmer House Hilton* (pp. 1–39). Chicago, IL: American Political Science Association.

[58] Magro, M.J. (2012). A Review of Social Media Use in E-Government. *Administrative Sciences,* 148–161.

[59] MCIT (2008). *Impact Assessment of E-Governance Projects.* New Delhi: Landscape Publishing Solutions.

[60] NCRB (2016). *Crime in India-2016.* New Delhi: National Crime Records Bureau-Ministry of Home Affairs (India).

[61] Norris, D.F. (2010). E-government... not e-governance... not e-democracy: Not now! Not ever? *Proceedings of the 4th International Conference on Theory and Practice of Electronic Governance* (pp. 339–346). New York: ACM.

[62] Norris, P. (2001). Digital Divide: Civic Engagement, Information Poetry and Internet Worldwide. New York: Cambridge University Press.

[63] Norris, P. (2004). Draft chapter for the UN World Public Sector Report:Draft chapter for the UN World Public Sector Report. Harvard University.

[64] O. Sæb., J. and Nyvang, T. (2009). The Role of Social Networking Services in eParticipation. *ePart Conference*, (pp. 46–55).

[65] OECD (1997). National Innovation System. *Organisation for Economic Co-operation and Development.* Paris: Organisation for Economic Co-Operation and Development.

[66] Ono, H. (2003). Gender and the Internet. *Social Science Quarterly*, 110–121.

[67] Palmer, N.A. and Perkins, D.D. (2012). Technological Democratization: The Potential Role of ICT in Social and Political Transformation in China and Beyond. *Neal A Palmer & Douglas D Perkins, Technological Democratization: The Potential Role of ICT in Social and Perspective on Global Development and Technology*, 456–479.

[68] Pimbert, M. and Wakeford, T. (2001). Deliberative Democracy and Citizen Empowerment. *PLA Notes*.

[69] Poushter, Jacob, Bell, J. and Oates, R. (2005). Internet Seen as Positive Influence on Education but Negative on Morality in Emerging and Developing Nations. Pew Research Centre.

[70] Räsänen, P. (2006). Information Society For All? Structural Characteristics of Internet Use in 15 European Countries. *European Societies*, 59–81.

[71] Reuters (2018). *The World's Most Dangerous Countries for Women-2018.* London: Thompson Reuters Foundation.

[72] Rue, F.L. (2011). Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression. United Nations.

[73] Salamat, Hassan, Fudzee and Azhar (55–62). A Framework for Formulating Malaysia's Public Policy through e-Participation. *Knowledge Management International Conference,* 2012.

[74] Shahbaz, A. (2018). Freedom on The Net 2018: The Rise of Digital Authoritarianism. Freedom House.

[75] Shumpeter (1946). *Capitalism,Socialism and Democracy.* New York: Harper & Row.

[76] Sobaci, Z. (2010). What the Turkish parliamentary web site offers to citizens in terms of e-participation: A content analysis. Information Polity.

The International Journal of Government and Democracy in the Information Age, 227–241.

[77] Srinivasan, M.I. (2014). Playing second innings online: an ethnographic study on the Internet Usage among the urban Indian middle class senior citizens of Hyderabad, India. *Global Media Journal.*

[78] Tambouris, E. (2007). A Framework for Assessing eParticipation Projects and Tools. *40th Hawaii International Conference on System Sciences*, (pp. 1–10).

[79] Tambouris, E., Liotas, N., D. Kaliviotis and K.Tarabanis (2007). Framework for Scoping eParticipation. *8th Annual International Digital Government Research Conference.* Philadelphia, USA.

[80] Teorell, J. (2006). Political participation and three theories of democracy: A research inventory and agenda. *European Journal of Political Research*, 787–810.

[81] The White House (2009). Transparency and Open Government, Memorandum for Heads of Executive Departments and Agencies. Washington: The White House.

[82] Tufekci, Z. (2013). Engineering the Public: Big Data, surveillance and computational politics. *First Monday.*

[83] UN (2018). *United Nations E-Government Survey 2018.* New York: Department of Economic and Social Affairs-UN.

[84] Vissers, S. and Stolle, D. (2013). The Internet and new modes of political participation: Online versus offline participation. *Information, Communication & Society,* 937.

[85] Wenmoth, D. (2011). Participation Models: Citizen, youth, Online. Nonformality.org.

A Literature Analysis on Big Data Paving Its Way to E-Business

Suman Madan[1], Kirti Bhardwaj[2] and Shubhangi Gupta[3]

ABSTRACT: *Big data stands as a substantial buzzword considering the modern business management frontier today. A great outburst of interest occurred for big data analytics from various industries including academia and the e-business industry. E-businesses are constantly looking for ways to retain customers and improve the customer experience, not realizing that solution of this problem is right at their fingertips. Big data corresponds to exceedingly enormous datasets which are analyzed and scrutinized to reveal trends and subsequently, patterns in human behavior. The terminology of big data is not restricted or limited to data only, but it refers to using complex datasets to derive focus, direction, and for decision-making within an organization or firm. This outburst had forced e-business firms to use big data analytics in their value chain process for higher productivity. This survey paper focuses on the different sorts of data held and its inverse usage for e-business along with different ways of providing security and safety for data corresponding to anonymizations. It also highlights the issues in big data with respect to e-businesses and tries to trigger the maximum usage of big data applications effectively.*

Keywords: Big Data, E-business, Security, Models, Tools and Trends.

1. INTRODUCTION

The emerging web-based technologies are giving e-business firms metamorphic gains like real-time customer services, personalized offers, dynamic prices etc. [16]. The Big Data Analytics (BDA) is providing major help to e-business firms to use the available huge data more efficiently, improving the decision-making; converting one-time

[1] Asst. Prof., JIMS, Rohini, Delhi, India. E-mail: suman.madan@jimsindia.org
[2,3]Student, JIMS, Rohini, Delhi, India. E-mail: [2]archnabhardwaj1969@gmail.com; [3]shubhangig16@gmail.com

buyers to repeat customers, finding the lowest possible inventory level, bringing higher conversion rates etc. [1]. Big Data is thus proving that it is a game-changer when talked about retail and e-business. If businesses can successfully implement effective big data strategies, predict trends, optimise pricing and forecast demand then they will reap the rewards of better customer experiences and bigger profits. From the transaction cost aspect, BDA is benefitting the online businesses in terms of improving communication between buyers and sellers; reducing efforts of searching, bargaining and after sales monitoring while also increasing managerial efficiency and providing new recommendation algorithms.

Lots of researches have been done on this domain but the literature shows that it is largely informal, fragmented study only. This paper is result of research question: "How big data analytics can create more business value than traditional analytics for e-businesses?" To get more insights on this research issue, this paper focusses on providing an in-depth accumulative knowledge about big data and its types, along with tools that can be used to analyse it. The study aims to discuss and support both the qualitative and quantitative aspect of Big Data, its present and future scenarios focusing on its effect on e-business as an emerging and evolving technology. [4]

Figure 1 shows the journey of raw data to business intelligence (BI) helping organizations. BI cites the technology-driven process that analyses

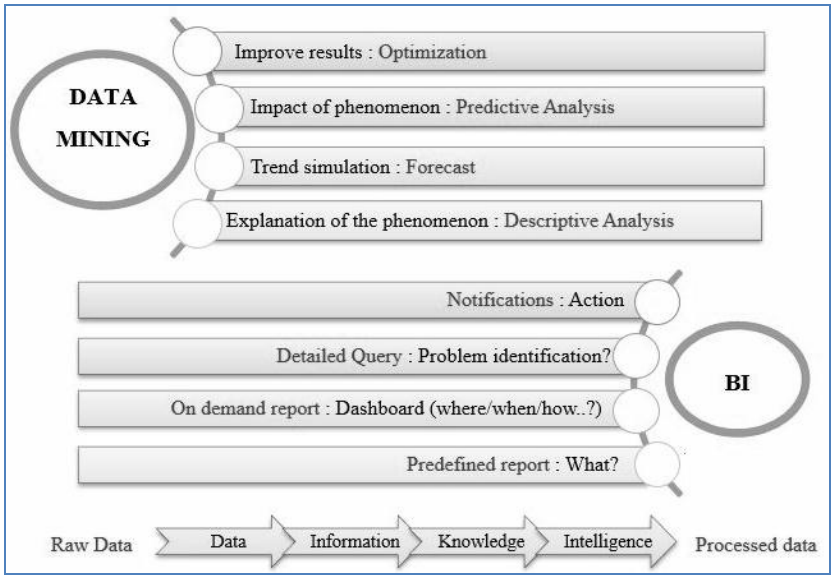


Fig 1: Process showing Conversion of Raw Data to Business Intelligence

data and produce appealable information for top management decision making task and thus in turn increasing market efficiency and lowering cost. BI is also helping e-businesses in predicting market trends, performing competitor analysis.

2. MOTIVATION

2.1 Literature Review

Ferreira, Pedrosa and Bernardino (2017) [6] proposed 4-level architecture for BI with e-business where in raw data of online business is transformed into knowledge by identifying purchase patterns of customers to improve CRM. The authors discussed cloud computing, big data and machine learning usage in data analytics, precisely. The issue raised by authors is organizational and related to user's data security and privacy, respectively.

Akter and Wamba (2016) [12] discussed good practices, nature of data and business value related to e-business based data for determining BDA. The authors highlighted few challenges in BDA like reaching directly to the business target along with internal changes in organization, identifying right and exact customer requirements to offer personalized offers, maintaining privacy of customer. But they didn't consider the security issue of customer's account details and need of skilled data analytical persons.

Aivalis, Boucouvalas and Gatziolis (2016) [12] proposed "log file analyser" application for accommodating technologies related to big data. The application allows the preloading of all e-business information along with customer information and then automatically posting advertisements in customer's social media accounts.

Chen and Laug (2016) [7] featured five V's corresponding to Big Data and e-business relationship and gave new research measurements that concentrate on real-time dynamic product pricing, analytics dependent on geographical constraints, customer behaviour and need of product managing and monitoring devices for taking care of mass requests. Sun, Strang and Zou (2015) [11] give the association between big data analytics and Business intelligence. Likewise, the authors propose how to improve business intelligence via big data analytics by giving new service based design or model.

2.2 Challenges

The aim of the paper is to understand the following [2]:

- To analyse big data characteristics and its need in E-Business
- Exploring the opportunities for E-Business by using Big Data Analytics
- Addressing the Big Data challenges
- To foresee future trends.

3. BIG DATA CHARACTERISTICS AND NEED IN E-BUSINESS

Since the markets both international and national are becoming customer-oriented, Big Data provides the benefits (Figure 2(a)) of time saving and less cumbersome buying, delivery and business, in general, especially for those who want to spend minimal time on regular chores like stocking groceries, etc.

This new data is characterized by what is commonly known as the 3Vs (Figure 2(b)):

- *V for Volume*: The growth rate of 50% a year, the volume of data available is growing exponentially with data cross-referencing being the key to determining the relevance of generated information.
- *V for Variety*: The various format types like text, photograph, etc, must now be added to a variety of internal and external providers; objects or people, while also encompassing the possible uses associated with raw data.
- *V for Velocity*: Given the rapid obsolescence of real-time data, its integration with other data to get a timely insight must be done rapidly.[2, 7]

These 3 V's of big data led to the addition of the following 2 V's:

- *V for Veracity*: Data obtained is limited in volume but controlled in terms of coherence and quality and so, its use hinges on the ability to neutralize these weaknesses without modifying the original information which leads to data reliability that has become an essential criterion as per the gigo13 principle application to Big Data.
- *V for Value*: Value assessment of a data item on the spot is difficult but one must be beware of falling into a restrictive pattern considering that a data source which may appear internally useless can have a monetizable value for another partner.[2, 7]

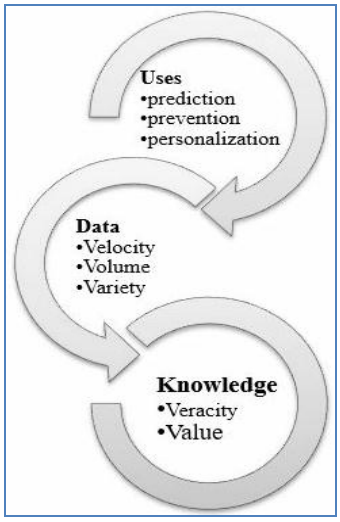


Fig. 2(a): Big Data and Advantages **Fig. 2(b):** Big Data Facilitating E-business

The knowledge thus derived is useful for the 3P's of any e-business:

- ***P for Prediction:*** Prediction is a huge part of big data uses as we need to anticipate the data to acquire adequate knowledge in order to predict problems, behaviours, tastes etc. Through big data we can collect and reconcile the data to apply analysis and predictive modelling [2, 7].
- ***P for Personalization:*** Personalising at the detailed level is the second category of uses. Big data basically helps us to create a solution for a specific group of people or an individual. For Example, creating a personalised website that is based on the information automatically builds the page and different processes are performed. The solution uses analysis and interactions for a customized contribution [2, 7].
- *P for Prevention:* The third use is for Prevention which implies that big data is used to identify the risks and dangers involved and prevent them [2, 7].

4. BIG DATA OPPORTUNITIES IN E-BUSINESS

In the current business scenario where data is considered as an asset worthy of monetary value, enterprises look for business transforming abilities from IT suppliers by embracing big data for:

1. *Risk Management:* This is a very important aspect as incorrect risk analysis can often lead to inadequate and unfit decision making and

higher expenses. Every risk needs to be identified with a systematic approach to prevent the company from threats due to data breaches. Analysis of big data can establish balance between opportunity and hazard. [3]

2. *Driving Research and Development*: R&D pertaining to Big Data is not so much about velocity as it is about viability, variety and sometimes even volume. Data analytics is used to visualize relationships and patterns. Genetic data mining helps in combining the real-world outcome of data with clinical data. It helps in recognising the research failures faster and design more efficient data with low costs.[3]
3. *Analysing Consumer Behaviour:* With the boom of World Wide Web, E-businesses and social media platforms have definitely revised the scene of accessible consumer behaviour data. E-commerce sites have now exchanged money registers and Point-of-Sale (POS) frameworks in a manner which enables recording of each and every move of the buyer, even when no purchase is being made. [3].

5. TYPES OF DATA

The different and clearly intricate, specialists from various fields like retail, finance, advertising and media describe data in three classes as indicated by securing focuses and proprietorship:

1. *First party:* It implies information possessed by the business, gathered about a business' clients, for instance, value-based information, statistic information gathered directly from clients by and large for business purposes, for either in-house investigation or deal to third parties. This information is shut. It isn't openly shared or uninhibitedly accessible. First party information can be viewed as a yield that may have an incentive in its very own privilege and can fill in as a reason for the restrictive upper hand. [15]
2. *Second party:* It means derived data where ownership is uncertain like data collected in collaboration with another company. This information is shared between organizations according to authoritative understanding indicating an extent of time and practices. Responsibility for information is dubious yet for the most part the organization gathering the information stores it. For instance, Google AdWords information, what number of site hits, what number of snaps, how often a specific inquiry term was entered. The nature of this data is also typically closed. [15]

3. *Third party:* It means data collected and owned by different parties. For example: government datasets, credit scores and audience ratings. This information incorporates open information just as information of a closed or business nature gathered from the web through web scratching or information gathered and prepared by information merchants. Third party information sellers bundle information as indicated by customer request, giving data on explicit objective gatherings. Third party information is a frequently expensive information. Plans of action at that point rotate around how to utilize such a contribution to make new items or improve business effectiveness in novel and imaginative ways [15].

6. BIG DATA BUSINESS MODELS

The business of big data allows many complexities as many organizations participate in various spectrums of gathering, handling, and selling of data. Rising in this market are "as a service" models in which organizations make available software, platforms, analytics, or consulting "as a service," to draw in their clients in different ways.

1. *Monetizing first party data:* In progressive firms, first party information is utilized to aid internal business choices on an incredibly fine grained scale. While the essential plan of action for these organizations is retail, first party data affects pretty much every critical choice. One clear approach to adapt restrictive first party data is to treat it like any market-worthy product and offer it to different sellers to whom this data may be directly or indirectly useful. Therefore, first party information is considered as a result or output in its own right.
2. *Data analytics as a 'Service':* The worth of data lies in the proceedings that are a result of its analysis and not in its intrinsic benefits. In this way, a typical business architecture model is made for organizations in the big data spectrum which is nothing but the provision of analytics as a service where data analytics is an integral part of the aforementioned model. Further, in this model, the analytics firm takes its own proprietary data, data supplied by its client or some third-party source of data as an input and generates data summary, analysis, insight, proposal, or some other product derived from that data, as an output. Analytics often result in reporting insights on a client's targeted audience segment based on aggregated behavioral data for particular groups. Since processing and analysis

of technology is being more affordable and the consumers have begun utilizing big data frequently, big data analytics is likewise being more accessible for individual use. Hence, a new variety of consumer facing analytics firms is emerging. For instance: Mappings, a portable application that empowers its customers to report their degrees of satisfaction and receive feedback, gathers individual personal information, examines it, and reports it in a usable structure to clients.

3. *Three distinct classes of big data business models.*
   (a) *Data users:* These are organizations that utilize data either in the process of business decisions, or as a contribution to different services and products. Such associations face problems like what data is to be generated, what data is required remotely, and by what means can this data be utilized to make an incentive inside the business? These organizations require tangible and human resources to exploit the information out of this data.
   (b) *Data suppliers:* These are organizations that either create data that is of native worth and therefore commercial or fulfill some sort of brokerage rule by providing access to an aggregation of first-party and third-party data. Such associations need not have any specialization in the supply of data as numerous associations are finding that they hold information that is of impressive worth when third parties put it to a use other than that for which it was really gathered. The key inquiries include – what data is accessible, what benefits the data might have and for whom, and by what means should data be conveyed to expand its worth?
   (c) *Data facilitators:* These organizations implement varied services including counsel on the best way to exploit big data, provision of physical architecture, and the arrangement of outsourced analytics services. These organizations have a significant job, currently, as many firms are redesigning to make data a key contributing element to their organizations.

7. BIG DATA AND SECURITY TECHNIQUES

Big data is gathered by all kinds of data and is, mostly, anonymous in nature, this process is called anonymisation. Anonymization is a privacy preserving technique that manipulates the information, make the data identification tough for anybody but the owners henceforth, making it different from that of data encryption. Anonymization of data eradicates

identifying features like social security numbers or names from the database. Since Different vulnerabilities can be linked with different types of anonymizations, there are various methods available to anonymize data, such as shuffling, encryption, substitution, date and number variance and nulling some fields. Some anonymization techniques are discussed for ambiguous data in database.

1. *Data Hiding:* It overrides a data value by swapping it with a value '0'. It is also known as Black marker anonymization. For example, while considering hospital database, an age of a patient might not be essential for processing, so it is swapped with constant '0'.
2. *Hash Calculation:* It can find a hash value either of a single field or several fields. It takes a variable input and produces fixed size hash of input. The SHA or MD5 can be used. For example, we can calculate hash of the first name and the last name.
3. *Shifting:* Shifting shifts a field or data value by specific value. It adds some offset to data value. Shift value is the only key to shift function, so that is kept secret. For example, an offset value 10 is added in age field.
4. *Data Truncation:* In this 'n' least significant bits are removed from numerical field. It conserves the info even if data at the last is lost. For example, the phone number of surgeons is truncated, and the first 4 digits are shown.
5. *Data Permutation:* It is a replacement technique. It swaps the original value with a new unique value. The selection of substitution value is random. They may result in non-collision. For example, first and last name are permuted.
6. *Data Enumeration:* It is also a replacement technique. It retains the sequential order of the events. It is beneficial for applications that require strict sequencing order. Example: salary field is enumerated along with preserving the execution order.
7. *IP Prefix-Preserving:* In this method the n-bit prefix of IP-address is preserved. Two anonymized IP addresses match on prefix of n-bits, iff two real IP addresses are same on prefix of n-bits. The IP address is prefix preserved here. Prefix preserving anonymization belongs to Typed Transformation, which uses single anonymized value for each unique value of original data. The tool TCPdPriv uses prefix preservation anonymization. Crypto PAN is an approach developed by Fan *et al.* for creating prefix preserving anonymized addresses without the use of prefix table. [8, 13].

8. TOOLS FOR ANALYSING BUSINESS

Following are certain data analytics tools that enterprises use to analyze their businesses based on their use cases:

1. *Apache Hadoop:* Hadoop tool is a collection of open-source software utilities that use a network of computers for forming racks of servers which are further used to solve problems involving big data via the MapReduce programming model. Moreover, NoSQL databases are mainstream tools for archiving data on these racks. Hadoop tasks or jobs are programmed in Java which requires a lot of complexity henceforth making it efficient yet complex to understated.[9, 10]
2. *Pentaho Business Analytics:* Pentaho is a software platform that started out as a report generating engine, in the beginning. It makes it easy to absorb information from new sources while having the facility to attach the tool to various NoSQL databases including MongoDB, Cassandra, etc. Once a connection is established, the information can be represented as if it was originally derived from SQL databases.[9, 10]
3. *Karmasphere Studio and Analyst:* Karmasphere Studio is an IDE that makes it simpler to make and run tasks in Hadoop. While setting up the work environment, the tools show the condition of test data at every single step making the client aware about how the temporary data will look like as it is sliced, analyzed and then reduced. [9, 10]
4. *Talend Open Studio:* Talend offers an Eclipse-based IDE for executing data processing tasks with Hadoop that are intended to help with data integrity, data quality and data management, all with subroutines tuned into these tasks. Talend Studio enables us to develop tasks by dragging and dropping little icons onto a canvas. It also has an added feature of adding proxy while getting a RSS, feed if required. [9, 10]
5. *Tableau Desktop and Server:* Tableau Desktop is a specialized visualization tool that simplifies viewing data in new innovative ways while slicing it. Tableau offers an interactive and dynamic mechanism so that we can slice data which makes it different from other such visualization tools.[9, 10]

9. ADDRESSING BIG DATA CHALLENGES

The following challenges apply to big data collection, cleansing and processing in the e-business domain:

1. *Guaranteeing the Authenticity and Accuracy of the acquired Data:* Today is the time of prescriptive investigation which like each other

innovation relies upon the legitimacy and precision of the information feed. One of the real challenges for most e-organizations is to comprehend which information is to be utilized for which examination like in Southeast Asia, where provincial dialects are numerous and various, making an interpretation of the gathered information into English is another hardship all alone.

2. *Detecting bots:* Due to the volume and type of traffic that they generate, Spiders and crawlers; collectively called bots which are automated programs visiting the websites including typical bots like web search engines (such as Google), site monitoring software, and price and email harvesters, are dramatically changing click-stream patterns at the website level.
3. *Lack of Talent:* Considering that Big Data is a comparatively new technology, determining the perfect fit and its successful implementation is often difficult.
4. *Security of the Data:* Some of the major Big Data security challenges include user authentication, restricted access or access constraints, meeting various compliance guidelines, appropriate encryption in data in-transit.

10. ADVANTAGES AND DISADVANTAGES OF BIG DATA

Advantages	*Disadvantages*
• Builds a better customer and customer gets an expertise opinion. • Predictive analytics results can be used to predict customer choice and his area of interest. • Big Data aids in personalising emails and increasing conversion rates. • Product Pricing can be updated in real-time so as to keep up with the competition. [10]	• This industry is limited by infrastructure specially for developing countries like India. The bandwidth available for usage restricts the type of content and services that can be accessible online. • While operational costs are minimal, setup costs are high, dedicated computer servers. • Highly talented and specialised workers are required for setup and maintenance • The inability to view products physically often means a lack of trust between buyer and seller • As this is still a comparatively new industry, regulations governing the industry are often ambiguous or simply not present. [10]

11. FUTURE BIG DATA TRENDS

The evident growth of e-business industry is resulting in companies focusing on new trends in e-business for sales to increase multi-fold by aligning e-business operations to customer expectations. Some of the trends wherein big data will prove to be a big promoter are [2, 5, 10]:

1. *Mobile Taking the Center Stage:* A responsive mobile site is the need of the hour as convenience and accessibility in terms of optimised checkout, content, site performance, content and online marketing are an integral part of the e-business technology trends leading to more visitors to the site inadvertently.
2. *Voice Search will be a Necessity:* According to estimates 50% of the searches will be voice search by 2020. The reason of which could be that people do not want to spend the time to type a query. Soon, businesses would have to focus on voice optimisation and not just text like the US-based outdoor gear and apparel brand which achieved a 24 percent increase in income by enabling voice-enabled search. Further, you could even optimise the local voice searches to cater to global markets. An essential element of this process would be the usage of structure and common language that reflects on the way people speak.
3. *Personalization Will Dominate Buying Cycles:* Personalization, another trend that would be on the surge in 2019, is based on the analysis of buying-behaviour history of buyers which helps in customer segregation and the examples of which are e-business companies like Amazon that uses user's browsing history to share recommendations which according to a recent report led to 35% increase in revenue for the Amazon recently.
4. *Payment Methods will Continue to Evolve:* This trend comprises emergence of specific payment methods and quantified categories of people like Kodak coin for photographers and musician for musicians since mobile payments, crypto currencies and digital wallets like Apple Pay, Google Pay, and Samsung Pay are leading the change in payment methodologies. Thus, not having adequate payments for your e-business portal will affect the shopping cart abandonment.
5. *Application of Augmented Reality:* A new way to display services and products can be augmented reality, that is expected to reach $ 133.78 billion by 2021 which provides a way to improve customer experience and can increase the sale profits at an extraordinary pace as the users get to explore and learn the different features of the product globally, as suggested by market research studies.

12. CONCLUSION

There's no uncertainty that Big Data has showed up as the new bleeding edge of development and rivalry in the field of e-business because of the difficulties and openings produced by the data upheaval and the outcomes that this information can create. Enormous Data really offers some incentive to e-organizations by utilizing the elements of individuals, procedures, and advancements to change information into experiences for money related ventures, basic leadership and reasonable answers for prior business issues.

This has turned into a thorough yet profitable methodology, managing information, sources, abilities, and frameworks so as to make an upper hand on the business front. Prominent e-business giants such as Amazon, Google, Facebook, and E-bay have already embraced Big Data Analytics at a time when it is most in demand.

Through the taxonomy and simplification of the key qualitative and quantitative aspects of Big Data Analytics, this study helps in presenting a valuable beginning stage for the use of Big Data in developing e-business explore, continuously. This study represents a unique way of encapsulating all the best tools and techniques like Hadoop that build and shape Big Data Analytics capabilities.

This position paper investigates Big information in e-business by drawing on a precise survey of the writing. The paper presents a system that investigates the definitional angles, unmistakable attributes, types, business worth and difficulties of Big Data in the field of e-business. Close by the above mentioned, the investigation impressions that once Big Data and its extension are all around characterized; explanatory strategies and kinds of huge information are surely known; and the difficulties relating to it are appropriately tended to, the Big Data application will amplify and show new plans of action that may change how we see commercialization in today.

REFERENCES

[1] Arya, K., Kumar, T. and Jain, M.K. (2016). Big Data Analytics of Global E-Business Organisations: A Study, Survey and Analysis, IJSER. Https://www.ijser.org/researchpaper/Big-Data-Analytics-Of-Global-E-Commerce-Organisations-A-Study-Survey-And-Analysis.pdf

[2] Avinash, B.M. and Babu, S.H. (2018). Big data technologies for e-business—future opportunities, challenges ahead and growing trends. International journal of advanced Research in Computer Science, 9(2).

[3] Bohrer, C. (2018). 5 Ways Big Data and Analytics Will Impact E-Business in 2018, Data conomyblog. Https://dataconomy.com/2018/02/5-ways-big-data-analytics-will-impact-e-business-2018

[4] Chopra, A. and Madan, S. (2015). Big Data: A Trouble or a real solution, International Journal Computer Science issues, 12(2).

[5] E-Business Trends: A Service Enterprise Engineering Perspective, pennstate college of engineering, Service Enterprise Engineering (SEE 360) Initiative.

[6] Ferreira, T., Pedrosa, I. and Bernardino, J. (2017, April). Business Intelligence for E-commerce: Survey and Research Directions. In World Conference on Information Systems and Technologies (pp. 215–225). Springer, Cham.

[7] From Big Data to Big Business (2016). Business and Decision Whitepaper Https://www.businessdecision.us/index.php/2016/06/01/big-data-white-paper-big-data-big-busine-free-whitepaper.

[8] Goswami P. and Madan S. (2017). Techniques for Enhancing Sensitive Data Security in Cloud, *International Journal of Innovations and Advancement in Computer Science*, 6(8).

Https://www.apparatus.io/The-application-of-Big-Data-for-e-business#gref

[9] Kumar M, (2014), Big Data: Why Should Enterprises Adopt It, Happiest Minds Technologies Pvt. Ltd.

Https://www.happiestminds.com/whitepapers/Big-Data-Why-should-enterprises-adopt-it.pdf

[10] Madan S. and Goswami P. (2018). An Extensive Study on Statistical Data Anonymization Algorithms, 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering, 22–25.

[11] Pavithra, B. Niranjanmurthy M., Kamal Shaker J. and Martien Sylvester Mani F. (2016). The Study of Big Data Analytics in E-Commerce, *International Journal of Advanced Research in Computer and Communication Engineering*, 5.

[12] Madan S. (2016). Big Data Security, International Journal of Computer Science and Information Technology and Security, 6(3).

[13] Sun, Z. and Wang, P.P. (2017). Big data, analytics and intelligence: An editorial perspective. *Journal of New Mathematics and Natural Computation*, 13(2), 75–81.

[14] Tripathi M.R., Jena R.K. and Elammari M.M. (2015). Advances in Computer Science and Information Technology (ACSIT), *Journal of the Krishi Sanskriti.* Http://www.krishisanskriti.org/acsit.html

[15] Verma A. (2018). Why is Big Data Analytics So Important, Whizlabs blog. Https://www.whizlabs.com/blog/big-data-analytics-importance

A MATLAB Programming Applied in Gradient Search Method to Solve Quadratic Optimization Problem

L.N. Das[1] and Sanyam Gupta[2]

ABSTRACT: *In the applied Optimisation technique subject there are certain methods mentioned in Operations Research books to solve Quadratic optimisation problem. In this book chapter we formulate an unconstrained nonlinear programming problem and determine the optimal solution using gradient search technique. Also we discuss the post optimality analysis by varying the coefficient of the objective function and the obtained optimal solution values respectively. The solution process is determined with MATLAB Program coding and the post optimality discussion examples are also mentioned in this chapter.*

Keywords: Unconstraint Nonlinear Program, Gradient Search Method, Post Optimality Analysis.

1. INTRODUCTION

Optimization problem with objective or constraint function's variables having degree 2 is named quadratic programming problem. Unconstrained Quadratic programming is a special case of unconstrained nonlinear objective functioned optimization problem having many applications in mechanical production engineering, finance and economics. There are few discussions about the solution determining procedure of the quadratic programming is discussed in the optimization technique literature.

In the linear programming problem the objective function and constraint variables are of first degree. The iterative solution process for linear programing problem is obtained by using the simplex method. However, the optimization problem having either objective function or the constraint

[1,2]Department of Applied Mathematics, Delhi Technological University, Delhi.
E-mail: [1]lndas@dce.ac.in; [2]sanyam2804@gmail.com

function variables with the degree other than 1 is named nonlinear programming problem or NLP. Since a long time the researchers are investigating an effective iterative solution procedure for solving the NLP problem. The NLP problems are either constrained or unconstrained. The Karush-Kuhn-Tucker (KKT) constraint qualification algorithm [1] is one of the method for solving NLPP having objective function as well as constraint functions. The gradient search method is one among the methods for solving unconstraint NLPP. The algorithm of gradient search method is mentioned below.

1.1 Gradient Search Method

As per the definition of the book citation [1], we review the gradient search algorithm in this chapter as following manner.

Let $f(x,y)$: $S \times S \rightarrow R$ be a continuous nonlinear function defined in $S \times S$ where S ⊆ R. Select initial solution $x' = (x,y)$ & epsilon $(\varepsilon) > 0$ for given nonlinear function $f(x,y)$ and go for the following steps:

(a) Express $f(x' + h\nabla f(x'))$ as function of h by setting
$x_j = x_j' + h\left(\frac{\partial f}{\partial x_j}\right)_{x=x'}$, for $j = 1, 2, \ldots, n$ and h is real number name step length of the iteration
And substitute this expression in$f(x,y)$.

(b) Use calculus to find $h = h^*$ that maximizes $f(x' + h\nabla f(x'))$ over $h \geq 0$.

(c) Reset the value $x' = x' + h^*\nabla f(x')$.
Then go for stopping rule i.e.
Evaluate $\nabla f(x')$ at $x = x'$ and check if
$\left|\frac{\partial f}{\partial x_j}\right| \leq \varepsilon$ for all $j = 1, 2, \ldots, n$
If so, stop with the current x' as the desired approximation of an optimal solution x^*. Otherwise perform another iteration.

In this book Chapter we have solved the NLPP mentioned in Section 2 manually, and after the discussion of MATLAB coding in Section 3 we mention the MATLAB program to solve this problem with the help of computer having window operating system and we obtained the NLPP computerised solution. In Section 4 we discussed the post optimality analysis, where the parameters range and solution values are discussed in the screenshots as the program output. Finally, in the conclusion, we summarised the chapter's discussed topics.

2. PROBLEM FORMULATION

We assume the unconstrained NLPP maximize $f(x,y)$ subject to $x,y \in R^+ \cup \{0\}$ where,

$$f(x,y) = aXY + bY - cX^2 - dY^2 \quad \text{... (1)}$$

here a,b,c,d are parameters having real values. In order to apply gradient search method to solve the unconstrained NLPP we assume f is continuous function defined on $R^+ \cup \{0\} \times R^+ \cup \{0\} \to R$ with initial solution $x = 0, y = 0$ and epsilon is 0.1. Initially we assume that $a = 1, b = 1, c = 1, d = 1$ then the optimum value of the objective function is calculated as per the gradient search algorithm. Let us solve this problem manually and the solution procedure is mentioned below:

Maximize $f(x,y) = xy + y - x^2 - y^2$, with initial values $x' = (0,0)$ and $\varepsilon = 0.1$.

$$\frac{\partial f}{\partial x} = y - 2x$$

$$\frac{\partial f}{\partial y} = x + 1 - 2y$$

Gradient $\nabla f(x') = f(0,0) = (\frac{\partial f}{\partial X}, \frac{\partial f}{\partial Y})_{(0,0)} = (0,1)$

Begin the first iteration

$$x_1 = 0 + h.0 = 0$$

$$x_2 = 0 + h.1 = h$$

$$f\big(x' + h\nabla f(x')\big) = f\big((0,0) + h(0,1)\big) = f(0,h) = h - h^2$$

Since $f(0,h^*) = \max_{h \geq 0} f(0,h) = \max_{h \geq 0}\{h - h^2\}$

And assume $\frac{d}{dh}(h - h^2) = 1 - 2h = 0$

$$\Rightarrow h^* = \frac{1}{2}$$

$$x' = (0,0) + \frac{1}{2}(0,1) = (0,\frac{1}{2})$$

For this new value of x'

Gradient $\nabla f(x') = f\left(0,\frac{1}{2}\right) = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})_{(0,\frac{1}{2})} = (\frac{1}{2}, 0)$

Begin the second iteration,

$$x_1 = 0 + h.\frac{1}{2} = \frac{h}{2}$$

$$x_2 = \frac{1}{2} + h.0 = \frac{1}{2}$$

$$f(x' + h\nabla f(x')) = f\left(\left(0, \frac{1}{2}\right) + h\left(\frac{1}{2}, 0\right)\right) = f\left(\frac{h}{2}, \frac{1}{2}\right) = \frac{1}{4}(h + 1 - h^2)$$

Since $f\left(\frac{h^*}{2}, \frac{1}{2}\right) = \max_{h \geq 0} f(\frac{h}{2}, \frac{1}{2}) = \max_{h \geq 0}\{\frac{1}{4}(h + 1 - h^2)\}$

And assume $\frac{d}{dh}\left(\frac{1}{4}(h + 1 - h^2)\right) = \frac{1}{4}(1 - 2h) = 0$ to determine the maximum value of h,

$$\Rightarrow h^* = \frac{1}{2}$$

$$x' = \left(0, \frac{1}{2}\right) + \frac{1}{2}\left(\frac{1}{2}, 0\right) = (\frac{1}{4}, \frac{1}{2})$$

Similarly we can do further iterations those are summarized in the following table.

Table 1

Iteration	x'	$\nabla f(x')$	$x' + h\nabla f(x')$	$f\left(x' + h\nabla f(x')\right)$	h^*	$x' + h^*\nabla f(x')$
1	(0, 0)	$(0, 1)$	$(0, h)$	$h - h^2$	$\frac{1}{2}$	$\left(0, \frac{1}{2}\right)$
2	$\left(0, \frac{1}{2}\right)$	$\left(\frac{1}{2}, 0\right)$	$\left(\frac{h}{2}, \frac{1}{2}\right)$	$\frac{1}{4}(h + 1 - h^2)$	$\frac{1}{2}$	$\left(\frac{1}{4}, \frac{1}{2}\right)$
3	$\left(\frac{1}{4}, \frac{1}{2}\right)$	$\left(0, \frac{1}{4}\right)$	$\left(\frac{1}{4}, \frac{1}{2} + \frac{h}{4}\right)$	$\frac{1}{16}(5 + h - h^2)$	$\frac{1}{2}$	$\left(\frac{1}{4}, \frac{5}{8}\right)$
4	$\left(\frac{1}{4}, \frac{5}{8}\right)$	$\left(\frac{1}{8}, 0\right)$	$\left(\frac{1}{4} + \frac{h}{8}, \frac{5}{8}\right)$	$\frac{1}{64}(21 + h - h^2)$	$\frac{1}{2}$	$\left(\frac{5}{16}, \frac{5}{8}\right)$
5	$\left(\frac{5}{16}, \frac{5}{8}\right)$	$\left(0, \frac{1}{16}\right)$	–	–	–	–

As $\nabla f(x') = \left(0, \frac{1}{16}\right) = (0, 0.0625)$

i.e. $|\frac{\partial f}{\partial x}|_{(0, \frac{1}{16})} = 0 < .01 = \varepsilon$

and $|\frac{\partial f}{\partial y}|_{(0, \frac{1}{16})} = 0.0625 < .01 = \varepsilon$

Therefore point of maximum is $\left(\frac{5}{16}, \frac{5}{8}\right) = (0.3125, 0.625)$ and hence the objective function maximum value is 0.

3. A MATLAB PROGRAM FOR SOLVINGUNCONSTRAINED NLPP USING GRADIENT SEARCH METHOD

We fit the problem solution algorithm in the MATLAB program to get the optimal solution.

```
% Gradient Search Algorithm:
% Written by Sanyam Gupta
clc
clear
format long
% Define the Variables
syms XY;
% Enter the Function
f = X × Y + Y – X^2–Y^2;
% Initial Values:
x(1) = 0;
y(1) = 0;
% Ephisilon Value
e = .1;
% Numbering of Iterations
i = 1;
% Compute the Gradient
df_dx = diff(f, X); % Derivative of f with respect to X
df_dy = diff(f, Y); % Derivative of f with respect to Y
% Find the Gradient
G = [subs(df_dx, [X, Y], [x(1), y(1)]) subs(df_dy, [X, Y], [x(1), y(1)])]
% Search Direction
D = –(G)
% Maximize Condition:
while norm(G) > e
 I = [x(i), y(i)]'
```

```
% Define Step size
 syms h;
 t = subs(f, [X,Y], [x(i) + D(1) × h, y(i) + h × D(2)]);
 dg_dh = diff(t, h); % Derivative of g with respect to h
 h = solve(dg_dh, h); % Optimal Step Length
 x(i + 1) = I(1) + h × D(1); % Redefine Value of x
 y(i + 1) = I(2) + h × D(2); % Redefine Value of y
 i = i + 1; % Redifine Iteration
 G = [subs(df_dx [X, Y], [x(i), y(i)]) subs(df_dy, [X, Y], [x(i), y(i)])]; % Redefine Gradient
 D = –(G); % New Search Direction
end

% Outputs:
fprintf('Initial Objective Function Value is %d\n\n', subs(f, [X, Y], [x(1), y(1)]));
if (norm(G) < e)
 fprintf('Result : Maximum succesfully obtained\n\n');
end

fprintf('Number of Iterations for Convergence are %d\n\n', i);
fprintf('Point of Maximum is [%d, %d]\n\n', x(i), y(i));
fprintf('Objective Function Maximum Value Post-Optimization is %d\n\n', subs(f, [X, Y], [x(i), y(i)]));
% Result in Table Form:
Iter = 1:i;
X_coordinate = x';
Y_coordinate = y';
Iterations = Iter';
T = table(Iterations, X_coordinate, Y_coordinate);
% Plots:
% fcontour(f, 'Fill', 'On');
hold on;
grid on;
plot(x, y,'×-r');
disp(T);
```

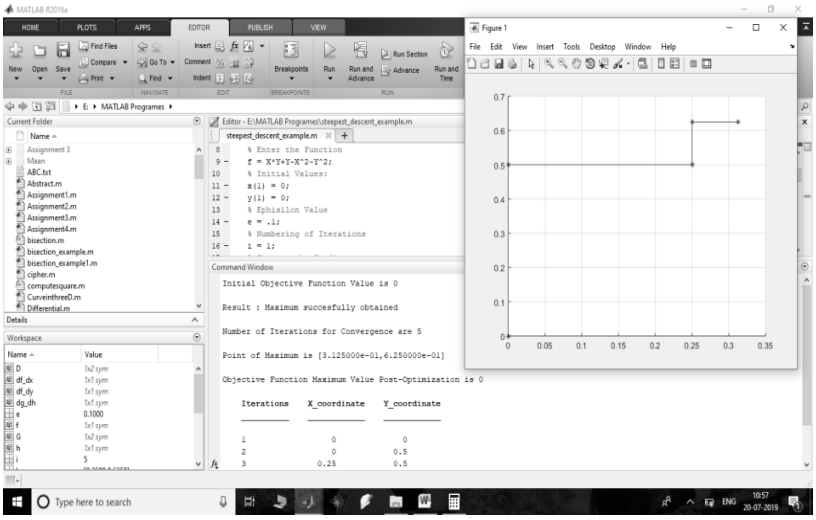


Fig. 1: Output when $a = 1, b = 1, c = 1, d = 1$

From the above screenshot we see outcome of the MATLAB program is $x = 3.125000e - 01$ and $y = 6.250000e - 01$ for the coefficients value $a = 1, b = 1, c = 1, d = 1$ with initial solution$x = 0$ & $y = 0$. The objective function value is 0. In the next section we study the range of a, b, c & d in different cases with the initial solution $x = 0, y = 0$ and epsilon $= 0.1$.

4. POST OPTIMALITY

In the above section we considered the function variables coefficients as $a = 1, b = 1, c = 1, d = 1$ for the optimization problem objective function defined in (1).

Case 1: Let us increase the variable coefficient a and keep $b = 1, c = 1, d = 1$ then up to $a = 1.42$ the result optimum value to the objective function is 0. Hence there is no change in the objective function value. At $a = 1.43, b = 1, c = 1, d = 1$ the objective function value is 1. In this manner, if we go on increasing value of $a = 1.82, b = 1, c = 1, d = 1$ the value of objective function is 1 for initial solution $(x = 0, y = 0)$ with value of epsilon 0.1 using the gradient search algorithm. If we set $a > 1.82, b = 1, c = 1, d = 1$, the objective function value changes. So the range of the coefficient a is (1.43, 1.82) for objective function value 1 with consideration $b = 1, c = 1, d = 1$ and the initial solution $(x = 0, y = 0)$ & value of epsilon 0.1.

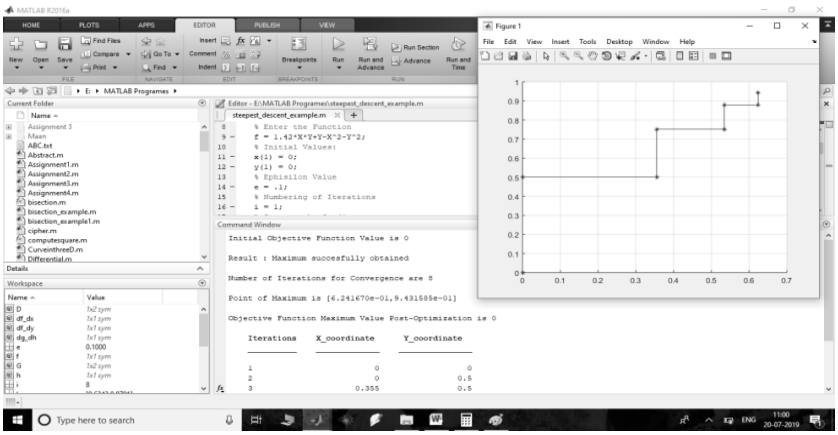


Fig. 2: Output when $a = 1.42, b = 1, c = 1, d = 1$

The above screenshot shows the optimal solution $x = 6.241670e - 01, y = 9.431585e - 01$ for the coefficients $a = 1.42, b = 1, c = 1$ & $d = 1$ with initial value $x = 0, y = 0$ and epsilon 0.1. The objective function value is 0.

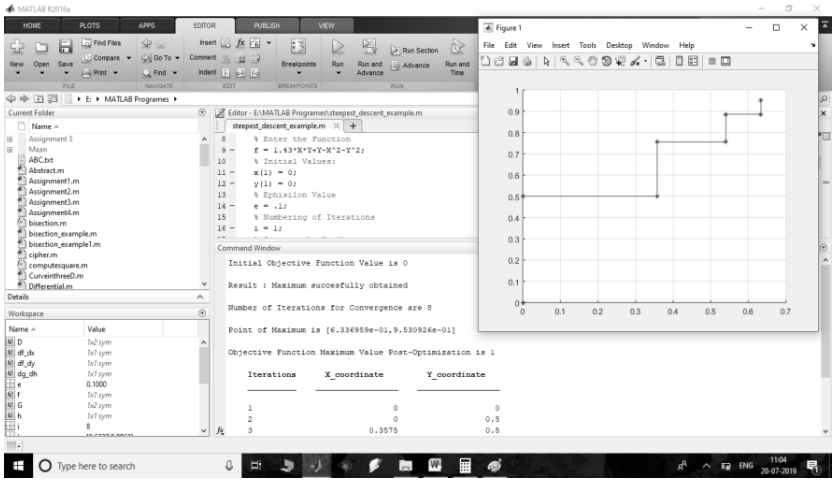


Fig. 3: Output when $a = 1.43, b = 1, c = 1, d = 1$

The above screenshot shows the optimal solution $x = 6.336959e - 01, y = 9.530926e - 01$ for the coefficients $a = 1.43, b = 1, c = 1$ & $d = 1$ with initial value $x = 0, y = 0$ and epsilon 0.1. The objective function value is 1.

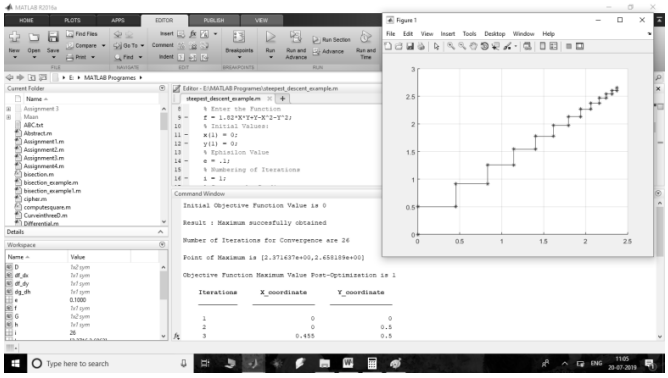


Fig. 4: Output when $a = 1.82, b = 1, c = 1, d = 1$

In the above screenshot the gradient search algorithm plotted graph has an ascent direction and explains the convergence of the solution search algorithm.

Case 2: Let us assume $a = 1, c = 1, d = 1$ and we have determine from the MATLAB program that the range of b is $(1.23, 2.12)$ for objective function value 1 with initial solution $(x = 0, y = 0)$ and value of epsilon 0.1.

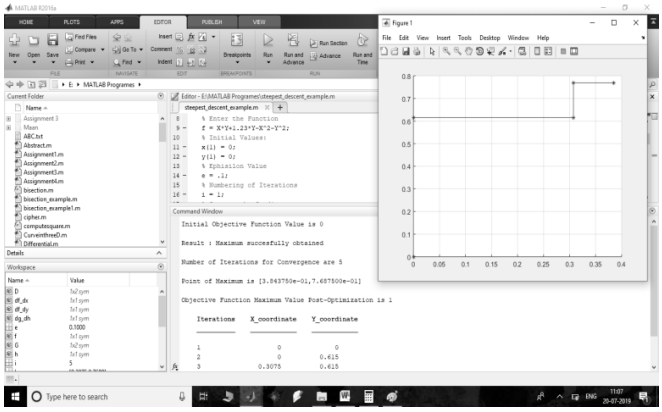


Fig. 5: Output when $a = 1, b = 1.23, c = 1, d = 1$

The screenshot explains the objective function value is 1 for $a = 1, b = 1.23, c = 1$ & $d = 1$with $x = 0, y = 0$ and epsilon 0.1. Further we observe $b = 1$ up to $b = 1.22$ the objective function value is 0 with the increment of 0.01 to the b value causes the value of the objective function different.

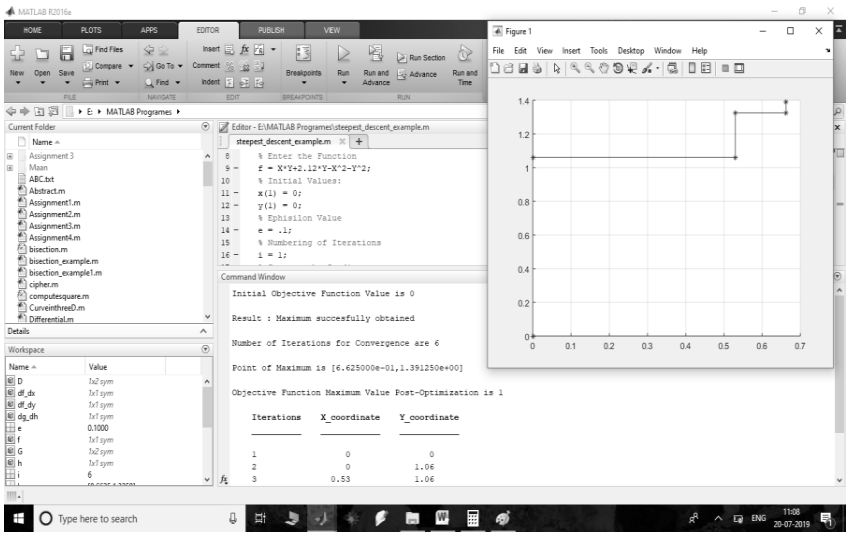


Fig. 6: Output when $a = 1, b = 2.12\ c = 1, d = 1$

In this screenshot we observe the objective function value is 1 and remains unchanged for $a = 1$, b belongs to (1.23, 2.12), $c = 1, d = 1$ for the initial solution $x = 0, y = 0$ and epsilon 0.1.

Case 3: When the parameter $\boldsymbol{a = 1.43}$ the range of $\boldsymbol{b}$ is within (1, 1.71) for $\boldsymbol{c = 1, d = 1}$, and objective function value is 1 with $\boldsymbol{x = 0, y = 0}$ and epsilon is 0.1. The screenshots are mentioned below

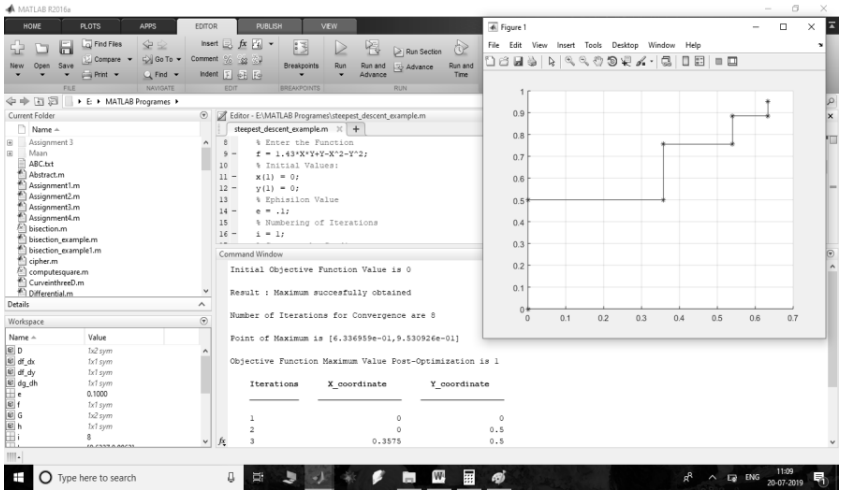


Fig. 7: Output when $a = 1.43, b = 1\ c = 1, d = 1$

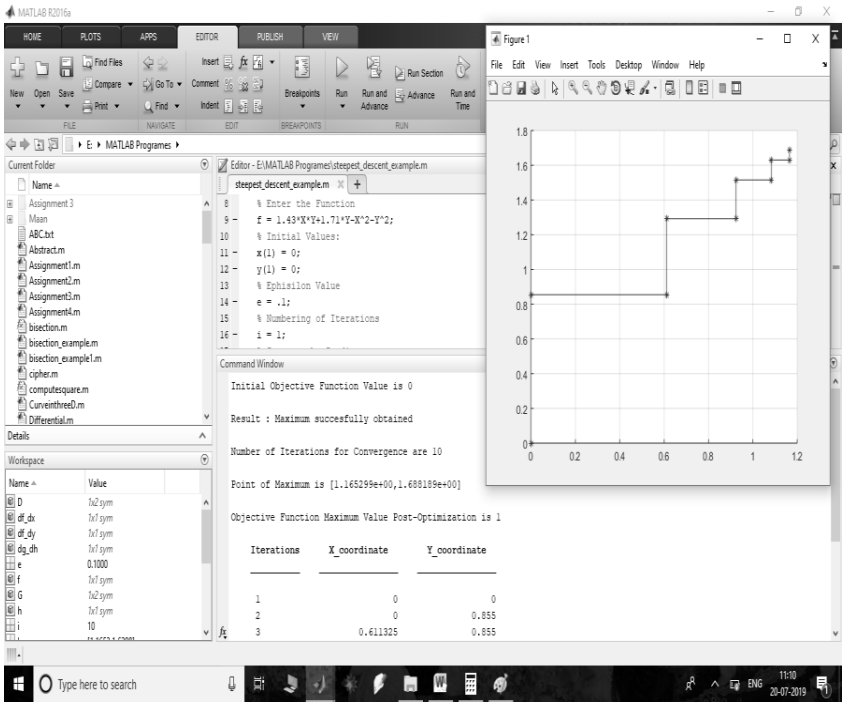


Fig. 8: Output when $a = 1.43, b = 1.71\ c = 1, d = 1$

Case 4: In this post optimality case, we observe if we change the value of c or d other than 1, and values of $a = 1, b = 1$ the objective functional value remain 0. If we assume $a = 1.43, b = 1.71$ and value of c or d is other than 1, the objective function value is 0 for initial solution $x = 0, y = 0$ and epsilon is 0.1.

5. CONCLUSION

The objective function mentioned in equation (1) is a general expression with coefficients a, b, c, d, where a, b, c, d are definite real numbers. We have observed the objective function value and optimal solution of the NLP problem are different, if we set suitable values to the parameter for obtaining the optimal solution by using gradient search algorithm with the initial solution $x = 0, y = 0$ and epsilon 0.1. Further if we assume x and y values are other than $(0, 0)$ the objective function value will be same but the gradient changes. For instance, if we assume $x = 1, y = 1, a = 1.43, b = 1, c = 1, d = 1$, then the optimal function value remains 1. This screenshot is mentioned below

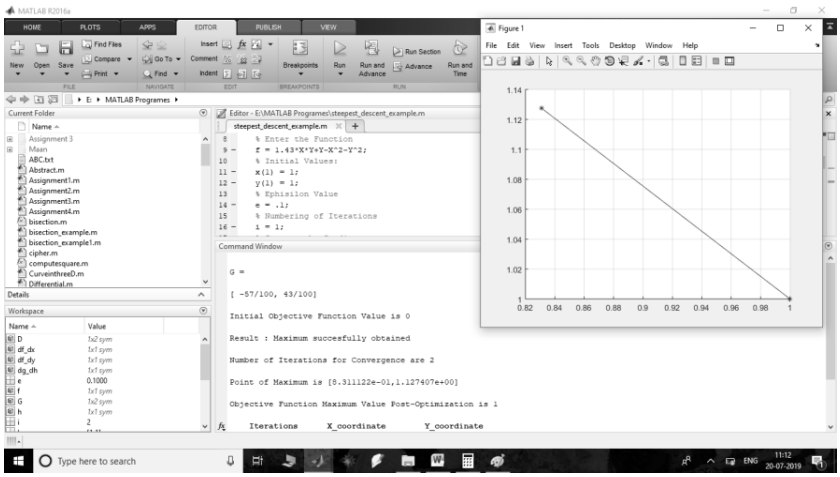


Fig. 9: Output when $x = 1, y = 1, a = 1.43, b = 1, c = 1, d = 1$

REFERENCES

[1] Hiller, F.S. and Liberman, G.J. (2017). *Introduction to Operations Research*, New York, McGraw-Hill.

[2] Chandra, S. and Jayadeva, Mehra, A. (2011). *Numerical Optimization with Applications*, Delhi, Narosa.

[3] Sharma, S. (2017). *Applied Nonlinear Programming*, Delhi, New Age International Private Limited.

[4] Shetty, C.M., Bazaraa, M.S. and Sherali, H.D. (2017). *Nonlinear Programming: Theory and Algorithms,* Delhi, Wiley India.

Postmodern Techno Culture and the Need for Feminist Intervention: Study of Mobile Applications for Women Safety in India

Rubina[*] and Madhav Govind[#]

ABSTRACT: *Women safety is a pressing issue especially in India and lately various innovators are trying to come up with ICT based solutions to empower women so that they can fight against this menace. Use of mobile phone-based applications for identifying rural and urban landscape on various safety parameters intended for women safety is one such digital system innovation. Though our previous research on similar mobile phone application "My Safetipin" shows that rather than helping the cause of women safety the mobile application was jeopardizing women safety in India. Feminist intervention in post-modern techno culture claims that more female participation not only as user of technological artifact but as decision maker at every stage of decision-making process can make the product better for females which in turns improve the society. Thus, this research tries to explore our previous case study of "My Safetipin" mobile application from a fresh perspective of analyzing female participation not from user perspective but as decision makers perspective to look at possible reasons for failure of this e-participation platform.*

Keywords: Feminist Technological Studies, Technological Determinism, Women Safety, Digital Exclusion, ICT.

1. INTRODUCTION

There exists missing perspective of technology into women's studies and at the same time theme of gender takes back seat in postmodern

[*] Research Scholar, Centre for Studies in Science Policy, JNU
E-mal: rubinaakhtar1612@gmail.com
[#] Professor, Centre for Studies in Science Policy, JNU.
E-mail: m_govind@mail.jnu.ac.in

techno-culture.[1] Techno culture tries to answer how technology pervade culture, both in terms of popular culture and gendered applications. Today we have the mobile phone, i.e., since it became a common occurrence, it is now the primary form of contact for most people, although originally only usable to make phone calls. It has become a vessel for many other parts of technology. Recent research on technology states the difference between 'Objective usability' and 'User experience', that is, what a computer can do and what the user can do with a computer, respectively. Use of android and I phone give user an amalgamation of features, combining many pieces of tech into one, and is designed for accessibility and ease of use, it becomes effectively part of both 'material' and popular culture in one swoop. Now 'digital' programs and applications have become a standard and recognizable medium to create work on par with traditional techniques.

The archetype thought is that women's use of technology is mostly viewed as dystopic. The emergent techno-feminism theory[2] defends a relationship in which technology is cause and the consequence of gender relations, at the same time. For years, feminists initiated a critique of the inherently patriarchal nature of technology and of technoscience. Socialist feminist tried to be more contextual in their work, pushing Marxist analysis beyond class to ask why and how modern technology became a male domain and to address the gendering impact of modern divisions of labor.[3] Even in a postmodern society, interaction between human and machines are studied as "man-machine system" making the whole technological academic discourse as a masculine discourse.[4]

Studies in past have showed that most of the problem that exist with technological solutions centered around women in India, are rooted in the fact that there is lack of involvement of female in the whole technology development cycle.[5] Even other previous academic work on gender politics raised concern regarding lack of female stakeholder presence in decision making process as far as technology implementation in society is concerned. It showed that technological determinism is not only non-neutral but at various instances forces female to bear most of the social

[1] Berg & Merete, 1995.
[2] Wajcman, 2004.
[3] Bray, 2012.
[4] Cockburn & Fiirst-Dilic, 1994.
[5] Manazir *et al*, 2019.

cost of technological advancement which leads to their own social oppression.[6] Since women safety is a pressing issue in India, we take technological solution in form of mobile phone applications for women safety as our theme of study. Taking further our previous study, for a robust policy level solution which can bring in feminist intervention for not only technological advancement but for social emancipation of society, without excluding females. Through this research we want to study the condition of female engagement at various stages of technological solution development cycle, barriers if any available in female intervention in decision making as well as possible solution beyond the technological determinism.

2. WHAT IS 'MY SAFETIPIN' MOBILE APPLICATION?

'My Safetipin' is a mobile application which was released on 9th March 2016 in India available on Android and Apple's Internetwork Operating System (iOS). It is basically meant to be a personal location guide of public places using Global Positioning System for the safety of females. It helps its user in taking a safer decision based on the safety score of an area thus notifying user how safe or unsafe any location is, to roam around. Also, user can notify their family and friends in case they are in an unsafe location and helps others to keep track of user. The application user can also rate an area based upon the nine parameters based upon their own understanding of these parameters thus scores are cumulative and could vary from user to user. In India, this mobile application is operational in 9 cities namely New Delhi, Gurgaon, Noida, Mumbai, Bengaluru, Chennai, Chandigarh, Guwahati, Karnal in English and Hindi languages. Thus "My Safetipin" currently is available to only 2.64% female population of India including those of rural and urban areas both.[7]

2.1 Research Question

1. What is the extent of women participation at conceptualization and development stages of technological solutions for women safety in form of mobile phone applications?

---

[6] Harding, 1995. [to curb unwanted pregnancy distribution of contraceptive devices for female as condition of aid by donor agency was favored, while neither male contraceptive nor women's education and social welfare was improved.]

[7] Manazir, Govind, & Rubina, 2019.

2. How does barrier at each developmental stage relate to the overall extent of women's participation with respect to such mobile phone-based applications?
3. How can female participation at various stages improve the development and implementation of women safety applications in India?

2.1.1 *Literature Review*

As per Marxian concept, it was accepted without much criticism that technological determinism is the most important influencer of cultural, political and economic aspects of any society.[8] Though it is true that technological determinism helps in development of technological artifacts but very less emphasis is given upon the fact that technology is not the sole factor of changes in society, but one of the many factors which should be seen in totality.[9]

Further women suffer with Digital Exclusion in India, where even the use of internet and mobile phones are majorly male dominated in both rural as well as urban demographics.[10] Female safety is an important issue today and even Sustainable Development Goals - 5 emphasizes "Women Equality and Empowerment", and mobile phones have emerged as a positive contributor towards women empowerment.[11] With advent of Information, Communication and Technology various forms of mobile phone applications have come up worldwide for women safety, and in India too they have large user base. These mobile phones based e-participation platforms aims at using available technologies to support active citizenship, enabling broader participation for reaching a wider audience and enabling more informed citizens' contributions through providing accessible and understandable information to target audience.[12] Technically e-participation platforms like "My Safetipin" mobile phone applications are nothing but forms of Digital System Innovation using advancement in technologies like GIS mapping, advanced internet connectivity technologies like 4G etc. But as far as technological solution's in form of e-participation platform for women safety are concerned, first we need to understand different stages where feminist intervention in decision making is not up to the mark.

---

8. Marx, 1959.
9. Berg & Merete, 1995.
10. IAMAI, 2017.
11. UN, 2018.
12. Manazir, Govind, & Rubina, 2019.

2.1.2 *Research Findings*

Epistemologically this study will be empiricist form of enquiry where we will follow mixed method including primary and secondary data research followed by case study of "My Safetipin" women safety application being used in India. Taking further our previous research on e-participation platform namely "My Safetipin"[13] where we analyzed the e-participation platform which claimed to resolve various issues related to women safety. We found that rather than resolving women safety concerns in rural and urban landscape mobile interface was not only not designed properly for the functions which it claimed to make moreover it was counterproductive to female users in case of distress and urgency.[14]

To understand level of involvement and skillset of all employees especially females in development of "My Safetipin" and taking leads from previous research which states more involvement of stakeholders like females in Digital System Innovation meant for women safety leads to better product[15] we conducted analysis of people involved with the product development as well as management and maintenance.

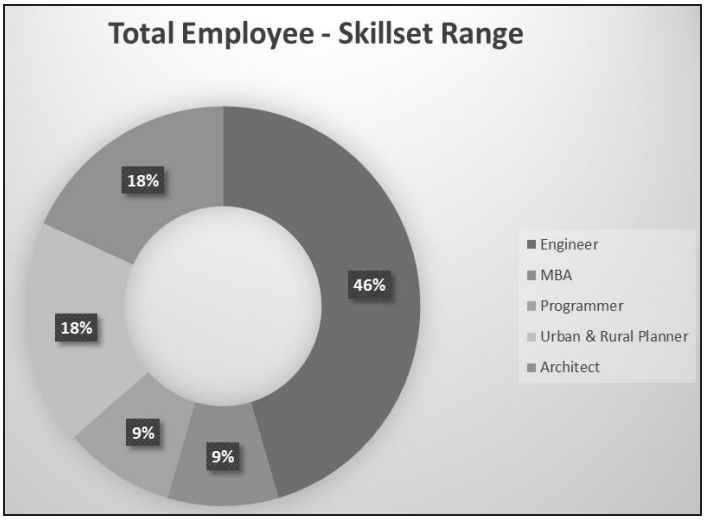


Fig. 1: Total Employee Differentiated based on their Skillset[16]

[13.] Manazir, Govind, & Rubina, 2019.
[14.] SafetiPin, 2019.
[15.] Vissers & Stolle, 2013.
[16.] SafetiPin, 2019.

We found that among total employees working with the organization in India, almost half of the total strength were engineers, followed by architect, urban and rural planners, MBA's and lastly programmers. Worth mentioning percentage of programmers and product development expert were least.

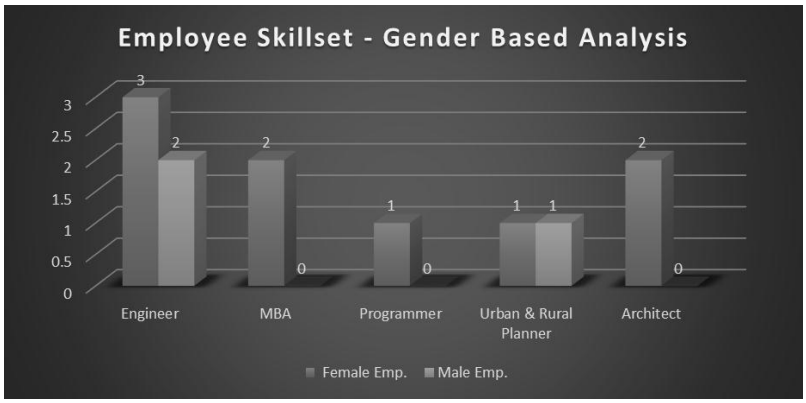


Fig. 2: Variation in Employees Skillset based on Gender[17]

Out of the total employee strength, females were most in every department and were in majority as far as Digital System Innovation development is concerned.

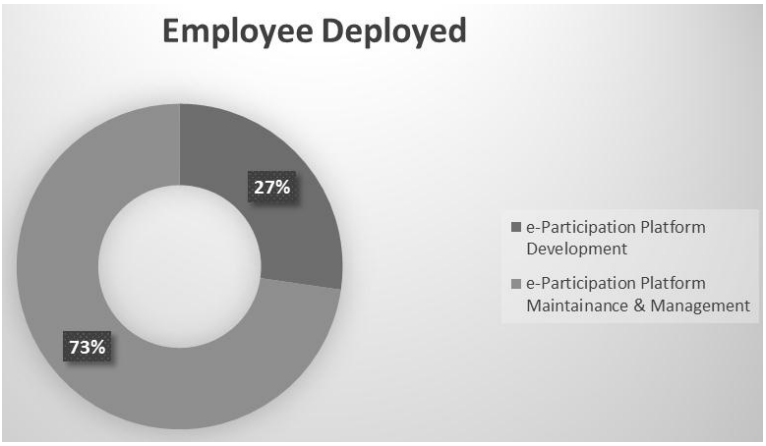


Fig. 3: Variation in Total Strength of Employees Deployed based on Major Roles Assigned

[17.] SafetiPin, 2019.

Also, while almost 3/4th of the employee strength was focused over e-Participation Platform maintenance and management only 1/4th of the total employee strength was meant for platform development and upgradation.

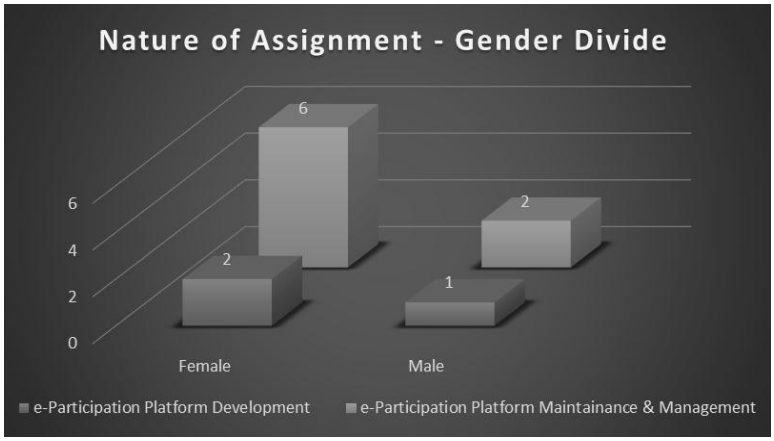


Fig. 4: Variations in Major Roles Assigned to Employees based on Gender[18]

On closer analysis we found that only 2 female staff are designated full time for e-participation platform development i.e. from Digital System Architecture development to final product launch along with only one male employee meant for managing GIS mapping, while remaining majority staff was meant for e-participation platform maintenance and management.

Thus, our two important findings were as follows:

1. Though majority of females were part of the organization as well as part of e-participation platform development team. Yet, there total strength especially of development was comparatively very low in comparison with the large female user base they want to serve to.
2. Majority of issues associated with the failure of functionality of e-participation platform is associated with poor Digital System Architecture designing process which shows all the stakeholders were not considered.
3. Digital Exclusion of females in India has not been considered which is visible by the fact that the mobile application has various issues and even after so much time of product launch there has been no upgradation in product glitches.

[18] SafetiPin, 2019.

3. CONCLUSION

This research aims to open academic debate regarding women participation in Science Technology and Society not only as a reader or observer but as a participant in decision making who can influence technology and in turn society for more inclusive and developed social system. Thus, our findings reveal that though it is true that female intervention is important in post-modern techno culture, but more than that involvement of females in decision making process i.e. Digital System Architecture Designing in case of e-participation platforms like 'My Safetipin' is very important. Also, corporations promoting such innovation should hire more and more females in decision making and product development roles.

REFERENCES

[1] Berg, A.-J. and Merete, L. (1995). Feminism and Constructivism: Do Artifacts Have Gender? *Science, Technology, & Human Values*, 332–351.

[2] Bray, F. (2012). Gender and Technology. In N. Kumar, *Gender and Science : Studies across Cultures* (pp. 37-60). New Delhi: Cambridge University Press India Pvt. Ltd.

[3] Cockburn, C. and Fiirst-Dilic, R. (1994). *Bringing technology home: Gender and rechnology in a changing Europe.* Milton Keynes, England: Open University Press.

[4] Harding, S. (1995). Just Add Women and Stir? In UN, *Missing Links - Gender Equity in Science and Technology for Development* (pp. 295-307). New York: United Nations Commission on Science and Technology Development.

[5] IAMAI (2017). *Internet in India* . New Delhi: KANTAR -IMRB.

[6] Manazir, Sharique. H., Govind, M. and Rubina. (2019). My Safetipin Mobile Phone Application: Case Study of E-participation Platform for Women Safety in India. *Journal of Scientometric Studies*, 47–53.

[7] Marx, K. (1959). *Das Kapital :A critique of political economy Vol (I-III).* Chicago: H. Regnery.

[8] NCRB (2016). *Crime in India – 2016.* New Delhi: National Crime Record Bureau.

[9] SafetiPin (2019, 14 08). Safetipin supporting safer cities . *About Safetipin; our team*, p. 1. Retrieved from http://safetipin.com/about/ourTeam

[10] UN (2018). *Turning Promises into Action: Gender Equality in the 2030 Agenda for Sustainable.* Milan: UN Women.

[11] Vissers, S. and Stolle, D. (2013). The Internet and new modes of political participation: Online versus offline participation. *Information, Communication and Society*, 937–955.

[12] Wajcman, J. (2004). *TechnoFeminist.* Cambridge : Polity Press.

[13] Winner, L. (1980). Do Artifacts have Politics? *Daedalus*, 121–136.

Recommender System: Principles, Techniques and Application

Sushma Malik[1], Yukta Pandey[2] and Amanpreet Kaur[2]

ABSTRACT: *In today's internet era, human beings are fully dependent on the internet for their daily tasks. Internet is used to search everything and anything meaningful that the user requires. With the growth of internet, a number of business opportunities are created and are made available to the users easily. 24×7 online services of internet have generated large amount of choices, but users are now faced with the problem of information overload. To alleviate the problem of information overload, internet needs the filtering system that could help the user to discover the information or product which they search. Such kind of filtering systems are called the Recommender Systems. This kind of system acts as software that reduces the user's choices and provides them the most appropriate options as per user's requirements. Do you know how Amazon closely knows what kind of products the user need and display a number of options for the selection? Or how can Netflix know what kind of movies or series users are interested in watching? Or how does Facebook give the options of "people you may know". All of this objective can be achieved through the Recommender System that collect the user information, analyze it and produce an immersive user experience. We can say that the Recommender System helps the user in selection process which becomes easier and time saving. The major function of Recommender System is choosing the most appropriate recommendation for the user from abundant available options. A lot of extensive research is happening in this field which is tried to give more precise and relevant recommendations to a user. Recommender systems are employing in a variety of domains, and are most commonly used to generate the playlist for video and music for various sites like Netflix, YouTube and Spotify. Various E-Commerce sites like Amazon, Flipkart, and Paytm Mall etc are using the product recommenders system to recommend the products to users on their*

[1] Assistant Professor (IT), Institute of Innovation in Technology and Management, G.G.S.I.P. University, New Delhi, India.

[2] Students, BCA 2nd Year, Institute of Innovation in Technology and Management, G.G.S.I.P. University, New Delhi, India.

interest. Recommender System is also used for the content recommenders for social media platforms like Facebook and Twitter. This chapter gives review on some significant works in the field of Recommender System on the basis of its three techniques namely collaborative filtering, content based filtering and hybrid filtering. The present chapter also includes the various issues and application domains of Recommender System.

Keyword: Recommender System (RS), Collaborative Filtering (CF), Content based Filtering (CB), Cold Start Problem, E-governs, E-commerce.

1. INTRODUCTION

The World Wide Web or simply the internet has brought numerous changes in the way we are live and communicate with each other. The explosive growth in the amount of available digital information and the numbers of visitors to the internet have created potential challenges of information overload on the users who is searching the relevant and reliable information on the web. This situation increased the demand of Recommender systems more than ever before. RS has the ability to predict whether a particular user would prefer an item or not on the basis of user profile [1]. RS can be defined as a program which is used to recommend the most appropriate items (products or services) to the particular users by predicting a user's interest in an item based on related information about the items, the users and also the interaction between items and users.

The main objective of the RS is to minimize the information overload by retrieving the most relevant information and items and services from the huge amount of data and providing the personalized services to the users. The most important feature of RS is its ability to "guess" a use's preferences and interests by analyzing the behavior of that user and /or the behavior of other users to generate personalized recommendations [2].

2. PHASES OF RECOMMENDATION PROCESS

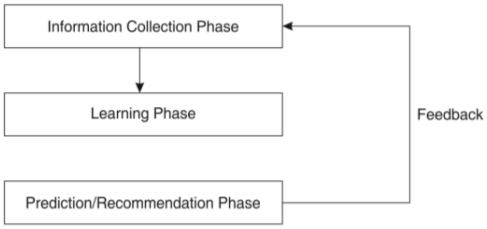


Fig. 1: Phases of Recommendation Process

2.1 Information Collection Phase

In this phase, relevant information of users are collect to create the user profile which include the user's attributes or content of the resources that user access on the web. For a good recommendation, user profile has been well designed. RS using the different types of inputs like explicit, implicit and hybrid feedback. The most convenient and high quality data collection method is explicit input method because in this data provided by the users regarding their interest in items or services. In implicit method, data collected indirectly through user behavior on the web and in hybrid method, data can be obtained through the combination of both explicit and implicit method.

2.2 Learning phase

In this phase, the learning algorithm is applied to filter and apply the user's features on the collected data from the information collection phase.

2.3 Prediction/Recommendation Phase

This phase recommends or predicts the items that user may prefer. This can be done either directly based on the dataset which are collected in information phase or through the system's observed activities of the user [1].

3. ISSUES IN RECOMMENDATION SYSTEM

This section is include the various issues related to the RS like:

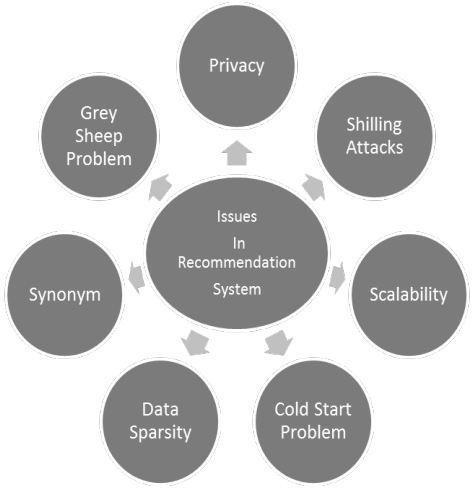


Fig. 2: Issues in Recommendation System

Cold Start Problem: The cold star problem arise when the new products are added into the catalogue or new user enter into the system. So the main reason for it is too little rating data is available in the initial state of the RS because neither new user taste can be predicted before its purchase nor the new items can be rated [2, 5, 7].

Synonym: This problem arise when same item is represented with more different names or having similar meanings. At that time RS cannot easily identify that the term symbolize different item or the same item. For example RS will treat "movie" and "film" differently. The excessive usage of synonym terms decrease the performance of the RS [7].

Shilling Attacks: what happens if number of users and competitor giving the false ratings to some products either to increase its popularity or to decrease into the RS? Such kind of situation break the user trust as well decreases the quality and performance of the RS [5].

Privacy: For the better recommendation to the user, RS need the personal information of the user but it may lead to issues of user data privacy and security. Users are unwillingly fill the personal data into the RS that suffers from the data privacy issues. Some time user feed the wrong data into the RS for that reason which affect the performance of the RS [5, 7].

Scalability: In many recommender techniques the problem of scalabilty arises when the number of users of items increases tremendously. Computation normally grows linearly with the number of users and items. Hence, the techniques which takes limited number of dataset fails here. A techniques that can increase the dataset as per the requirement is needed in a recommender system for it to perform efficiently [6, 7].

Grey Sheep Problem: Grey Sheep problem occur with people who do not completely agree or disagree with a group of people. This is some weird customer whose choices are out of the line. A user for instance may like first season of some web series but dislike the other seasons. Recommendation to such customers is difficult because their likes or dislikes are not certain [6].

Diversity: Primary factor on which most of the recommender system work upon is that similar user is likely to have similar taste. Many of the recommender system allows us to connect with social media and suggests us the shows or products that are brought by our friends on the social media platform or the similar type of products we have viewed before. It is not always necessary that we want to view the same kind of items that we have viewed before. This makes the recommendations monotonous.

Diversity is also important as say, a user has already brought TV set but now he does not want to buy that item again but many recommender system will keep showing the user different TV sets because he had bought it once [6].

Data Sparsity: This problem arises because the users generally rate only a limited number of products. There are thousands of products available most of which are generally not used by the users and hence, they are not rated by the users. Moreover, there are also users who buys a product but do not rate them or gives feedback. For instance, take your own example how many times have you watched a movie on an online platform and rated it? Or how many times have you brought a product online and rated it? There are rarely people who rate every product they see or buy. This leads to sparse user item matrix, inability to locate successful neighbours, and hence generation of weaker recommendation [5, 6].

4. RECOMMENDATION TECHNIQUES

The broad classifications of RS are including traditional methods such as collaborative filtering-based, content-based, knowledge-based, and hybrid methods and recently developed advanced methods, such as fuzzy set-based, social network-based, trust-based, context awareness-based, and group recommendation.

4.1 Collaborative Filtering

Due to human psychology, a user taking suggestions about something they own from their friends and family members which helps the user to take a decision. CF based technique help the user to take decision basede on the views of other users who have similar interest [6]. This technique can be divided into two categories like: user-based and item-based.

In the user-based CF approach, a user will obtain recommendations of items based on interest of the users having the similar taste. In the item-based CF approach, a user will receive recommendations of items that are similar to those they have loved in the past [1].

4.2 Content based Filtering

This technique using items features and interest of user to offer recommendations. Item features like keywords are used to describe items and user profile designate the items liked by the user. The recommendation engine, recommends the items that are similar with the items preferred in the past by the user [1, 6].

4.3 Hybrid Filtering

This technique is basically used to overcome the drabacks of traditional recommendation techniques like CF and CB techniques and to achive the higher performance. For that it combines the best features of two or more recommendation techniques. The most commonly used recommendation techniques for hybrid recommendation technique is CB and CF [6].

There are different possivble ways to combine the CB and CF for hybrid technique like:

1. CF and CB techniques are executed separately and after that combine their result.
2. Using the properties of CB in CF technique and vice a versa.
3. Design a combination model which combines both CB and CF features [4].

5. APPLICATION DOMAINS OF RECOMMENDER SYSTEMS

1. **e-Government:** e-Government initiative have been launched by almost all the governments across the globe. e-Government uses innovative system via information and communication technology to serve it's citizens by enhancing the management of the public sector. e-Government increases the transparency and accountability of government towards citizens. The major problem that hinders the services of e-Government is Information Overloading. Recommender System is an effective solution to this problem. It gives user the specific information and services they need rather than whole set of data [8].
   (a) **Government to citizen Services Recommendation:** It is provide citizens the access to information easily and quickly. It basically focuses on interaction between government and citizens. It enables citizens the services like tax payment, license registration, and other government services. It also gives citizens transparent information about various policies, government laws, and voting information helping an individual to make a decision.
   (b) **Government to Business Services Recommendation:** It is provides interaction between government and business enterprises, including profit and non-profit organizations. It enables the organizations to make contract bids, policies, and rules and regulations. Such services are useful for both the government and the business enterprises providing them with more business

benefits. It gives the organizations the opportunity to work with government with cost saving deals and effective results.

(c) **Government to Government Services Recommendation:** It focuses on interaction of government with agencies related to them to work efficiently for the betterment of it's citizens. Such services improves the transparency speed and consistency of the information.

2. **e-Business:** e-Business is any kind of business or commercial transaction that includes information sharing over internet. Internet has become that powerful tool that it has left behind the shop to shop walk in trend of buying items. Users now prefer sitting at home scrolling over the screen trend of shopping. But the problem with online shopping is that choosing and selecting your choice of product from such a large amount of item is a challenge in itself. To solve this problem Recommder System came into use. Recommender System filters the products according to many criteria depending on the algorithm of that particular recommender and shortlists the products according to the users taste, and thus saving their time. These products can be based on top sellers of that particular site, or analysis of the behavior of the customer through their past purchases. There are many sites like Amazon.com, ebay, flipkart, etc. that uses recommender system. Amazon and flipkart asks the user to check the specific category from a wide range of category to recommend products. Similarly, other sites uses many other different techniques to satisfy their customers needs [5, 8].
3. **e-Library:** e-Library is a collection of online database of digital objects that can include text, image audio or video. In simple words we can say that, it enables users to access digitalize data from anywhere across the world through internet. E-Library provides users with a wide range of information but at the same time the time cost and difficulty in finding proper information increases. Recommender System solves this hindrance by structuring the large content available in the database. Recommender System tends to reduce the information overload and retain the customer by recommending them the subset of data from the universal data according to their preference. This preference can be based on the past history of purchases, feedback given by the user to other books, or the books frequently visited. This recommendation acts like a filter saving users' time from traversing the whole set of data [5].

4. **E-tourism recommender systems:** Internet and mobile devices provide great opportunities to the tourists to access tourism information, but the impressive increase in the number of available tourism choices make it difficult for tourists to choose which option they prefer. This overload problem is solved by the E-tourism recommender systems which are designed to provide suggestions for tourists on the basis of their interest [5].
5. **E-learning recommender systems:** In today internet era, learning process is also changed from tradition method to E-learning method. E-learning RS have become more popular in educations institutions. The main objective of this kind of RS is to assist the users/learners to choose the courses, subjects and learning material as well as users learning activities [5, 7].

6. CONCLUSION

From the last few decades, RS open new oppertunities of retriving the personalized users information on the internet. RS have been used to solve the overload problem by suggesting related and relevant items and services to the users which are not readily available to the users on the system. This paper discuss the phases of RS process like, information collection phase, learning and predicting phase of the RS. It also discuss the traditional recommendation techniquies like Colllaborative, content based and hybrid techniques with several issues which are faced by the RS. This paper also discuss the various application domains of Recommender System like e-Government, e-Business, E-Tourism, E-Learning and many more. Several RS are designed that are based on CF, CB and Hybrid recommendation approches but each have some limitations which are the challenges for research work. Thus, the existing RS needs improvement for present and future requirments of improved recommendation qualities.

REFERENCES

[1] Isinkaye, F.O., Folajimi, Y.O. and Ojokoh, B.A., "Recommendation systems: Principles, methods and Evaluation", *Egyptian Informatics Journal* (2015).

[2] Lu, Jie; Wu, Dianshuang; Mao, Mingsong; Wang, Wei and Zhang, Guangquan, "Recommender System Application Developments: A Survey", https://www.sciencedirect.com/science/article/pii/S0167923615000627 (2017).

[3] Madhukar, Mani, "Challenges and Limitation in Recommender Systems", *International Journal of Latest Trends in Engineering and Technology, (IJLTET)* (2014).

[4] Cortés-Cediel, María E.; Cantador, Iván and Gil, Olga. "Recommender systems for e-governance in smart cities: State of the art and research opportunities", DOI: 10.1145/3127325.3128331, Research Gate (2017).

[5] Kohar, Mukta and Rana, Chhavi, "Survey Paper on Recommendation System", (IJCSIT) *International Journal of Computer Science and Information Technologies,* Vol. 3(2), 2012, 3460–3462.

[6] Sharma, Richa; Vinayak, Sharu and Singh, Rahul, "Guide Me: A Research Work Area Recommender System", *I.J. Intelligent Systems and Applications* (2016).

[7] Khusro, Shah; Alil, Zafar and Ullah, Irfan, "Recommender Systems: Issues, Challenges, and Research Opportunities", https://www.researchgate.net/publication/294860665 (2016).

[8] Zhao, Zhou; Cai, Deng; He, Xiaofei and Zhuang, Yueting, "User Preference Learning for Online Social Recommendation", *IEEE Transactions on Knowledge and Data Engineering* (2016).

Adaptive Sliding Mode Disturbance Observer Control Base Synchronization in a Class of Fractional Order Chua's Chaotic System

Ayub Khan[1] and Uzma Nigar[2]

ABSTRACT: *In this paper, we have investigated synchronization of two identical fractional order (FO) Chua's chaotic system with disturbance using the technique of Adaptive Sliding Mode. Out of many control techniques, it is well observed that classical sliding mode control (SMC) technique has high precise tracking, low sensitivity to external disturbances, and predicated inaccuracy. However, in real applications, some disadvantage of SMC that may cause breakage in the system: chattering problem. To overcome this problem in SMC, some modification has been applied on SMC and proposed an adaptive SMC technique. A non-linear disturbance has many effects, such as increasing the instability and modeling inaccuracy. To resists this problem, a disturbance observer (DOB) is introduced. A DOB is used to estimate the parameter uncertainties, unmodeled dynamics, etc. In this paper firstly, we design, a non-linear FO DOB to tackle the unknown external bounded disturbances (BOD). Based on SMC, we design a sliding mode surface including the FO non-linear DOB for synchronization. Further, using the Lyapunov stability theory (LST), the design appropriate control method has been establishing that the state of master and slave chaotic system is synchronized with different initial conditions. FO Chua's chaotic system with external unknown BOD is synchronized via adaptive SMC for the first time. The introduced technique has many applications in secure communication and neural network.*

Keywords: FO Chua's Chaotic System, Adaptive SMC, FO Disturbance Observer, MATLAB.

[1,2] Department of Mathematics, Jamia Millia Islamia, New Delhi, India.
E-mail: [1]akhan12@jmi.ac.in; [2]uzmanigarkhan@gmail.com

1. INTRODUCTION

Henri Poincare, a French mathematician, which led to the theory of chaos [1]. A chaotic system is nonlinear dynamical systems that display complex and unpredictable behavior and highly sensitive to initial conditions. E.N. Lorenz, best known as the founder of modern chaos theory, gives the introduction of chaos phenomenon in the weather model in 1963 in [2]. In [3], Pecora and Carrol have first introduced the idea of chaos synchronization. Several types of synchronization have been introduced such as anti synchronization [4], complete synchronization [5], projective synchronization [6], hybrid synchronization [7], hybrid projective synchronization [6], etc.

Nowadays, various synchronization method have been applied to study the chaotic behaviour. These are active control [8], adaptive control [9], sliding mode control (SMC) [10], adaptive SMC [11], etc. SMC technique has a significant characteristic, and strong robustness against external disturbances and uncertainty. A nonlinear external disturbance has many effects, such as modeling inaccuracy and energy fluctuations, which can not be easily ignored. To overcome this problem, DOB has been designed. In [11] Shao firstly introduced the synchronization of FO chaotic system with unknown external BOD in 2016.

We present the adaptive SMC technique for the problem of synchronization in FO Chua's chaotic system with unknown external BOD based on Lyapunov stability theory. Hartley *et al.* [12] introduced fractional-order Chua's system. This paper is categorized as follows: Section 2, contains the preliminaries and some essential property of fractional calculus. Section 3, includes the system description of the FO Chua's system . Section 4, contains design methodology of the DOB of FO chaotic system . Section 5, contains an example of an adaptive sliding mode synchronizes of two identical FO Chua's chaotic system studied. Section 6 includes the numerical simulation. Section 7 consiste the application in secure communication, and Section 8 provides conclusions.

2. PRELIMINARIES

Definition 1: The Caputo's derivative of FO $'\alpha'$ of function $f(t)$ can be described as:

$$cD_{\omega}^{\alpha} f(\omega) = \frac{1}{\gamma(n-\alpha)} \int_{c}^{\omega} \frac{f^{n}(\Psi)}{(\omega - \Psi)^{\alpha - n + 1}} d\Psi$$

where $n-1<\alpha<n$ and $\gamma(\alpha)=\int_0^\infty \Psi^{\alpha-1}e^{-\Psi}d\Psi$ is the gamma function.

Property 1: If $\Psi(t)$ is a constant function and the order $q>0$, the Caputo FO derivative fulfill the given condition:

$$D^q\Psi(t)=0$$

Property 2: The Caputo fractional derivative satisfies the linear property:

$$D^\alpha[p_1\Psi_1(t)+p_2\Psi_2(t)]=p_1D^\alpha\Psi_1(t)+p_2D^\alpha\Psi_2(t)$$

Lemma 1: Suppose $\Psi(t)\in R$ be a continuous and derivable function. Then, we have: $\frac{1}{2}D^\alpha\Psi(t)^2\leq\Psi(t)D^\alpha\Psi(t)$, for any time $t\geq t_0$ and $0<\alpha<1$.

Lemma 2: Let the FO system satisfies $D^\alpha\Psi(t)\leq-p_0\Psi(t)+p_1$, then there exists a constant $t_1>0$ in which for all $t\in(t_1,\infty)$, we get $\|\Psi(t)\|\leq\frac{2p_1}{p_0}$.

Assumption: Let $\psi_i(t)$ be the unknown external disturbance for i = 1, 2, ..., n, then the Caputo's derivative of $\psi_i(t)$ is bounded, i.e., $|D^\alpha\psi_i(t)|\leq\xi_i$, where $\xi_i>0$ is an unknown constant.

3. SYSTEM DESCRIPTION

The FO Chua's chaotic system as a master system.

$$\begin{aligned}D^\alpha x_{11m}(t) &= a_1\Big[x_{12m}-x_{11m}-n_1x_{11m}\\ &\quad -\frac{1}{2}(n_0-n_1)(|x_{11m}+1|-|x_{11m}1|)\Big]\\ D^\alpha x_{12m}(t) &= x_{11m}-x_{12m}+x_{13m}\\ D^\alpha x_{13m}(t) &= -b_1x_{12m}-c_1x_{13m}\end{aligned}\quad\ldots(1)$$

where x_{11m},x_{12m},x_{13m} represent state vectors, a_1,b_1,c_1,n_0,n_1 is variable parameter, parameter values chosen as $a_1=10.1911,b_1=10.3035,c_1=0.1631,n_0=-1.1726$ and $n_1=-0.7872$.

Corresponding slave system can be written as:

$$\begin{aligned}D^\alpha y_{11s}(t) &= a_1[y_{12s}-y_{11s}-n_1y_{11s}-\frac{1}{2}(n_0-n_1)(|y_{11s}+1|\\ &\quad -|y_{11s}-1|)]+\psi_1+u_1\\ D^\alpha y_{12s}(t) &= y_{11s}-y_{12s}+y_{13s}+\psi_2+u_2\end{aligned}$$

$$D^{\alpha}y_{13s}(t) = -b_1 y_{12s} - c_1 y_{13s} + \psi_3 + u_3 \quad \ldots (2)$$

where, $\psi_1 = cost, \psi_2 = cos2t, \psi_3 = cos0.5t$ are the unknown bounded disturbances (BOD), and u_1, u_2, u_3 are the controllers. This paper is to estimate the unknown external disturbances $\psi_i(t)$.

4. FO DISTURBAN OBSERVER DESIGN

We create an FO nonlinear DOB as follows: [13]

$$\begin{aligned}\delta_1(t) &= \psi_1(t) - \sigma_1 y_{11s}(t)\\ \delta_2(t) &= \psi_2(t) - \sigma_2 y_{12s}(t)\\ \delta_3(t) &= \psi_3(t) - \sigma_3 y_{13s}(t)\end{aligned} \quad \ldots (3)$$

where, $\sigma_i > 0$ is constant. In (3), using Caputo's derivative and (2), we obtain,

$$\begin{aligned}D^{q_1}\delta_1(t) &= D^{q_1}\psi_1(t) - \sigma_1(a_1[y_{12s} - y_{11s} - n_1 y_{11s} - \frac{1}{2}(n_0 \\ &\quad - n_1)(|y_{11s} + 1| - \quad |y_{11s} - 1|)] + \delta_1 + \sigma_1 y_{11s}) \\ &\quad - \sigma_1 u_1\\ D^{q_2}\delta_2(t) &= D^{q_2}\psi_2(t) - \sigma_2(y_{11s} - y_{12s} + y_{13s} + \delta_2 + \sigma_2 y_{12s}) \\ &\quad - \sigma_2 u_2\\ D^{q_3}\delta_3(t) &= D^{q_3}\psi_3(t) - \sigma_3(-b_1 y_{12s} - c_1 y_{13s} + \delta_3 \\ &\quad + \sigma_3 y_{13s}(t)) - \sigma_3 u_3\end{aligned} \quad \ldots (4)$$

Estimate $\delta_i(t)$ to calculate the external disturbances written as:

$$\begin{aligned}D^{q_1}\hat{\delta}_1(t) &= -\sigma_1(a_1[y_{12s} - y_{11s} - n_1 y_{11s} - \frac{1}{2}(n_0 - n_1)(|y_{11s} \\ &\quad + 1| - |y_{11s} - 1|)] + \sigma_1 y_{11s}(t)) - \sigma_1\hat{\delta}_1(t) \\ &\quad - \sigma_1 u_1\\ D^{q_2}\hat{\delta}_2(t) &= -\sigma_2(y_{11s} - y_{12s} + y_{13s} + \sigma_2 y_{12s}(t)) - \sigma_2\hat{\delta}_2(t) \\ &\quad - \sigma_2 u_2\\ D^{q_3}\hat{\delta}_3(t) &= -\sigma_3\big(-b_1 y_{12s} - c_1 y_{13s} + \sigma_3 y_{13s}(t)\big) \\ &\quad - \sigma_3\hat{\delta}_3(t) - \sigma_3 u_3\end{aligned} \quad \ldots (5)$$

$\hat{\delta}_i(t)$ is the estimate of δ_i. Eq.(3), disturbance $\psi_i(t)$ can be estimated as:

$$\begin{aligned}\hat{\psi}_1(t) &= \hat{\delta}_1(t) + \sigma_1 y_{11s}(t)\\ \hat{\psi}_2(t) &= \hat{\delta}_2(t) + \sigma_2 y_{12s}(t)\\ \hat{\psi}_3(t) &= \hat{\delta}_3(t) + \sigma_3 y_{13s}(t)\end{aligned} \quad \ldots (6)$$

Define disturbance estimation error as:

$$\tilde{\psi}_1(t) = \psi_1 - \hat{\psi}_1, \tilde{\psi}_2(t) = \psi_2 - \hat{\psi}_2, \tilde{\psi}_3(t) = \psi_3 - \hat{\psi}_3$$

Using Eqs. (3) and (6), we get

$$\begin{aligned} \tilde{\delta}_1(t) &= \delta_1(t) - \hat{\delta}_1(t) = \psi_1(t) - \hat{\psi}_1(t) = \tilde{\psi}_1(t) \\ \tilde{\delta}_2(t) &= \delta_2(t) - \hat{\delta}_2(t) = \psi_2(t) - \hat{\psi}_2(t) = \tilde{\psi}_2(t) \\ \tilde{\delta}_3(t) &= \delta_3(t) - \hat{\delta}_3(t) = \psi_3(t) - \hat{\psi}_3(t) = \tilde{\psi}_3(t) \end{aligned} \quad \dots (7)$$

By subtracting Eqs. (5) from (4), the fractional derivatives $\tilde{\delta}_i$ are proposed as:

$$\begin{aligned} D^\alpha \tilde{\delta}_1(t) &= -\sigma_1 \tilde{\delta}_1(t) + D^\alpha \psi_1(t) \\ D^\alpha \tilde{\delta}_2(t) &= -\sigma_2 \tilde{\delta}_2(t) + D^\alpha \psi_2(t) \\ D^\alpha \tilde{\delta}_3(t) &= -\sigma_3 \tilde{\delta}_3(t) + D^\alpha \psi_3(t) \end{aligned} \quad \dots (8)$$

Lyapunov function can be chosen as:

$$V_{\psi_i}(t) = \frac{1}{2}\tilde{\psi}_i^{\,2}(t) = \frac{1}{2}\tilde{\delta}_i^{\,2}(t), \forall i = 1,2,3 \quad \dots (9)$$

Using the Caputo derivative in (9) and using Lemma 1, we have

$$D^\alpha V_{\psi_i(t)} < \tilde{\delta}_i(t) D^\alpha \tilde{\delta}_i(t) \quad \dots (10)$$

Substituting (8) into (10), it has

$$D^\alpha V_{\psi_i}(t) \le \tilde{\delta}_i(t)\left(-\sigma_i \tilde{\delta}_i(t) + D^\alpha \psi_i(t)\right) \le -\sigma_i \tilde{\delta}_i^{\,2}(t) + \tilde{\delta}_i(t) D^\alpha \psi_i(t) \quad \dots (11)$$

We observing that

$$(\tilde{\delta}_i(t) - D^\alpha \psi_i(t))^2 = \tilde{\delta}_i^2(t) + (D^\alpha \psi_i(t))^2 - 2\tilde{\delta}_i(t) D^\alpha \psi_i(t) > 0 \quad \dots (12)$$

$\tilde{\delta}_i(t) D^\alpha \psi_i(t) < \frac{1}{2}\tilde{\delta}_i^2(t) + \frac{1}{2}(D^\alpha \psi_i(t))^2$, from (11) and assumption 1 , we get

$$\begin{aligned} D^\alpha V_{\psi_i}(t) &\le -\sigma_i \tilde{\delta}_i^2(t) + \frac{1}{2}\tilde{\delta}_i^2(t) + \frac{1}{2}\xi_i^{\,2} \le -(\sigma_i - \frac{1}{2})\tilde{\delta}_i^2(t) + \frac{1}{2}\xi_i^{\,2} \\ &= \mu_0 V_{\Theta_i}(t) + \mu_1 \end{aligned} \quad \dots (13)$$

where $\mu_0 = 2\sigma_i - 1$ and $\mu_1 = 0.5\xi_i^2$. If $\sigma_i > 0.5$, This confirms that $\widetilde{\Theta}_i(t), \tilde{\psi}_i(t)$ are bounded using lemma 2 and equation (13), we achieve

$$|V_{\psi_i}(t)| \le \frac{\xi_i^{\,2}}{2(\sigma_i - 0.5)} \quad \dots (14)$$

which means

$$|\tilde{\psi}_i(t)| \leq \sqrt{\frac{{\xi_i}^2}{\sigma_i - 0.5}} \quad \ldots (15)$$

From (15). It indicate that $\tilde{\psi}_i(t)$ is bounded above, where $\tilde{\psi}_i(t)$ represents disturbance estimation error. $|\tilde{\psi}_i(t)|$ fulfill $|\tilde{\psi}_i(t)| \leq k_i$, where $k_i > 0$.

5. EXAMPLE OF ADAPTIVE SMCISTURBANCE OBSERVER

Error: $e_{11}(t) = y_{11s} - x_{11m}, e_{12}(t) = y_{12s} - x_{12m}, e_{13}(t) = y_{13s} - x_{13m}$

Applying Eqn. (1) and (2) in Eqn. (16), the error dynamics becomes:

$$\begin{aligned} D^{\alpha} e_{11}(t) = {} & a_1 e_{12} - a_1 e_{11} + 0.7872 a_1 e_{11} + 0.1927 a_1 [(|y_{11s} \\ & + 1| - |y_{11s} - \quad 1|) - (|x_{11m} + 1| - |x_{11m} \\ & - 1|)] + cost + u_1 \end{aligned}$$

$$\begin{aligned} D^{\alpha} e_{12}(t) &= e_{11} - e_{12} + e_{13} + cos2t + u_2 \quad \ldots (16) \\ D^{\alpha} e_{13}(t) &= -b_1 e_{12} - c_1 e_{13} + cos0.5t + u_3 \end{aligned}$$

To achieve FO error system, the sliding surface $s_i(t)$ is considered as:

$$s_i(t) = e_{1i}(t), i = 1, 2, 3 \quad \ldots (17)$$

Applying the Caputo derivative in (17), we have:

$$D^{\alpha} s_i(t) = D^{\alpha} e_{1i}(t), i = 1, 2, 3 \quad \ldots (18)$$

By using the adaptive SMC method, the desired controller is composed as:

$$\begin{aligned} u_1(t) = {} & -a_1 e_{12} + a_1 e_{11} - 0.7872 a_1 e_{11} - 0.1927 a_1 [(|y_{11s} + 1| \\ & - |y_{11s} - 1|) \ - (|x_{11m} + 1| - |x_{11m} - 1|)] \\ & - \eta_1 s_1 - \hat{k}_1 sign(s_1(t)) - \hat{\psi}_1 \\ u_2(t) = {} & -e_{11} + e_{12} - e_{13} - \eta_2 s_2 - \hat{K}_2 sign(s_2(t)) - \hat{\psi}_2 \quad \ldots (19) \\ u_3(t) = {} & b_1 e_{12} + c_1 e_{13} - \eta_3 s_3 - \hat{k}_3 sign(s_3(t)) - \hat{\psi}_3 \end{aligned}$$

where $\text{sign}(s) = \frac{|s|}{s}$ and $\eta_i > 0$ are constants. $\hat{k}_i$ is estimated value of $k_i$is:

$$\begin{aligned} D^{\alpha} \hat{k}_1 &= \gamma_1 (|s_1(t)| - \hat{k}_1) \\ D^{\alpha} \hat{k}_2 &= \gamma_2 (|s_2(t)| - \hat{k}_2) \quad \ldots (20) \\ D^{\alpha} \hat{k}_3 &= \gamma_3 (|s_3(t)| - \hat{k}_3) \end{aligned}$$

where $\gamma_i > 0, i = 1,\ 2,\ 3$. Using (19) in (16), The error dynamics can be written as:

$$\begin{aligned} D^{\alpha}e_{11}(t) &= -\eta_1 s_1 - \hat{k}_1 sign(s_1(t)) + \psi_1 - \hat{\psi}_1 \\ D^{\alpha}e_{12}(t) &= -\eta_2 s_2 - \hat{k}_2 sign(s_2(t)) + \psi_2 - \hat{\psi}_2 \\ D^{\alpha}e_{13}(t) &= -\eta_3 s_3 - \hat{k}_3 sign(s_3(t)) + \psi_3 - \hat{\psi}_3 \end{aligned} \quad \ldots (21)$$

If the control input is constructed as (19) for error dynamics (16), then the $s_i(t)$ is bounded and stable, and is given in the form of:

$$|s_i(t)| \leq \tau, \text{where } \tau > 0 \quad \ldots (22)$$

Using equation (17) and (22), we obtain

$$e_{1i}(t) \leq \tau \quad \ldots (23)$$

According to the above discussions, it is clear that if the $s_i(t)$ is bounded, then the error $e_{1i}(t)$ is also bounded.

Theorem 1: For the synchronization error system (16) with $0 < \alpha < 1$, the sliding surface is designed according to (17). The unknown BOD is estimated by using the scheme FO non-linear DOB (5) and (6). Then, $e(t)$ is eventually bounded and stable under the adaptive SMC technique as (19) and (20).

Proof: The Lyapunov functional V(t) is selected as:

$$V(t) = \sum_{i=1}^{3} \frac{1}{2} s_i(t)^2 + \sum_{i=1}^{3} \frac{1}{2} \tilde{\psi}_i(t)^2 + \sum_{i=1}^{3} \frac{1}{2} \left(\frac{1}{\sqrt{\gamma_i}} (\hat{k}_i - k_i)\right)^2 \quad \ldots (24)$$

Using property 2 in equation (24), one achieves

$$D^{\alpha}V(t) = \frac{1}{2}\left(\sum_{i=1}^{3} D^{\alpha}s_i(t)^2 + \sum_{i=1}^{3} D^{\alpha}\tilde{\psi}_i(t)^2 + \sum_{i=1}^{3} D^{\alpha}\left(\frac{1}{\sqrt{\gamma_i}}(\hat{k}_i - k_i)\right)^2\right) \quad \ldots (25)$$

using $\tilde{k}_i = \hat{k}_i - k_i$ and Lemma 1 in equation (25) can be written as:

$$D^{\alpha}V(t) \leq \sum_{i=1}^{3} s_i(t) D^{\alpha}s_i(t) + \sum_{i=1}^{3} \frac{1}{2} D^{\alpha}\tilde{\psi}_i^2(t) + \sum_{i=1}^{3} \frac{1}{\sqrt{\gamma_i}} \tilde{k}_i D^{\alpha}\left(\frac{1}{\sqrt{\gamma_i}} \tilde{k}_i\right) \quad \ldots (26)$$

On applying property 2 in equation (26), we obtain

$$D^{\alpha}V(t) \leq \sum_{i=1}^{3} s_i(t)D^{\alpha}s_i(t) + \sum_{i=1}^{3} \frac{1}{2}D^{\alpha}\tilde{\psi}_i(t)^2 + \sum_{i=1}^{3} \gamma_i^{-1}\tilde{K}_i D^{\alpha}\tilde{K}_i \quad \ldots (27)$$

Using (18) and substituting (21) into (27), we achieve

$$D^{\alpha}V(t) \leq \sum_{i=1}^{3} s_i(t)\left(-\eta_i s_i + \tilde{\psi}_i(t) - \hat{k}_i sign\left(s_i(t)\right)\right) + \sum_{i=1}^{3} \frac{1}{2}D^{\alpha}\tilde{\psi}_i^2(t) + \sum_{i=1}^{3} \gamma_i^{-1}\tilde{k}_i D^{\alpha}\tilde{k}_i \quad \ldots (28)$$

Using property 1 and $\tilde{k}_i = \hat{k}_i - k_i$, we get

$$D^{\alpha}\tilde{k}_i = D^{\alpha}\hat{k}_i \quad \ldots (29)$$

where K_i is a constant parameter

In view of equations (20) and (28), one can achieve

$$\sum_{i=1}^{3} \gamma_i^{-1}\tilde{k}_i D^{\alpha_i}\tilde{k}_i = \sum_{i=1}^{3} \tilde{k}_i\left(|s_i(t)| - \hat{k}_i\right) \leq \sum_{i=1}^{3} \tilde{k}_i|s_i(t)| - \frac{1}{2}\sum_{i=1}^{3} \tilde{k}_i^2 + \frac{1}{2}\sum_{i=1}^{3} {k_i}^2 \quad \ldots (30)$$

After substituting the above inequality in (28), we get

$$D^{\alpha}V(t) \leq \sum_{i=1}^{3} s_i(t)\left(-\eta_i s_i(t) + \tilde{\psi}_i(t) - \hat{k}_i sign\left(s_i(t)\right)\right) + \frac{1}{2}\sum_{i=1}^{3} D^{\alpha}\tilde{\psi}_i^2(t) + \sum_{i=1}^{3} \tilde{k}_i|s_i(t)| - \frac{1}{2}\sum_{i=1}^{3} \tilde{k}_i^2 + \frac{1}{2}\sum_{i=1}^{3} {k_i}^2 \quad \ldots (31)$$

We notice that (31) can be rewritten as:

$$D^{\alpha}V(t) \leq -\sum_{i=1}^{3} \eta_i s_i^2(t) + \sum_{i=1}^{3} |s_i(t)|\tilde{\psi}_i| + \sum_{i=1}^{3} \tilde{k}_i|s_i(t)| - \sum_{i=1}^{3} \frac{1}{2}\tilde{k}_i^2 + \sum_{i=1}^{3} \frac{1}{2}{k_i}^2 - \sum_{i=1}^{3} \hat{k}_i|s_i(t)| + \sum_{i=1}^{3} \frac{1}{2}D^{\alpha}\tilde{\psi}_i^2(t) \quad \ldots (32)$$

Using the condition $|\tilde{\psi}_i(t)| < k_i$ and $\sum_{i=1}^{3} \tilde{k}_i|s_i(t)| - \sum_{i=1}^{3} \hat{k}_i|s_i(t)| = -\sum_{i=1}^{3} k_i|s_i(t)|$. Furthermore, (32) can be obtained as

$$D^{\alpha}V(t) \leq -\sum_{i=1}^{3} \eta_i s_i^2(t) - \sum_{i=1}^{3} \frac{1}{2}\tilde{k}_i^2 + \sum_{i=1}^{3} \frac{1}{2}k_i^2 + \sum_{i=1}^{3} \frac{1}{2}D^{\alpha}\tilde{\psi}_i^2(t) \quad \ldots (33)$$

Considering equation (11) and (33), we get

$$D^{\alpha}V(t) \leq -\sum_{i=1}^{3} \eta_i s_i^2(t) - \sum_{i=1}^{3} \frac{1}{2}\tilde{k}_i{}^2 + \sum_{i=1}^{3} \frac{1}{2}k_i^2 + \sum_{i=1}^{3} -(\sigma_i - \frac{1}{2})\tilde{\psi}_i{}^2(t) + \sum_{i=1}^{3} \frac{1}{2}\xi_i^2 \leq -p_2V(t) + p_3 \quad \ldots (34)$$

where $p_2 = min(2\eta_i, 1, 2\eta_i - 1)$ and $p_3 = \sum_{i=1}^{3} \frac{1}{2}\xi_i^2 + \sum_{i=1}^{3} \frac{1}{2}k_i^2$. On selected the value of parameters $\eta_i > 0$ and $\sigma_i > 0.5$, the error is bounded. Using Lemma 2 in (34), we achieve

$$|V(t)| \leq \frac{2p_3}{p_2} = \frac{\sum_{i=1}^{3} \xi_i^2 + \sum_{i=1}^{3} k_i^2}{p_2} \quad \ldots (35)$$

Which implies that

$$\| s(t) \| \leq \sqrt{\frac{2(\sum_{i=1}^{3} \xi_i^2 + \sum_{i=1}^{3} k_i^2)}{p_2}} \quad \ldots (36)$$

From the inequality (36), when $t \to \infty$, the synchronization error $e(t)$ and $s(t)$ are bounded. Thus, error dynamics (16) is stable and bounded from Lemma 2. Hence, bounded synchronization between two identical FO Chua's chaotic system is achieved successfully.

6. NUMERICAL SIMULATION

Numerical simulations are performed in the MATLAB software with $h = 0.001$. The initial states of the master system and slave system are considered as $(x_{11m}, x_{12m}, x_{13m}) = (1,\ 1,\ 1)$, $(y_{11s},\ y_{12s},\ y_{13s}) = (1.4,\ 0.4,\ 0.3)$. The initial states for the error systems are $(e_{11}, e_{12}, e_{13}) = (2.2,\ 3.1, -2.1)$ and $\hat{\delta}_1(0)$, $\widehat{\delta}_2(0)$, $\widehat{\delta}_3(0)) = (.1,\ .1,\ .1)$ and $(\hat{k}_1(0)$, $\hat{k}_2(0), \hat{k}_3(0)) = (.5, .5, .5)$, the designed parameters $(\sigma_1,\ \sigma_2,\ \sigma_3) = (50,\ 50,\ 50), (\gamma_1,\ \gamma_2,\ \gamma_3) = (.1,\ .1,\ .1), (\eta_1,\ \eta_2,\ \eta_3) = (50,\ 50,\ 50)$.

Disturbance assumed as $\psi_1 = cost, \psi_2 = cos2t,\ \psi_3 = cos0.5t$. Phase portrait of Chua's chaotic system in 3-D is shown in Figure 1(a). Synchronization of master system and slave system is illustrated in Fig 1(b-d). Errors $e_{11}(t)$, $e_{12}(t), e_{13}(t)$ are converging to zero as displayed in Figure 1(e). Time trajectories of the BOD ψ_i and disturbance estimation $\hat{\psi}_i$ are depicted in Figure 1(f-h).

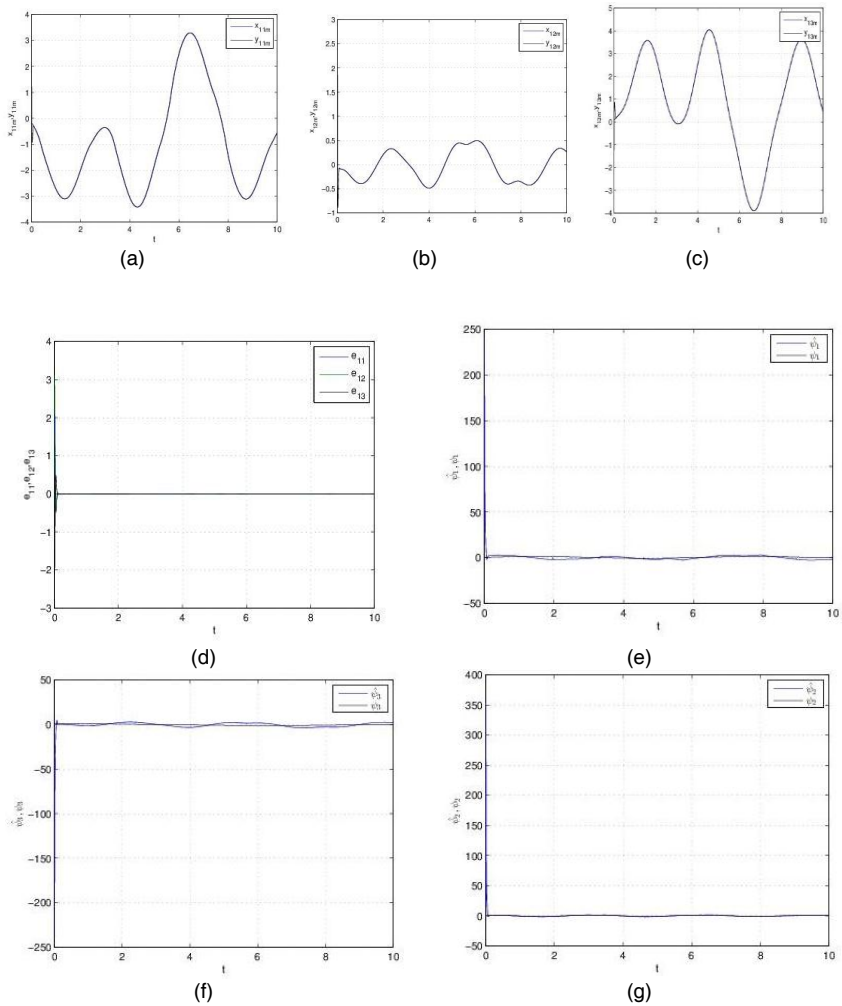


Fig. 1: (a) Synchronization of State Variable between $x_{11m}(t)$ and y_{11s}; **(b)** Synchronization of State Variable between $x_{12m}(t)$ and $y_{12s}(t)$; **(c)** synchronization of state variable between $x_{13m}(t)$ and $y_{13s}(t)$; **(d)** $\hat{\psi}_1(t)$ and $\psi_1(t)$; **(e)** $\hat{\psi}_2(t)$ and $\psi_2(t)$; **(f)** $\hat{\psi}_3(t)$ and $\psi_3(t)$.

7. APPLICATION IN SECURE COMMUNICATION

Across the last two decades, chaotic systems have noticed much application in science and engineering, especially in the area of secure communication. For secure transmission in chaos, the master system is fixed at the transmitter and a slave at the receiver. The central purpose is to protect the information signal masking it with the chaotic signals and transfer to be recovered later only by the approved receiver after taking out the requisite synchronization. In a signal masking method, the message signal $M(t)$ is added to the chaotic master signal, and at the receiver, the transmitted signal is of the form $S(t) = x_{12m}(t) + M(t)$ is assigned to a chaotic slave signal. The recovered message signal $\widehat{M}(t)$ is obtain when the transmitted signal is subtracted from the chaotic signal y_{12s} ie. $\widehat{M}(t) = S(t) - y_{12s}$. For the numerical simulation, we select the $M(t) = 5cos4t$.

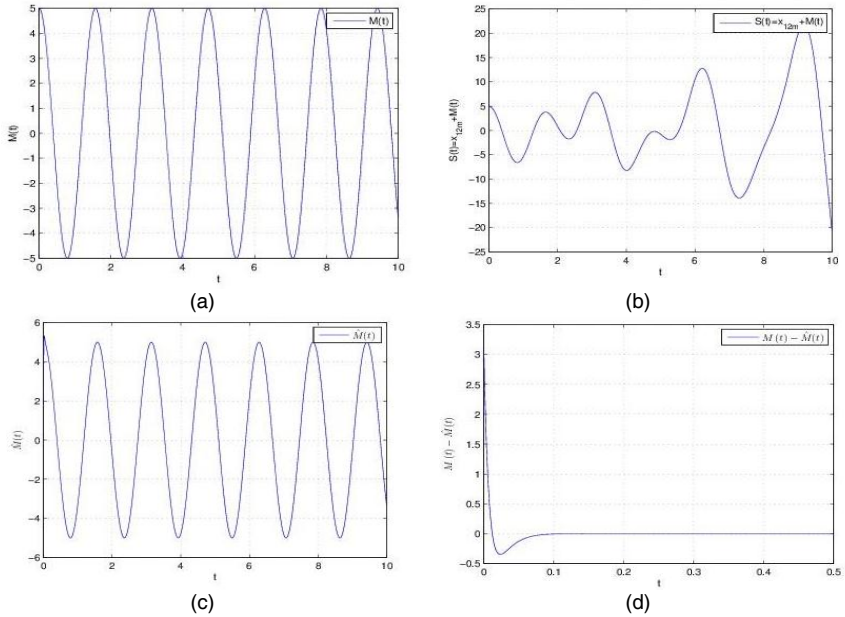


Fig. 2: (a) The Original Message Signal M(T); **(b)** Transmitted Message Signal S(t) **(c)** Recovered Signal $\widehat{M}(t)$; **(d)** Error Message Signal $M(t) - \widehat{M}(t)$

8. CONCLUSION

This work proposes a FO non-linear DOB based adaptive SMC synchronization for the identical FO Chua's chaotic system with unknown external BOD. A non-linear DOB has been applied to approximate the

external disturbance. The sliding surface based on SMC technique is performed to achieving the synchronization among two identical FO Chua's chaotic system had different initial values. It is shown that synchronization error is stable and bounded under the adaptive SMC technique.

REFERENCES

[1] Poincaré, H. (1890). Sur le problème des trois corps et les équations de la dynamique. *Acta mathematica*, 13(1):A3–A270.

[2] Lorenz, E.N. (1963). Deterministic nonperiodic flow. *Journal of the atmospheric sciences*, 20(2):130–141, 1963.

[3] Pecora, L.M. and Carroll, T.L. (1990). Synchronization in chaotic systems. *Physical review letters*, 64(8):821.

[4] Li, G.H. and Zhou, S.P. (2007). Anti-synchronization in different chaotic systems. *Chaos, Solitons and Fractals*, 32(2):516–520, 2007.

[5] Mahmoud, G.M. and Mahmoud, E.E. (2010). Complete synchronization of chaotic complex nonlinear systems with uncertain parameters. *Nonlinear Dynamics*, 62(4):875–882.

[6] Zhenyuan X.uaLiuxiaoGuoac Manfeng Huac and Yang, Y. (2008). Hybrid projective synchronization in a chaotic complex nonlinear system. *Mathematics and Computers in Simulation*, 79(1):449–457.

[7] Sudheer, K.S. and Sabir, M. (2009). Hybrid synchronization of hyperchaotic system. *Pramana*, 73(4):781.

[8] Bhalekar, S. (2014). Synchronization of non-identical fractional order hyperchaotic systems using active control. *World Journal of Modelling and Simulation*, 10(1):60–68.

[9] Khan, A. and Bhat, M.A. (2017). Hyper-chaotic analysis and adaptive multi-switching synchronization of a novel asymmetric non-linear dynamical system. *International Journal of Dynamics and Control*, 5(4):1211–1221.

[10] Vaidyanathan, S. and Sampath, S. (2012). Anti-synchronization of four-wing chaotic systems via sliding mode control. *International Journal of Automation and Computing*, 9(3):274–279.

[11] Shao, S., Chen, M. and Yan, X. (2016). Adaptive sliding mode synchronization for a class of fractional-order chaotic systems with disturbance. *Nonlinear Dynamics*, 83(4):1855–1866.

[12] Hartley, T.T, Lorenzo, C.F. and Qammer, H.K. (1995). Chaos in a fractional order chua's system. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 42(8):485–490.

[13] Chen, M., Chen, W. and Wu, Q. (2014). Adaptive fuzzy tracking control for a class of uncertain mimo nonlinear systems using disturbance observer. *Science China Information Sciences*, 57(1):1–13.

A Single Server Finite Capacity Markovian Feedback Queuing System with Encouraged Arrivals and Reverse Reneging

Bhupender Kumar Som[1] and Sunny Seth[2]

ABSTRACT: *In today's highly volatile uncertain environment, understanding customer behaviour in advance is of utmost importance for any firm in order to stay ahead of competitors. Firms need to design effective operational strategies for overall functioning of the system. Stochastic queuing model can be formulated for such systems by taking into consideration various contemporary challenges being faced by firms and these models then can be solved to derive performance measure of the system so as to give an overview of various characteristics involved in the model. The model developed in this paper is single server finite capacity Markovian queuing model incorporating feedback customers, encouraged arrivals and the concept of reverse reneging.*

Keywords: Encouraged Arrivals, Queuing Model, Reverse Reneging, Feedback Customers.

1. INTRODUCTION AND LITERATURE SURVEY

As it can be observed in the current scenario that in order to attract customers firms often announce sale which encourages more and more customers to join the system (let it be online or offline). Such customers are termed as *encouraged arrivals*. Encouraged arrival can be understood as a concept contrary to the concept of discouraged arrival existing in queuing literature, where arriving customer gets discouraged in joining the queue looking at more customers present in the system. Encouraged arrivals at times become a challenging issue if firm is not

[1] Professor, JIMS Rohini, Sector-5, New Delhi, India. E-mail: bksoam@live.com
[2] Assistant Professor, JIMS Rohini, Sector-5, New Delhi, India.
E-mail: sunnyseth2005@gmail.com

ready and well prepared in advance to entertain the large customer base. In queuing literature, a customer behaviour termed as *reneging* is defined for customers who join the system but leave without availing service after waiting for a time exceeding their threshold limit of patience. Chances of a customer getting impatient and leaving the queue are high in reneging. But in case of encouraged arrivals it has been observed that the patience level of customers is very high and customers are willing to wait for a very longer period of time to avail service, so in this case the concept opposite to reneging prevails and is termed as reverse reneging. Also due to encouraged arrivals as well as reverse reneging customer base of the system grows bigger and puts service facility under pressure and forces servers to service at a little fast pace. This at times leads to some sort of dissatisfaction amongst customers and certain dissatisfied customers may re-join the queue again for availing satisfactory service. Such customers are known as feedback customers. With all these challenges which firms often face these days, it is necessary for firms to understand the functioning of the system well in advance before they announce sale so that efficient strategies and policies can be formed. If firms don't consider such approaches in taking preventive measures in advance they may face chaotic situations which then may hamper their business to a greater extent. As now-a-days with social media in hand, customers don't hesitate to put any bad experience immediately on various social media platforms which hampers not only business but also deteriorate the brand image of the firm.

In [13] Som and Seth introduced the concept of encouraged arrivals and studied single server finite capacity queuing model. They also discussed economic analysis of the model and extended the work in [14] by studying reneging in the model for managing business effectively. They further studied two server heterogeneous model in [15] by incorporating reneging and retention in the model and derived various performance measures of the system and performed sensitivity analysis of the model. In [16] they studied finite capacity multi-server queuing model with encouraged arrivals along with the reneging, retention and feedback and discussed various special cases of the model. In [17] they extended the work by studying infinite capacity multi-server feedback queuing model and discussed encouraged arrivals in non-preemptive priority queuing model in [18] for managing healthcare facility.

Natvig [9] studied queues in which arriving customers gets discouraged by more customers present in queue and Reynolds [10] discussed multi-server queues with discouragement and studied steady state solution of

the model. Haight [6] studied queuing model with reneging and derived steady state distributions of number of customers in the system. Ancker and Gafarian [1, 2] studied queuing model with balking in which customer refuses to join the system and reneging in which customer joins the system but leaves if service does not begin by certain time which is random with negative exponential distribution. They also derived customer loss rate, mean time in queue for reneging and non-reneging customers and also discussed special case of no reneging. Later many authors studied reneging [11, 3, 4]. Wang *et al*. [22] reviewed of queuing models with customer impatience including balking and reneging behaviour of customers and also discussed research tendencies of models.

Takacs [20] studied single server queuing model with feedback customers in steady stage and derived distribution of queue length and also discussed Laplace-Stieltjes transform. Davignon and Disney [5] studied M/G/1 queuing model with feedback customers in steady state and derived busy period distribution. Santhakumaran and Thangaraj [12] studied single server queuing model with impatience and feedback customers. They derived the stationary distribution of arrival process and also discussed no feedback case.

Kumar and Som [7] introduced the concept of reverse reneging and studied queuing model in steady state, derived various measures of the model and further in [8] studied reverse reneging model with reverse balking. Reverse balking is a concept in which probability of a customer joining the system increases with more number of customers present in the system. Thangarajan and Sasikala [21] studied queuing model with controlling the arrival rates at faster and slower rate with reverse reneging. They derived system probabilities with faster and slower rate of arrivals and discussed the characteristics of the model with numerical illustrations. Recently Som and Seth [19] studied model with encouraged arrivals and reverse reneging and extended the work in this paper by incorporating the concept of feedback customers along with encouraged arrivals and reverse reneging in finite capacity single server queuing model.

2. MATHEMATICAL MODEL FORMULATION

It is assumed that Poisson arrivals occur with parameter $\lambda(1+\eta)$, where 'η' represents the percentage increase in number of customers. Service times are exponential with parameter μ. Reverse reneging rate is {N−(n−1)} ξ, where N is the capacity of the system, n is number of customers present in the system and ξ is exponentially distributed reneging

parameter. A feedback customer may re-join the queue with probability $q = 1 - p$and service discipline is FCFS.

Under the assumptions discussed above, following model is formulated:

$$\frac{d}{dt}P_0 = -\lambda(1+\eta)P_0(t) + (\mu p + N\xi)P_1(t)\,; n = 0 \quad \ldots (1)$$

$$\frac{d}{dt}P_n = \lambda(1+\eta)P_{n-1}(t) - \{\lambda(1+\eta) + \mu p + [N-(n-1)]\xi\}P_n(t) + \{\mu p + [N-n]\xi\}P_{n+1}(t)\,; 1 \le n < N \quad \ldots (2)$$

$$\frac{d}{dt}P_N = \lambda(1+\eta)P_{N-1}(t) - (\mu p + \xi)P_N(t)\,; n = N \quad \ldots (3)$$

Validity of the model is tested using Laplace transform and the model developed here is found to be valid.

In Steady-state, taking$P_n(t) = P_n$ and $\frac{d}{dt}P_n(t) = 0\ as\ t \to \infty$, model becomes:

$$0 = -\lambda(1+\eta)P_0 + (\mu p + N\xi)P_1\,; n = 0 \quad \ldots (4)$$

$$0 = \lambda(1+\eta)P_{n-1} - \{\lambda(1+\eta) + \mu p + [N-(n-1)]\xi\}P_n + \{\mu p + [N-n]\xi\}P_{n+1}\,; 1 \le n < N. \quad \ldots (5)$$

$$0 = \lambda(1+\eta)P_{N-1} - (\mu p + \xi)P_N\,; n = N \quad \ldots (6)$$

3. STEADY-STATE SOLUTION

Recursively solving model equations, we get;

$$P_n = \Pr\{n \text{ customers in the system}\} = \prod_{i=1}^{n}\frac{\lambda(1+\eta)}{\mu p + [N-(i-1)]\xi}P_0,;\ 1 \le n\ \le N \quad \ldots (7)$$

Using condition of normality $\sum_{n=0}^{N} P_n = 1$

$$P_0 = Pr\{\text{system is empty}\} = \left\{1 + \sum_{n=1}^{N}\left[\prod_{i=1}^{n}\frac{\lambda(1+\eta)}{\mu p + [N-(i-1)]\xi}\right]\right\}^{-1} \quad \ldots (8)$$

4. MEASURES OF PERFORMANCE

Having calculated probabilistic measures, various performance measures can be derived as follows:

1. Expected system length (L_S)

$$L_s = \sum_{n=1}^{N} nP_n$$

$$L_s = \sum_{n=1}^{N} n \left\{ \prod_{i=1}^{n} \frac{\lambda(1+\eta)}{\mu p + [N-(i-1)]\xi} \right\} P_0 \quad \ldots (9)$$

2. Expected system waiting time (W_S)

$$W_s = \frac{L_s}{\lambda(1+\eta)}$$

$$W_s = \frac{1}{\lambda(1+\eta)} \sum_{n=1}^{N} n \left\{ \prod_{i=1}^{n} \frac{\lambda(1+\eta)}{\mu p + [N-(i-1)]\xi} \right\} P_0 \quad \ldots (10)$$

3. Expected queue waiting time (W_q)

$$W_q = W_s - \frac{1}{\mu}$$

$$W_q = \frac{1}{\lambda(1+\eta)} \sum_{n=1}^{N} n \left\{ \prod_{i=1}^{n} \frac{\lambda(1+\eta)}{\mu p + [N-(i-1)]\xi} \right\} P_0 - \frac{1}{\mu} \quad \ldots (11)$$

4. Expected queue length (L_q)

$$L_q = \lambda(1+\eta) W_q$$

$$L_q = \sum_{n=1}^{N} n \left\{ \prod_{i=1}^{n} \frac{\lambda(1+\eta)}{\mu p + [N-(i-1)]\xi} \right\} P_0 - \frac{\lambda(1+\eta)}{\mu} \quad \ldots (12)$$

5. NUMERICAL ILLUSTRATION

The results of probabilistic measures and measures of performance can be translated in MS Excel and simulated results can be obtained by varying particular parameters. These simulated results can help the firm understand its functioning in advance. Firm can decide by looking at these simulated results that at which service rate firm should provide service for varying arrival rates so as to maintain the smooth and efficient operations in the system and keeping the average length of the system and/or average waiting time of customers in control.

Table 1: Variations in Various Performance Measures with Respect to Arrival Rate Taking μ = 6, ξ = 0.2, η = 0.5 and p = 0.7, N = 10

λ	*Ls*	*Lq*	*Ws*	*Wq*
2	1.078634	0.578634	0.359545	0.192878
2.1	1.209730	0.684730	0.384041	0.217375
2.2	1.357475	0.807475	0.411356	0.244689
2.3	1.523695	0.948695	0.441651	0.274984
2.4	1.710026	1.110026	0.475007	0.308341
2.5	1.917754	1.292754	0.511401	0.344734

λ	*Ls*	*Lq*	*Ws*	*Wq*
2.6	2.147634	1.497634	0.550675	0.384009
2.7	2.399716	1.724716	0.592522	0.425856
2.8	2.673202	1.973202	0.636477	0.469810
2.9	2.966361	2.241361	0.681922	0.515255
3	3.276523	2.526523	0.728116	0.561450
3.1	3.600175	2.825175	0.774231	0.607565
3.2	3.933144	3.133144	0.819405	0.652738
3.3	4.270853	3.445853	0.862799	0.696132
3.4	4.608619	3.758619	0.903651	0.736984
3.5	4.941944	4.066944	0.941323	0.774656

*Simulated data.

The table above shows how various performance measures vary with varying arrival rate of customers. As the arrival rate increases all performance measures increases as expected.

Table 2: Variations in Various Performance Measures with Respect to Service Rate Taking λ = 3, ξ = 0.2, η = 0.5 and p = 0.7, N = 10

μ	*Ls*	*Lq*	*Ws*	*Wq*
4.6	5.564117	4.585856	1.23647	1.019079
4.7	5.37134	4.413893	1.193631	0.980865
4.8	5.181184	4.243684	1.151374	0.943041
4.9	4.994341	4.075974	1.109854	0.905772
5	4.811427	3.911427	1.069206	0.869206
5.1	4.632976	3.750623	1.02955	0.833472
5.2	4.459438	3.594054	0.990986	0.798679
5.3	4.291182	3.442125	0.953596	0.764917
5.4	4.128493	3.295159	0.917443	0.732258
5.5	3.971583	3.153401	0.882574	0.700756
5.6	3.820592	3.01702	0.84902	0.670449
5.7	3.675596	2.886122	0.816799	0.641361
5.8	3.536614	2.760752	0.785914	0.6135
5.9	3.403614	2.640903	0.756359	0.586867
6	3.276523	2.526523	0.728116	0.56145

*Simulated data.

The table above shows how various performance measures vary with varying service rate of customers. As the service rate increases, all the performance measures decreases.

6. CONCLUSION AND FUTURE SCOPE

In this paper we developed a queuing model incorporating all the contemporary challenges being faced by firms in current scenario and then solved this queuing model for deriving various probabilistic measures and performance measures theoretically. These theoretical results are of great use for the firm encountering these issues. These results can be translated in MS Excel for generating simulations to give firms a fair idea about various parameters involved in the model. This paper provides any firm facing issues as discussed above with an insight to look into the future performance and decide the level of various operating parameters in advance so as to be prepared with any kind of situation which may arise in future by making use of the results of simulated results as presented in the numerical illustration section. And effective operational strategies then can be devised in advance for smooth functioning of such systems and situations of chaos can be controlled.

Further economic analysis of the above model can be performed and model can also be studied in transient state. Optimization of the model can be carried out for parameters like service rate, system capacity, etc.

REFERENCES

[1] Ancker Jr, C.J. and Gafarian, A.V. (1963). Some queuing problems with balking and reneging. I. *Operations Research*, *11*(1), 88–100.

[2] Ancker Jr, C.J. and Gafarian, A.V. (1963). Some queuing problems with balking and reneging—II. *Operations Research*, *11*(6), 928–937.

[3] Baccelli, F., Boyer, P. and Hebuterne, G. (1984). Single-server queues with impatient customers. *Advances in Applied Probability*, *16*(4), 887–905.

[4] Bae, J. and Kim, S. (2010). The stationary workload of the G/M/1 queue with impatient customers. *Queueing systems*, *64*(3), 253–265.

[5] D'Avignon, G. and Disney, R. (1976). Single-server queues with state-dependent feedback. *INFOR: Information Systems and Operational Research*, *14*(1), 71–85.

[6] Haight, F.A. (1959). Queuing with reneging, Metrika, 2(1), 186–197.

[7] Kumar, R. and Som, B.K. (2015). A finite single server queuing system with reverse reneging. *American Journal of Operational Research*, *5*(5), 125–128.

[8] Kumar, R. and Som, B.K. (2014). An M/M/1/N Queuing System with Reverse Balking and Reverse Reneging. *Advanced Modelling and Optimization*, *16*(2), 339–353.

[9] Natvig, B. (1975). On a queuing model where potential customers are discouraged by queue length. *Scandinavian Journal of Statistics*, 34–42.

[10] Reynolds, J.F. (1968). The stationary solution of a multiserver queuing model with discouragement. *Operations research*, *16*(1), 64–71.

[11] Robert, E. (1979). Reneging Phenomenon of Single Channel Queues, *Mathematics of Operations Research*, vol. 4(2), 162– 178.

[12] Santhakumaran, A. and Thangaraj, V. (2000). A single server queue with impatient and feedback customers. *International journal of information and management sciences*, *11*(3), 71–80.

[13] Som, B.K. and Seth, S. (2017). An M/M/1/N Queuing system with Encouraged Arrivals. *Global Journal of Pure and Applied Mathematics*, *17*, 3443–3453.

[14] Som, B.K. and Seth, S. (2017, July). Effective Business Management in Uncertain Business Environment Using Stochastic Queuing System with Encouraged Arrivals and Impatient Customers. In *Proceedings of International Conference on Strategies in Volatile and Uncertain Environment for Emerging Markets, IIT Delhi* (pp. 479–488).

[15] Som, B.K. and Seth, S. (2017). An M/M/2/N Queuing System with Encouraged Arrivals, Heterogenous Service and Retention of Impatient Customers. *Advanced Modelling and Optimization*, *19*, 97–104.

[16] Som, B.K. and Seth, S. (2018). M/M/c/N queuing systems with encouraged arrivals, reneging, retention and feedback customers. *Yugoslav Journal of Operations Research*, *28*(3), 333–344.

[17] Som, B.K. and Seth, S. (2017). An M/M/c Feedback Queuing System with Encouraged Arrivals. *Siddhant—A Journal of Decision Making*, *17*(3), 252–255.

[18] Som, B.K. and Seth, S. (2017, December). Stochastic Priority Queuing Model with Encouraged Arrivals for Managing the Healthcare Facility. Presented in *3rd International conference on advances in healthcare management, IIM Ahmedabad*.

[19] Som, B.K. and Seth, S. (2019). An M/M/1/N Encouraged Arrivals Queuing Model with Reverse Reneging. *Journal of Engineering Mathematics & Statistics*, *3*(2), 1–5.

[20] Takacs, L. (1963). A single-server queue with feedback. *Bell system Technical journal*, *42*(2), 505–519.

[21] Thiagarajan M. and Sasikala, S. (2016). The M/M/1/N Interdependent Queuing Model with Controllable Arrival Rates and Reverse Reneging. *Imperial Journal of Interdisciplinary Research*, *2(5),* 945 – 952.

[22] Wang, K., Li, N. and Jiang, Z. (2010, July). Queueing system with impatient customers: A review. In *Proceedings of 2010 IEEE International Conference on Service Operations and Logistics, and Informatics* (pp. 82–87), IEEE.

Study on Cyber Security and Malware Protection

Tripti Lamba[1] and Shivani Kandwal[2]

ABSTRACT: *Cyber Security deals with protecting digital information and information systems. It underlines on protecting limitless data kept up on the frameworks from opponents who might need to get to, degenerate, harm, pulverize or deny access to the equivalent. Hence, it is an important part of the IT field. The target of this paper is to absorb the ideas of Cyber Security alongside the unsafe impacts of Malicious Software and finding proper ways of protecting the system and simultaneously highlighting the cyber security on a global scale and especially in India. Malicious software, or Malware, is a Cyber Security threat that affects the user's computer system by exploiting the system's vulnerabilities. It is another primary threat to the security of information stored in the computer systems. Therefore, there is a need for malware protection. Although, the current security measures and solutions generally protect the computer system against known risks, yet they are attackable to unknown risks. Normally it is hard to predict how anti-malware software would react to new (unknown) malicious software. And for that reason, malware protection remains in the continuous development phase.*

Keywords: Cyber Security, Protection, Attack, Vulnerabilities Crime, Malware.

1. INTRODUCTION

Cybersecurity is the activity of defending frameworks, systems, and projects from advanced attacks. Normally these cyber-attacks targets accessing, manipulating, or destroying susceptible information [17]. Significance of the field is growing due to increasing dependence on computer systems, the Internet and wireless networks and because of the populous increase of "smart" devices. It is also one of the primary

[1] Associate Professor, IITM, Delhi, India. E-mail: triptigautam@yahoo.co.in
[2] IITM, Delhi, India. E-mail: ishivanikandwal@gmail.com

challenges of the modern world [22]. Cyber Security is needed because there are threats to the cyber world. One of the key contestants in that section is Malware. Malware is probably the greatest risk in PC security [5]. Malware protection is more important today than ever since cybersecurity malware threats like viruses and trojans are on a sky-rocketing high.

This paper is divided into various sections. Section 1 gives a brief introduction to the need for cybersecurity and malware protection. The literature review has been discussed in section 2. Section 3 describes the significance of cybersecurity. Cybersecurity objective is discussed in section 4.Major challenges to security measures like IoT threats; AI expansion etc. has been discussed in section 5. Application of Cybersecurity has been discussed in Section 6. Cybersecurity in the world has been explained in section 7. Section 8explains about malware, which is a major concern that is mainly spreading through the internet. Section 9 gives the conclusion of the paper.

2. LITERATURE REVIEW

It's the 21st century or as it is more famously known as the age of electronics. Everything from the most complex of tasks to the basic of works is now digital in nature, and because of that particular fact, day-to-day tasks are easy and convenient to perform. But not everything is all good; this digital age has its own drawbacks. As with the added benefits of digitization of data, come the threats of the digital world.

Cyber Crime is almost an integral part of this world and Cyber Security, a well thought and needed counter measure [1]. Cyber Security deals with safeguarding digital information and information systems technological solutions for shielding personal data and computer networks are essential but not sufficient to ensure security [2]. Cyber-Security, Cyber-Safety, and Cyber-Ethics should be integrated for a better approach. As of late Cybersecurity has emerged as a dug in subject for PC frameworks and foundation with a main issue of defending delicate data on the frameworks from opponents who might need to get to, degenerate, harm, wreck or disallow access to the equivalent. E-commerce, mobile and cloud computing and social networking are becoming prominent trends of cyber security by adopting new technologies [3].

Also, there are some challenges that are caused due to lack of coordination between security agencies and the critical IT infrastructure [2]. Going to

the opposite side of the coin, Malware is one of the essential security dangers looked by the Cyber Community. Malware, or Malicious Software, Exploits the framework's vulnerabilities to overrun itself in the framework and after that damage the information put away in the framework by the client [4].

Typically, it is difficult to anticipate how hostile to malware programming would respond to new (obscure) malevolent programming. What's more, hence, malware security stays in the ceaseless advancement stage so as to have the option to identify new malware or new variations of existing malware, unexpectedly, so is Malware [5].

In May 2019 the Facebook auxiliary said "a progressed digital entertainer" misused a security imperfection and introduced the malware by focusing on clients' cell phones through WhatsApp's call work, conceivably enabling programmers to get to private messages, area information and other data. As a preventive measure and prompt refreshed rendition of the application was discharged and clients were approached to refresh their application.[33]

3. THE SIGNIFICANCE OF CYBER SECURITY

Organizations send susceptible data across different networks and to other devices in the span of commerce, and cybersecurity explain the subject devoted to protecting the same information and the systems used to treat or save it [20].

In today's associated world, everyone gets the profit from exceptional cyber defense affairs. At an individual level, a cybersecurity attack can lead to everything from identity theft, to blackmail attempts, to the destruction of valuable data like family pictures. [10] Everyone relies on vital areas alike financial service companies, hospitals, power plants etc. Ensuring all these and other organizations is essential to keep our society functioning, hospitals, and financial service companies [9].

4. CYBER SECURITY OBJECTIVES

Security comprehensively has four objectives: confidentiality, availability, integrity, and non-repudiation. The first and foremost objective is confidentiality that is keeping information away from people who should not possess the same [27].

After privacy there comes accessibility which means guaranteeing that information spared in the system can be come to by the individuals who

ought to have it. Accessibility is a wide point moving toward things, for example, shortcoming toleration to shield against access control and forswearing of administration to guarantee that information is accessible to those qualified for access it.

In basic words, honesty means verifying that the data spared in the system can never be polluted or modified in a manner that isn't significant. Both classification and accessibility add to uprightness. Security is a huge enough errand simply attempting to meet the classification respectability accessibility goals. Non-renouncement (NR) implies the assurance that somebody can't expel the lawfulness of something.

5. CHALLENGES TO CYBER SECURITY MEASURES

The most troublesome test in cybersecurity is the regularly developing kind of security dangers. [8] Traditionally, associations and the administration have thought the greater part of their cybersecurity assets on outskirt security [32] to shield just their most fundamental framework components and safe-prepare for known dangers.

Today, this methodology is insufficient, since the dangers are abnormal state and change more all of a sudden than associations can stay aware of [21]. Accordingly, counseling associations energize increasingly proactive and versatile procedures to digital security [11].

5.1 Ransomware Evolution

Ransomware is the most despicable aspect of cybersecurity, IT, information experts, and administrators. Sensibly nothing is riskier than a developing infection that snaps onto client and business data that must be evacuated in the event that you satisfy the cybercriminal's unfortunate needs

5.2 A.I. Expansion

Artificial Intelligence might be capable of improve shield against incoming cyber-attacks. The record of present AI applications is already abundant and expanding as shown in Fig. 1. More racing and more reliable account scoring for banks, enhanced disease examination and treatment community for health care organizations are just a few samples. But it's not all genuine, the Artificial Intelligence, if configured, could do more impairment than good.

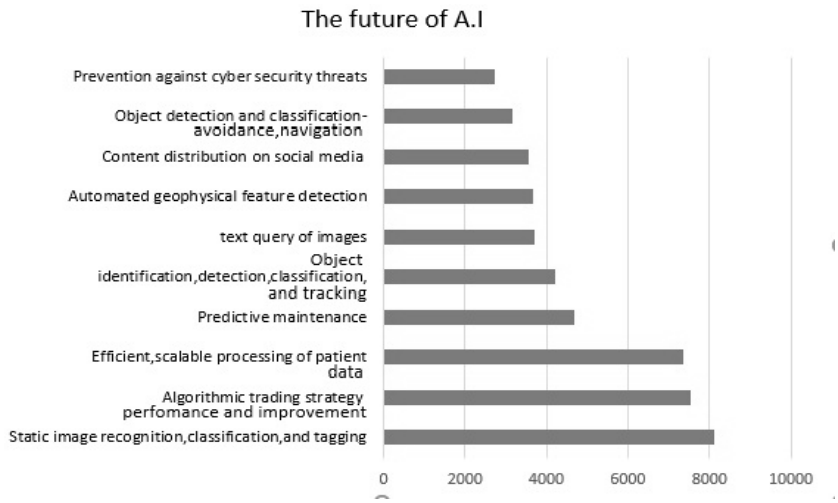


Fig. 1: Forecasted Cumulative Global AI Revenue (2016–2025) by Use Case (U.S dollar)

Source: https://www.statista.com/chart/6810/the-future-of-ai [35]

5.3 Serverless App Vulnerability

Serverless applications can welcome digital assaults. Client information is particularly in risk when clients get your accreditations off-server or locally on their gadget. Unfortunately, with the majority of the weakness that serverless applications speak to, they don't appear to go anyplace in the years to come. Current usage of Serverless Apps is shown in Figure 2.

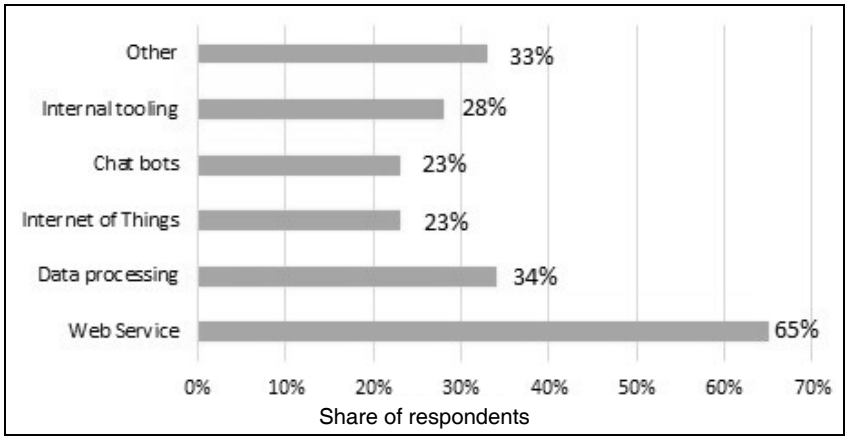


Fig. 2: Uses of Serverless Architecture Worldwide, as of 2016

Source: https://www.statista.com/statistics/658570/worldwide-serverless-architecture-survey-uses [36]

5.4 IoT Threats

Many people are constantly connected. Web of Things is ensuring that each and every gadget you possess is associated with the same. Obviously, the majority of that association conveys with it monstrous advantages, which is the thing that makes it so engaging in any case. The obstruction is that the majority of that interconnectedness makes clients exceptionally delicate to digital assaults. A considerable lot of the gadgets have extreme vulnerabilities.

6. APPLICATIONS OF CYBER SECURITY

- *Filtered Communication:* Include a firewall, against infection, hostile to spam, remote security, and online substance filtration.
- *Protection:* Cybersecurity arrangements give advanced security to your information that will guarantee your workers aren't in danger from potential dangers
- *Increased Productivity:* Viruses can hinder PCs to a slither, and making work for all intents and purposes unthinkable. Successful cybersecurity disposes of this probability, amplifying the potential yield.
- *Denies Spyware:* Spyware is a sort of digital pollution which is planned to observe on your PC activities and convey that information back to the digital lawbreaker.

7. CYBER SECURITY IN THE WORLD

Global Telecommunication Union (ITU) is the UN's specific organization for data and correspondence innovation. ITU's drive Global Cybersecurity Index (GCI) which is an index produced, analyzed and published by the union itself, measures the engagement of countries to cybersecurity to raise cybersecurity consciousness. Country's commitment evaluation is based on five pillars. These are (i) legal, (ii) technical, (iii) organizational, (iv) capacity building, and (v) cooperation [26]. The index holds value for each participating country with respect to the five pillars. A normalized score is generated with respect to the scores of five pillars. On the basis of the normalized score rank of the country is depicted in Figure 3.

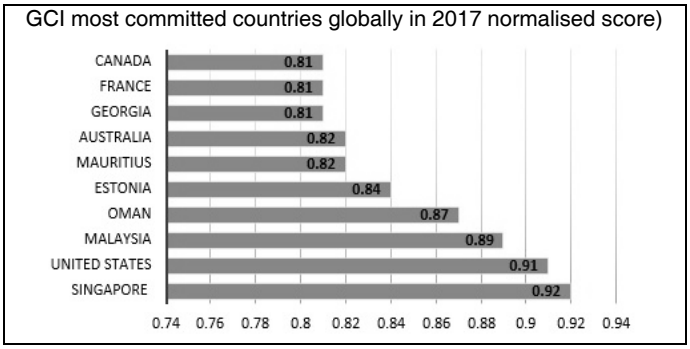


Fig. 3: Top Ten Most Committed Countries Towards Cyber Security, GCI (normalized score) 2017

Source: ITU-GCI 2017. [26]

7.1 Cybersecurity in India

India has been ranked 23rd out of 165 nations and has been mentioned under the maturing category of the index (2017) [27]. However, here lies an utmost concern as well: why India does not have a higher commitment towards cybersecurity? The answer to the question can be answered on the basis of the following four points:

- Lack of awareness among the masses: Despite various efforts from the government' people are still unaware [31].
- National level architecture for Cybersecurity: National level architecture that unifies the efforts of different sectors of the country
- Separation: Boundaries in cyberspace that make activities vulnerable to cyber attacks.
- Lack of uniformity in devices used for internet access: Devices with higher security norms are used by a smaller percentage of the population.

7.2 Measures for Cybersecurity in India

The Government of India legislative, technical and institutional measures for addressing issues related to cybersecurity. Some specific measures are listed below:

- Introduction to the "Digital Swachhta Kendra" (Cleaning of Botnet and examination of Malware) [30]
- National Cyber Security Coordinator (NCSC) under National Security Council Secretariat (NSCS) synchronizes with various offices for digital security related issues at the na-tional level.

- Information Technology Act, 2000: IT act has control arrangements to deal with digital dangers and assaults [28].
- The Indian Computer Emergency Response Team (CERT-In) the national nodal firm for responding to PC security occasions [29].
- National Critical Information Infrastructure Protection Center (NCIIPC) has been accus-tomed for insurance of basic data framework in the country.
- National Cyber Coordination Center (NCCC) has been worked to make essential situational consciousness of subsisting and potential cybersecurity dangers and encourage opportune information sharing for proactive, defensive and preventive activities
- All the new government sites and applications are inspected continuing to their facilitating and on normal premise succeeding facilitating.

8. INTRODUCTION TO MALWARE

Malware is one of the primary dangerous phenomena spreading over the internet. It is a document or code, ordinarily conveyed over a system, that taints, investigates, takes or directs for all intents and purposes any conduct an assailant need. It shifts in sort and abilities [16]. Actually, most Internet issues, for example, spam messages and forswearing of administration assaults have malware as their fundamental reason. With the expanding noteworthiness of malware in Internet assaults, much research was focused on creating systems to gather, study, and relieve malignant code. Thusly, so as to shield our frameworks from a dangerous demonstration of malware, it is important to gather and study malware found on the Internet.Objectives of a Malware Program are as follows:

- *Give Remote Control to an Attacker for utilization of the sullied Machine:* The malware looks at a framework through a far-off system association, things like this are not known to the injured individual [24].
- *Send Spam from the Infected Machine to Innocent Targets:* Negotiated applications are the most famous conveyance framework utilized by programmers to ship malware to clients' gadgets. Malware administrators will more often than not incline toward prominent applications to repackage or ruin, improving the prob-capacity that unfortunate casualties will download their maverick articulation.
- *Review the Infected User's Local Network:* Inserting malware into certifiable online promotion systems to focus on a wide range of end-clients. The promotions appear to be consummately typical and appear on an expansive scope of applications and website pages.

- *Steal Sensible Data:* Malware uncovers casualties of private and prohibitive data with the expectation of adapting held onto information through direct use or underground dispersion.

8.1 Different Types of Malware

- *Viruses:* Programs that imitate themselves all through a PC or system. At the very least, infections can debase or erase information, utilize the client's email to spread, or erase everything on a hard plate. [24]
- *Worms:* Self-duplicating infections that abuse security vulnerabilities to consequently in-wrinkle themselves crosswise over PCs and systems. Not at all like numerous infections, worms don't connect to existing projects or change documents. They normally go incon-spicuous until replication arrives at a scale that uses noteworthy framework assets or system transmission capacity [6].
- *Trojans:* Malware covered up in what gives off an impression of being certified programming. When actuated, Trojans will convey whatever activity they have been customized to take out. Dissimilar to infections and worms, Trojans don't imitate or duplicate through pollution [7].
- *Rootkits:* Programs that give special (root-level) access to a PC. Rootkits change and cover up in the working framework [23].
- *Remote Administration Tools (RATs):* Software that empowers a remote driver to control a framework. These instruments were essentially worked for real use, yet are currently utilized by danger on-screen characters. They are difficult to identify.
- *Spyware:* Malware that assembles data about the use of the tainted PC and conveys it back to the aggressor [18].
- *Botnets:* Short for "robot arrange," these are systems of tainted PCs under the specialist of single assaulting gatherings rehearsing order and-control servers [13].

8.2 Propagation Techniques

- *Web Browsing:* The most straightforward method for getting tainted is through drive-by-download. Using this strategy, malware engineers don't have any objective other than to contaminate whatever number mama chines as would be prudent [15]. Current programs like Chromium (a similar open-source venture on which Google created Chrome) incorporate two components which are structured remembering security [14].

- *USB Thumb Drives:* Thumb drives (or USB drives) are additionally used to develop malignant programming. The technique utilizes the Autorun highlight to begin malware when the capacity gadget is associated by the working system [12].
- *Email Spear Phishing:* Spear phishing is an email parodying trick exertion that objectives a standard particular association which is looking for unapproved access to classified information [16]. Lance phishing endeavors are not begun by discretionary assailants but rather are bound to be directed by culprits out for monetary increase, exchange confidences or military information [9]. Like email data utilized in ordinary phishing efforts, skewer phishing data seem to originate from a submitted source. Here falsehoods the disadvantage of information science (information mining) too [19].

8.3 Malware Protection

Malware security protection gives the essential layer of assurance for a PC or system [26]. The extended need for malware protection has to do with the widespread availability today of exceptional tools originally intended for cyber spying and cyber warfare [25]. Different ways of protecting users from Malware shown in Figure 4.

- *Access Control and Password Security:* The possibility of client name and secret phrase has been a significant method for protecting our data. This can be one of the main measures with respect to cybersecurity and assurance.
- *Authentication of Data:* The records that we get must be confirmed be before downloading that is it ought to be checked if the equivalent has been gone before from a trusted and a dependable source and that they are not altered. [6]
- *Malware Scanners:* This product more often than not checks every one of the records and archives that are available in the framework for vindictive code or damaging infections. Infections, worms, and Trojan steeds are instances of malignant programming.
- *Firewalls:* It a product program or bit of equipment that hides out programmers, infection es, and worms that endeavor to arrive at the machine over the Internet. All messages getting to the Internet go through the firewall present, which looks at each message and deters the individuals who don't coordinate the predetermined security criteria.

- *Anti-Virus Software:* Antivirus programming is a machine program that recognizes, anticipates, and requires exertion to square or take out malevolent programming programs, for example, infections and worms.

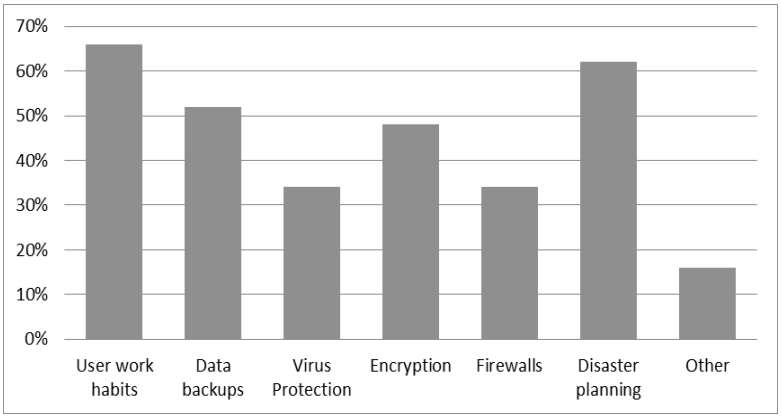


Fig. 4: Ways of Protection from Malware

Source: https://www.netaction.org/security/sect2.html [34]

9. CONCLUSION

Cybersecurity is a huge field that is ending up progressively significant in light of the fact that the world is ending up thickly interconnected as systems being utilized to do unsafe exchanges. This additionally results in malware being spread like out of control fire. Accordingly, Cyber-wrongdoing keeps on leaving behind various ways. Accordingly, so as to accomplish in general security, one should trust however check and realize that assurance is a significant part.

Despite the fact that, the present safety efforts and arrangements for the most part secure the PC framework against known dangers, yet they are attackable to obscure dangers. The most recent and troublesome advancements, alongside the new digital apparatuses and dangers that become known every day, are testing associations with learning of realizing how to verify their foundation as well as, which new stages and knowledge they require to do as such.

There is no perfect or durable answer for digital wrongdoing; it continues advancing as different violations do. In any case, the capacity to have a go at taking an absolute best to limit them is the security that stands in the middle of to guarantee a protected and secure future in the internet.

REFERENCES

[1] Pariekh, N., Patel, D. and Patel, R. (2017). A Disquisition of Cyber-Crime. *IJSTE - International Journal of Science Technology and Engineering*, 3.

[2] Gurusamy, V. and Hirani, B. (2019). Cyber Security for Our Digital Life. *Conference: National Conference On Innovations In Coputer Technolgy And Its Applications (NCICTC'18)*.

[3] Chukwunonso E., C., O. (2019). Malware Analysis and Mitigation in Information Preservation. *IOSR Journal of Computer Engineering (IOSR-JCE)*, *20*(4).

[4] Sen, S., Aydogan, E. and Aysan, A. (2018). Coevolution of Mobile Malware and Anti-Malware. *IEEE Transactions on Information Forensics and Security*, 13(10), pp.2563-2574.

[5] Catota, F., Morgan, M. and Sicker, D. (2019). Cybersecurity education in a developing nation: the Ecuadorian environment. *Journal of Cybersecurity*, *5*(1). doi: 10.1093/cybsec/tyz001

[6] Comer, D. (2015). *Computer Networks and Internets*. Pearson Education Limited.

[7] Gollmann, D. (2011). *Computer security*. Chichester: Wiley.

[8] IEEE Cybersecurity – Home of the IEEE Cybersecurity Initiative. (2019). Retrieved 14 August 2019, from https://cybersecurity.ieee.org.

[9] The Best Antivirus Protection For Home and Business | VIPRE. (2019). Retrieved 14 August 2019, from https://www.threattrack.com/malware-analysis.aspx.

[10] Journal of Cybersecurity *Academic. Oup* 2019, Retrieved 14 August 2019, from https://academic.oup.com/cybersecurity/CrossRef-CitedBy/5382610.

[11] Home - Springer (2019). Retrieved 14 August 2019, from https://link.springer.com/.

[12] USENIX Security '09 Technical Sessions. (2019). Retrieved 14 August 2019, from http://static.usenix.org/event/sec09/tech

[13] A. Saeed, I., Selamat, A. and M.A. Abuagoub, A. (2013). A Survey on Malware and Malware Detection Systems. *International Journal of Computer Applications*, 67(16), 25–31. doi: 10.5120/11480-7108

[14] Verma, A., Rao, M., Gupta, A., Jeberson, W. and Singh, V. (2019). A LITERATURE REVIEW ON MALWARE AND ITS ANALYSIS. *IJCRR*, *5*(16).

[15] The Best Antivirus Protection for Home and Business | VIPRE (2019). Retrieved 14 August 2019, from https://www.vipre.com

[16] Landage, J. and Wankhade, P. (2019). Malware and Malware Detection Techniques : A Survey. Retrieved 14 August 2019, from https://www.ijert.org/malware-and-malware-detection-techniques-a-survey-2

[17] Cyber security IEEE PAPER 2017 (2019). Retrieved 14 August 2019, from https://www.engpaper.com/cyber-security-2017.html.

[18] What is Cyber Security? Definition, Best Practices and More. (2019). Retrieved 14 August 2019, from https://digitalguardian.com/blog/what-cyber-security.

[19] Computer Security (2019). Retrieved 14 August 2019, from https://en.wikipedia.org/wiki/Computer_security

[20] En.wikipedia.org. (2019). *Malware*. [online] Available at: https://en.wikipedia.org/wiki/Malware [Accessed 14 Aug. 2019].

[21] What is Malware? Palo Alto Networks (2019). Retrieved 14 August 2019, from https://www.paloaltonetworks.com/cyberpedia/what-is-malware

[22] What is Malware? Palo Alto Networks. (2019). Retrieved 14 August 2019, from https://www.paloaltonetworks.com/cyberpedia/what-is-malware

[23] (2019). Retrieved 14 August 2019, from https://www.kaspersky.com/resource-center/preemptive-safety/what-is-malware-and-how-to-protect-against-it

[24] Byrnes, F. and Proctor, P. (2019). Information Security Must Balance Business Objectives | Objectives of Computer Security | InformIT. [online] Informit.com. Available at: http://www.informit.com/articles/article.aspx?p=26952 [Accessed 14 Aug. 2019].

[25] Artificial Intelligence Is A Threat To Cybersecurity. It's Also A Solution (2019). Retrieved 14 August 2019, from https://www.bcg.com/

[26] ITU: Committed to connecting the world (2019). Retrieved 14 August 2019, from https://www.itu.int/en/pages/default.aspx

[27] India ranks 23rd among 165 nations in cybersecurity index (2019). Retrieved 14 August 2019, from https://economictimes.indiatimes.com/tech/internet/india-ranks-23rd-among-165-nations-in-cybersecurity-index/articleshow/59478111.cms

[28] Cyber Security (2019). Retrieved 14 August 2019, from http://pib.nic.in/PressReleaseIframePage.aspx?PRID=1556474

[29] Indian - Computer Emergency Response Team (2019). Retrieved 14 August 2019, from https://www.cert-in.org.in/

[30] Das, M. and Niti, S. (2019). *Cyber Swachhta Kendra - PDF, Contact Details, Security Tools, Alerts*. [online] SarkariNiti. Available at: https://sarkariniti.com/cyber-swachhta-kendra/ [Accessed 14 Aug. 2019].

[31] OpenGov Asia (2019). *The current state of cyber security in India*. [online] Available at: https://www.opengovasia.com/the-current-state-of-cyber-security-in-india/ [Accessed 14 Aug. 2019].

[32] (2019). Retrieved 14 August 2019, from http://iitmjanakpuri.com/journals/Volume_8_Issue_1_January-June_2017.pdf

[33] WhatsApp users take note! Agent Smith malware hides your app; over 1.5 crore devices affected in India. (2019). Retrieved 17 August 2019, from

https://www.businesstoday.in/technology/news/whatsapp-infected-agent-smith-malware-devices-in-india-google-play-store-9apps/story/363376.html

[34] Audrie Krause, T. (2019). NetAction - Computer Security Practices in Nonprofit Organizations. Retrieved 17 August 2019, from https://www.netaction.org/security/sect2.html

[35] Armstrong, M. (2019). Infographic: The Future of A.I. Retrieved 17 August 2019, from https://www.statista.com/chart/6810/the-future-of-ai/

[36] Serverless architecture uses worldwide 2016 | Statista. (2019). Retrieved 17 August 2019, from https://www.statista.com/statistics/658570/worldwide-serverless-architecture-survey-uses

Efficiency Evaluation and Benchmarking of Air Carriers in India for the Year 2016–17 Using Data Envelopment Analysis (DEA)

Bharti Seth[1], Punita Saxena[2] and Shalini Arora[3]

ABSTRACT: *Paper aims at assessing and benchmarking the operational efficiency of 12 scheduled (domestic and international) passenger airlines in India during the year 2016–17. Efficiency scores (technical, pure technical and scale) are calculated using input-orientation data envelopment analysis models. Super-efficiency and cross-efficiency ranking methods under variable returns-to-scale are implemented to rank the airlines based on their performance to remove the biased nature of standard DEA models. Air India Express and Spice Jet are serving as benchmarks in Indian Airline Industry for inefficient airlines. Projections are discussed and improvements are mentioned which will help the decision makers to improve their companies performance.*

Keywords: Airline Efficiency, Data Envelopment Analysis, Super-Efficiency, Cross-Efficiency.

1. INTRODUCTION

Indian Air transportation started in the year 1911 and was one of the expensive modes of transport in India for the majority of Indian population. But, nowadays air transport is becoming affordable and hence demand in the market has increased. Aviation industry has become competitive and hence efficiency evaluation is necessary. Since 1911, the civil aviation industry of India has witnessed many revolutionary changes

[1] Research Scholar, Indira Gandhi Delhi Technical University for Women, Delhi, India. E-mail: bhartiseth@igdtuw.ac.in

[2] Associate Professor, Shaheed Rajguru College of Applied Sciences for Women, University of Delhi, India. E-mail: punitasaxena66@gmail.com

[3] Associate Professor, Indira Gandhi Delhi Technical University for Women, Delhi, India. E-mail: shaliniarora@igdtuw.ac.in

and has emerged as one of the rapidly growing industries in the country. India has moved from an underdeveloped to a developing economy and for continual growth of the industry, proprietors of the industry are responsible to implement the efficient decisions of policy makers which are in best interest of population. This paper will provide insight to the decision makers to tackle the challenges prevailing in the market.

It has been accepted world-wide that DEA is a cogent technique to assess the operational efficiencies of airlines. Efficiencies estimated are relative to the best performing airline. Best performing airline is assigned an efficiency score of unity and efficiency of other airlines vary between zero and one. This paper mainly focuses on assessing and benchmarking the performance of airlines operating in India. Techniques of DEA are used to obtain the ranks of 12 scheduled airlines operating in India based on their performance scores.

Rest of the paper is organised as follows. Section 2 presents the contribution of DEA in assessing the performance of airlines; Section 3 presents the models used in the study; Section 4 presents the data and explains the variables used for measuring the performance; Section 5 discusses the findings obtain; Section 6 comprises of concluding remarks.

2. LITERATURE REVIEW: DEA IN AVIATION INDUSTRY

This section agglomerates the work done globally by the researchers in the context of evaluating the performance of operational airlines using the non-parametric technique of DEA.

The term DEA was introduced by Charnes *et al*. [3]. They developed the mathematical programming technique to examine the technical efficiency of a set of homologous decision making units (DMUs) that was later extended by Banker *et al*. [2].

In last four decades (1978–2018), DEA has been applied in the transport sector (land, air and water) and there exists a voluminous literature in this sector. The research efforts related to the evaluation of airlines performance using the techniques of DEA are discussed below.

Precisely, Schefczyk [18] for the first time discovered that DEA is a robust tool for benchmarking of operational performance of airlines using non-financial data. Study analyzed the efficiency of 15 large international airlines for the year 1990 using DEA technique and regression analysis. Good, Roller and Sickles [9] assessed the performance of largest European and American airlines during the year 1976–1986 using two techniques

namely the stochastic frontier model based on regression analysis and full frontier model based on linear programming methods (DEA). The study suggested that European airline industry will be benefitted by deregulation and privatization but the pace of restructuring and reorganization of the industry should be gradual. Tofallis [23] ascertained that Input Efficiency Profiling (IEP) model is more powerful than conventional DEA models while deriving input-specific efficiency ratings. IEP model is applied on the data set of 14 major airlines and results are compared to conventional models of DEA. Fethi *et al*. [8] applied the stochastic DEA on a panel of 17 European Airlines. They used the Land, Lovell and Thore [13] model to find the effect of liberalisation on the efficiency of airlines. After allowing the stochastic error, they concluded that larger airline companies were efficient in 1995 and they resembled those who were efficient in 1993 but not those in 1991. Scheraga [19] studied the sample of 38 large international airlines for the years 1995 and 2000. As both the years were dynamic years in the history of airlines, therefore author investigated whether the operational airlines efficiency implies the financial mobility. He used the non-oriented and oriented DEA models to calculate the efficiency scores and concluded that in a dynamic environment, airlines must have some process of greater capabilities which can allow developing a fund flow so that an airline can operate smoothly. Michaelides *et al*. in [17] estimated the technical efficiency of a panel of 24 largest network airlines over the period 1991–2000. They observed that airlines were technically efficient with scores ranging from 51% to 97% approximately and following the constant returns-to-scale. They compared the approaches of DEA and Stochastic Frontier Analysis (SFA) and concluded the results obtained from both the approaches were approximately similar. Lee and Worthington [14] investigated the efficiency scores of 53 domestic, international and budget airlines in the year 2006 using the conventional DEA models. Study reveals that the exclusion of budget airlines impinged the relative efficiency scores. Efficiency scores of airlines are not much affected by relaxing the input variable of non-flight assets. Zhu in [26] measured the airline performance using a two-stage process. Study applied the centralized model of Liang, Cook and Zhu [15] to a set of 21 airlines and concluded that centralized model is more powerful than standard DEA models. Wen-Min Lu *et al*. studied the relationship between the corporate governance and operating efficiency of 30 airlines operating in US in [16]. They found that low cost airlines are more efficient then the full–service ones but less efficient marketers. Their findings demonstrated that corporate governance has

significantly influenced airline performance. Joo and Fowler in [11] measured the relative efficiency of 80 airlines in Asia, Europe and North America. They found that airlines in Europe are less efficient than airlines in other two regions. Also the efficiency differences between the airlines in Europe and the airlines in other two regions are statistically significantly in terms of technical and purely-technical efficiency scores but not significant between the airlines in Asia and North America. Kottas and Madas in [12] employed the techniques of DEA in assessing the effect of alliance group membership on 30 major international airlines for the period 2012–16. They concluded mainly three facts. Firstly an alliance group membership is not associated with efficient airline. Secondly high freight traffic revenue sharing airlines are more efficient than the airlines with lower freight traffic revenue share and finally the mean efficiency difference regarding air carriers operating in Asian and European continent are statistically significant over the air carriers operating in America. There are other researchers also who have worked in the aviation sector in evaluating the operational efficiency of airlines using the techniques of DEA.

However, the research is still at an infancy stage when we talk about the operational efficiency of an Indian Airlines using the DEA techniques. The efficiency scores of all airlines operating in India during 2006–2010 were studied in [10]. Authors applied variable returns-to-scale and IEP models of DEA over the data and found that budget and low cost airlines were more efficient when compared to the full service airlines in the private sector. Further, private low-cost airlines and all public sector airlines were efficient. They found that greater inefficiency is associated with the input of operating cost. Selvan *et al*. in [21] analysed the performance of airlines operating in India for the period of 2010 to 2014. They used two-stage DEA to evaluate the overall efficiency and applied the super-efficiency DEA model to rank the airlines. They found that production efficiency has higher impact on the overall efficiency than marketing efficiency. They also suggested that airlines in India must emulate the LCC model to be highly efficient.

From the available literature it is justifiable that DEA is an efficient tool in evaluating the operational efficiency of airlines operating in India. This paper aims at:

- Evaluating the technical, pure technical and scale efficiencies of airlines operating in India.

- Ranking the airlines under study by computing super and cross efficiencies.
- Providing the benchmarks for the decision makers of inefficient airline companies to improve.

3. MODELS

3.1 Conventional Models

Conventional models are the standard DEA models, also known as the radial-measures of efficiency. They can be either input oriented or output oriented. Charnes, Cooper and Rhodes (CCR) model is at head when we list down all the models in DEA literature. It was proposed in [3] and was named after the three authors (Charnes, Cooper and Rhodes). While the Banker, Charnes and Cooper (BCC) model was proposed in [2] and it is an extension of the CCR model. CCR model follows the assumption of constant returns-to-scale (CRS) while the production frontier of BCC model is the set spanned by the convex hull of organisational units under study which leads to variable returns-to-scale. Let there be 'n' DMUs whose efficiencies have to be compared. Let $X = \{x_{ir}\}$ and $Y = \{y_{jr}\}$ be $m \times n$ and $s \times n$ input and output matrix respectively. Let $v_i, i = 1, 2, \dots, m$ be input weights and $u_j, j = 1, 2, \dots, s$ be output weights. A DMU whose efficiency is being evaluated is called as **target DMU** and designated as$DMU_o, o = 1, 2, \dots, n$. The efficiency is calculated by CCR model as follows:

$$\max_{v,u} \theta = \frac{u_1 y_{1o} + u_2 y_{2o} + \cdots + u_s y_{so}}{v_1 x_{1o} + v_2 x_{2o} + \cdots + v_m x_{mo}}$$

$$\text{subject to } \frac{u_1 y_{1r} + u_2 y_{2r} + \cdots + u_s y_{sr}}{v_1 x_{1r} + v_2 x_{2r} + \cdots + v_m x_{mr}} \leq 1 \ (\mathrm{r} = 1, 2, \dots, \mathrm{n})$$

$$u_1, u_2, \dots, u_s \geq \epsilon > 0$$

$$v_1, v_2, \dots, v_m \geq \epsilon > 0$$

Here, ϵ is a non-Archimedean constant.

Objective is to obtain the weights v_i and u_j that maximize the efficiency of target DMU. Constraints indicate that efficiency of any other DMU should not exceed 1. Constraints also signify that optimal objective value is at most 1.

The BCC model for calculating the efficiency of D_o is formulated below.

$$\max_{v,u,u_o} \theta = u_1 y_{1o} + u_2 y_{2o} + \cdots + u_s y_{so} - u_o$$

subject to $\frac{u_1 y_{1r} + u_2 y_{2r} + \cdots + u_s y_{sr} - u_o}{v_1 x_{1r} + v_2 x_{2r} + \cdots + v_m x_{mr}} \leq 1$ (r $= 1, 2, \ldots,$ n)

$$v_1 x_{1o} + v_2 x_{2o} + \cdots + v_m x_{mo} = 1$$

u_o unrestricted

$$u_1, u_2, \ldots, u_s \geq 0$$

$$v_1, v_2, \ldots, v_m \geq 0$$

The BCC model differs from CCR model only in the adjunction of the convexity constraint. CCR efficiency will always be less than the BCC efficiency as the feasible region of BCC model is the subset of the feasible region of CCR model.

3.2 SBM Model

Slack-Based Measure (SBM) of efficiency model which was introduced in [24] is a measure of efficiency which is unit-invariant and also it is monotonically decreasing in each input and output slacks. The SBM model is represented as follows:

$$\min_{\lambda, s^-, s^+} \rho = \frac{1 - \frac{1}{m}\sum_{i=1}^{m} \frac{s_i^-}{x_{io}}}{1 + \frac{1}{s}\sum_{r=1}^{s} \frac{s_r^+}{y_{ro}}}$$

Subject to

$$x_o = X\lambda + s^-$$

$$y_o = Y\lambda - s^+$$

$$\lambda \geq 0, s^- \geq 0, s^+ \geq 0$$

Here $s^-$and s^+ are the slacks associated with input and output respectively. The SBM efficiency ρ^* will be equal to 1 only if the slacks are zero in all inputs and outputs. The objective of the model is to minimize the ratio of mean input and output mix inefficiencies. The model is also congruent with the standard DEA model.

3.3 Cross Efficiency Model

This model was first proposed in [22] and then later on in 1994, Doyle and Green re-examined this model in [6]. In this model we calculate the efficiency of each DMU using the weights of inputs and outputs chosen by other DMUs. Thus, cross-efficiency score of a particular DMU reflects that how it is rated by other existing DMUs. Mathematically, cross-

efficiency of DMU_k as measured by DMU_o is defined as the ratio of weighted output and weighted input, where the weights of input and output of DMU_o and levels of input and output of DMU_k are used. Mathematically, it can be expressed as follows:

$$E_{ok} = \frac{\sum_{r=1}^{s} u_{ro} y_{rk}}{\sum_{i=1}^{m} v_{io} x_{ik}}$$

A cross-efficiency matrix is a square matrix of order n and E_{ok} is the entry in row-o and column-k. Diagonal entries of this matrix are called the self-rated efficiencies and this score represents how highly the DMU can rate itself. Entries along any o-row represent the score of all the DMUs when optimal weights of DMU_o are used, that is, how DMU_o rates the other DMUs. Along the k-column the entries represent how DMU_k is rated by all other DMUs. All theses scores help us in ranking the DMUs under study.

3.4 Super Efficiency Model

The foundation of this model was laid in [1]. The model was examined further and infeasibilities and modified version of this model are discussed in [25, 20, 4]. Super-efficiency DEA model calculates the efficiency score while excluding the DMU under study from the reference set. This model helps in ranking even those DMUs which are fully efficient according to most of the DEA models. Super-efficiency model can be applied using the radial and non-radial measures of efficiency. In our study we have used the non-radial slack-based measure of efficiency [25] for calculating the super-efficiency score of efficient DMUs.

4. DATA

Our data set consists of a group of 12 Indian airlines national and private namely: Air India, Air India Express, Alliance Air, Air Asia, Go Air, Indigo, Jet Airways, Jet Lite, Spice Jet, True Jet, Vistara and Zoom Air. Data related to efficiency indicators of scheduled Indian airlines of both service types: domestic and international for the year 2016–17 has been collected from the website which is maintained by the office of Directorate General of Civil Aviation (DGCA). From this source, we identified three input variables and four output variables. It is found that negative correlation prevails among selected input and output variables. Table 1 presents the statistics on input and output variables for the year 2016–17 of selected Indian airlines. Definitions of the selected variables based on handbook on civil aviation statistics for the year 2016–17 available at the website of DGCA are defined below.

Table 1: Descriptive for Input and Output Variables of Airlines under Study

	CASK	*EASK*	*FASK*	*LF*	*FS*	*RPK*	*PR*
Max	10.795	15.909	5.277	91.59	131	46288.31	182434
Min	2.389	0.183	1.528	68.8	1	3.883	14
Average	4.726	1.858	2.465	79.82	40.5	14125.92	56388.21
SD	2.336	4.251	1.118	6.55	45.19	17170.4	68998.15

4.1 Input Variables

Cost per available seat kilometres (CASKs): It is calculated as the ratio of operating expense for the year and available seat kilometres. It is chosen as an input variable as lower CASK implies higher efficiency of an airline.

$$\text{CASK} = \frac{\text{Operating Expense (In Million)}}{\text{ASK (In Million)}}$$

Expenditure on personnel per available seat kilometres (EASKs): It is calculated by dividing the total amount on personnel an airline is spending including their salaries, wages, bonuses and other allowances by the available seat kilometres. This factor is considered as an input variable because lower EASK signifies how efficiently airline company personnel are working.

$$\text{EASK} = \frac{\text{Expenditure on personnel (In Million)}}{\text{ASK (In Million)}}$$

Flight operation Expense per available seat kilometres (FASKs): Lower expense on flight operations implies the lower expenditure on flight movements which indicates a more efficient aircraft. FASK is calculated by dividing the flight operation expense by available seat kilometres. Flight operation expense includes the expenditure on repairing, hiring, insurance and other movements of aircraft.

$$\text{FASK} = \frac{\text{Flight Operation Expense (In Million)}}{\text{ASK (In Million)}}$$

4.2 Output Variables

Revenue Passenger Kilometre (RPK): It is calculated as the sum of the products obtained by multiplying the number of revenue passengers carried on each flight by the flight stage distance. Higher RPK value indicates the higher sale of available seats and therefore it can be considered as an output variable while evaluating the efficiency of an airline.

Load Factor (LF): Load factor indicates that how successful an airline is in selling its available seats. It is calculated by dividing the RPK value by

the available seat kilometres. However higher load factor does not indicate that airline escalating towards efficient frontier. An airline has to maintain a break-even load factor to surmount over all the expenditures in providing the services

Fleet Size (FS): As mentioned in [26], FS can be taken as an output variable and is equal to the total number of aircrafts an airline company owns or leases. Higher fleet size indicates the more seat capacity to meet the market demands.

Passenger Revenue (PR): It is the revenue collected from the sale of tickets and ex-baggage charges. This factor plays a valuable role because an efficient airline has to generate more revenue for its survival.

5. RESULTS AND DISCUSSIONS

5.1 Efficiency Analysis

Efficiency score of the CCR model is called the technical efficiency (TE) and that of the BCC model is called the purely technical efficiency (PTE). TE and PTE scores are also called the CCR score and BCC score respectively. Nature of TE and PTE is global and local respectively. Scale efficiency (SE) is defined as the ratio of CCR score and BCC score. Table 2 shows the efficiency scores in input-orientation and returns-to-scale of 12 scheduled airlines for the year 2016–17. Five airlines namely Air India, Air India Express, Indigo, Jet Airways and Spice Jet are fully efficient according to CCR and BCC model as their TE, PTE and SE scores are all equal to unity so they needs to be further investigated to obtain the ranks. There is one airline namely Go Air which is locally efficient but not globally as it follows the decreasing returns-to-scale. According to SE scores, all the airlines are above average and this may mean that the inefficiencies of airlines (Alliance Air, Jet Lite, True Jet, Vistara and Zoom Air) are mainly due to their inefficient management.

5.2 Benchmarks

Table 2 shows that according to the input-oriented CCR model, Spice Jet and Air India Express are serving as benchmarks for inefficient units Alliance Air and Zoom Air respectively. For other inefficient units (Air Asia, Go Air, True Jet, Vistara and Jet Lite) Spice Jet and Air India Express both are in reference set.

Input-oriented BCC Model identifies that Air India Express is only serving as a benchmark for both of the inefficient units namely Alliance Air and

Zoom Air. Spice Jet and Air India Express are in reference set for other inefficient units. Further it is found that Air India, Indigo, Go Air and Jet Airways are although efficient units but they are not in the reference set of any inefficient unit.

Table 2: Efficiencies and Returns-to-Scale of Airlines in India for the Year 2016–17

S. No.	*DMU*	*CCR Scores*		*BCC Scores*			*SE*
		TE	*Reference*	*PTE*	*Reference*	*RTS*	
1.	Air India	1	Air India	1	Air India	Constant	1
2.	Air India Express	1	Air India Express	1	Air India Express	Constant	1
3.	Alliance Air	0.537196	Spice Jet	0.631661	Air India Express		0.85044
4.	Air Asia	0.825325	Air India Express, Spice Jet	0.931174	Air India Express, Spice Jet		0.886326
5.	Go Air	0.867612	Air India Express, Spice Jet	1	Go Air	Decreasing	0.867612
6.	Indigo	1	Indigo	1	Indigo	Constant	1
7.	Jet Airways	1	Jet Airways	1	Jet Airways	Constant	1
8.	Jet Lite	0.655842	Air India Express, Spice Jet	0.659082	Air India Express, Spice Jet		0.99508
9.	Spice Jet	1	Spice Jet	1	Spice Jet	Constant	1
10.	True Jet	0.295379	Air India Express, Spice Jet	0.298401	Air India Express, Spice Jet		0.98988
11.	Vistara	0.647246	Air India Express, Spice Jet	0.669106	Air India Express, Spice Jet		0.96733
12.	Zoom Air	0.493715	Air India Express	0.53435	Air India Express		0.92396

5.3 Rankings

Cross-efficiency and super-efficiency methods are applied to the panel of data and Table 3 shows the efficiency scores with ranks. Cross-efficiency scores under variable returns-to-scale are calculated and super-efficiency scores using slack-based measure of efficiency under variable returns-to-scale are calculated. In Table 2 almost half of the airlines under evaluation

are fully efficient but we can now analyse these fully efficient airlines further through cross-efficiency and super-efficiency scores. Air India Express is maintaining the first position throughout in all the scores. Although True Jet is efficient with respect to SE (0.98988) but its low PTE score forces it to be inefficient with lowest cross-efficiency and SBM scores. According to both the ranking methods, it is found that Air India is ranked low among all the efficient airlines which signifies its absence from the reference sets of inefficient units.

Table 3: Cross-Efficiency and Super-Efficiency Scores with Ranks

S. No.	*DMU*	*SBM Super Efficiency*	*Ranking*	*BCC Cross Efficiency*	*Ranking*
1.	Air India	1.017453	5	0.832369682	7
2.	Air India Express	1.5329	1	1.19209462	1
3.	Alliance Air	1	10	0.333709007	11
4.	Air Asia	1	7	0.903502888	5
5.	Go Air	1.005183	6	0.953641339	4
6.	Indigo	1.317317	2	1.017452027	3
7.	Jet Airways	1.043206	3	0.844982037	6
8.	Jet Lite	1	9	0.673950726	8
9.	Spice Jet	1.028995	4	1.018600658	2
10	True Jet	1	12	0.319217574	12
11.	Vistara	1	8	0.652377606	9
12.	Zoom Air	1	11	0.383297825	10

5.4 Correlation of Ranks

Ranks for airlines are obtained using the cross-efficiency method under variable returns-to-scale and super-efficiency method using slack-based measure under variable returns-to-scale. We need to find out that the ranks obtained from both the methods are independent of each other or not. Spearman's Rank Correlation coefficient (ρ) was computed and was observed to have a p-value of 1 at 5% level of significance. This justifies that the two ranking structures are independent of each other.

5.5 Slacks and Projections

Table 4(a, b, c) represents the input excess and output shortfalls obtained from CCR model with projections. LF row in Table 4(a, b, c) implies that every airline is successful in selling their seats. According to the results, no airline needs to improve this LF output variable. However, it is ironical that Alliance Air and True Jet airlines are capable of selling their available seats while maintaining the break-even load factor but still have very high projection percentage. Inefficient airlines need to reduce CASK in the range (13%–75%), EASK in the range (13%–98%), FASK in the range (13%–70%). It is found that Go Air is the one that needs a minimum reduction of approx 13% in all the input variables. The projection percentage of EASK for Zoom Air is very high (99%) as they got very small time-window to business in the market. It is found that True Jet and Alliance Air are spending more on their employees.

Table 5(a, b, c) represents the input excess and output shortfalls obtained from BCC model with projections. According to BCC model, the reduction percentage range of CASK, EASK and FASK are (7%–78%), (19%–99%) and (20%–71%) respectively. It is found that Air Asia requires least reduction in all the three input variables. BCC model also suggest that Alliance Air and Zoom Air are spending more on their employees. True Jet is the one which is spending more on flight operations. Further it is found in contrast with CCR results that Alliance Air needs to improve its load factor by 10% approx. In future this can be further investigated.

Table 4(a): Slacks and Projections (%) with Respect to Input and Output Variables using CCR Model

	DMU	*Air India*		*Air India Express*		*Alliance Air*		*Air Asia*	
	TE	***1***		***1***		***0.5372***		***0.82532***	
		S	*P%*	*S*	*P%*	*S*	*P%*	*S*	*P%*
Inputs	CASK	0	0.00%	0	0.00%	3.12737	–75.25%	0	–17.47%
	EASK	0	0.00%	0	0.00%	0.51585	–79.93%	0.10443	–43.74%
	FASK	0	0.00%	0	0.00%	0	–46.28%	0	–17.47%
Outputs	LF	0	0.00%	0	0.00%	0	0.00%	0	0.00%
	FS	0	0.00%	0	0.00%	25.8065	234.60%	19.91345	248.92%
	RPK	0	0.00%	0	0.00%	10985.9536	999.90%	7522.80899	268.73%
	PR	0	0.00%	0	0.00%	40638.5703	999.90%	29670.9122	333.03%

Table 4(b): Slacks and Projections (%) with Respect to Input and Output Variables Using CCR Model

	DMU	*Go Air*		*Indigo*		*Jet Airways*		*Jet Lite*	
	TE	***0.86761***		***1***		***1***		***0.65584***	
		S	*P%*	*S*	*P%*	*S*	*P%*	*S*	*P%*
Inputs	CASK	0.01011	–13.53%	0	0.00%	0	0.00%	0	–34.42%
	EASK	0	–13.24%	0	0.00%	0	0.00%	0.17055	–61.42%
	FASK	0	–13.24%	0	0.00%	0	0.00%	0	–34.42%
Outputs	LF	0	0.00%	0	0.00%	0	0.00%	0	0.00%
	FS	11.85162	51.53%	0	0.00%	0	0.00%	21.94732	274.34%
	RPK	3833.38381	47.52%	0	0.00%	0	0.00%	8019.72754	334.63%
	PR	14113.6857	45.55%	0	0.00%	0	0.00%	27702.044	237.89%

Table 4(c): Slacks and Projections (%) with Respect to Input and Output Variables Using CCR Model

	DMU	*Spice Jet*		*True Jet*		*Vistara*		*Zoom Air*	
	TE	***1***		***0.29538***		***0.64725***		***0.49372***	
		S	*P%*	*S*	*P%*	*S*	*P%*	*S*	*P%*
Inputs	CASK	0	0.00%	0	–70.46%	0.23759	–40.26%	0	–50.63%
	EASK	0	0.00%	0.11521	–81.07%	0	–35.28%	7.68547	–98.94%
	FASK	0	0.00%	0	–70.46%	0	–35.28%	0.78255	–68.24%
Outputs	LF	0	0.00%	0	0.00%	0	0.00%	0	0.00%
	FS	0	0.00%	21.63463	540.87%	22.95235	176.56%	20.25076	999.90%
	RPK	0	0.00%	9258.069	999.90%	8435.467	265.66%	8114.337	999.90%
	PR	0.00001	0.00%	34097.14	999.90%	32002.85	256.94%	30165.13	999.90%

Table 5(a): Slacks and Projections (%) with Respect to Input and Output Variables Using BCC Model

	DMU	*Air India*		*Air India Express*		*Alliance Air*		*Air Asia*	
	PTE	***1***		***1***		***0.63167***		***0.93117***	
		S	*P%*	*S*	*P%*	*S*	*P%*	*S*	*P%*
Inputs	CASK	0	0.00%	0	0.00%	4.42985	–77.87%	0	–6.88%
	EASK	0	0.00%	0	0.00%	0.78572	–88.08%	0.0473	–18.78%
	FASK	0	0.00%	0	0.00%	0	–36.83%	0.27907	–20.34%
Outputs	LF	0	0.00%	0	0.00%	7.11405	10.34%	0	0.00%
	FS	0	0.00%	0	0.00%	12	109.09%	31.0535	388.17%
	RPK	0	0.00%	0	0.00%	8462.694	999.90%	9858.492	352.17%
	PR	0	0.00%	0	0.00%	29580.1	959.39%	39525.07	443.63%

Table 5(b): Slacks and Projections (%) with Respect to Input and Output Variables Using BCC Model

	DMU	*Go Air*		*Indigo*		*Jet Airways*		*Jet Lite*	
	PTE	***1***		***1***		***1***		***0.65908***	
		S	*P%*	*S*	*P%*	*S*	*P%*	*S*	*P%*
Inputs	CASK	0	0.00%	0	0.00%	0	0.00%	0.05913	–35.54%
	EASK	0	0.00%	0	0.00%	0	0.00%	0.18547	–63.46%
	FASK	0	0.00%	0	0.00%	0	0.00%	0	–34.09%
Outputs	LF	0	0.00%	0	0.00%	0	0.00%	0	0.00%
	FS	0	0.00%	0	0.00%	0	0.00%	20.50185	256.29%
	RPK	0	0.00%	0	0.00%	0	0.00%	7716.662	321.99%
	PR	0	0.00%	0	0.00%	0	0.00%	26423.42	226.92%

Table 5(c): Slacks and Projections (%) with Respect to Input and Output Variables Using BCC Model

	DMU	*Spice Jet*		*True Jet*		*Vistara*		*Zoom Air*	
	PTE	***1***		***0.2984***		***0.66911***		***0.53435***	
		S	*P%*	*S*	*P%*	*S*	*P%*	*S*	*P%*
Inputs	CASK	0	0.00%	0	–70.16%	0.64327	–46.59%	0	–46.57%
	EASK	0	0.00%	0.11126	–80.40%	0.09468	–53.69%	8.3181	–98.85%
	FASK	0	0.00%	0.02028	–70.54%	0	–33.09%	0.84697	–65.63%
Outputs	LF	0	0.00%	0	0.00%	0	0.00%	5.77358	8.23%
	FS	0	0.00%	22.44641	561.16%	13.45007	103.47%	22	999.90%
	RPK	0	0.00%	9428.27	999.90%	6443.167	202.92%	8782.581	999.90%
	PR	0.00001	0.00%	34815.21	999.90%	23597.41	189.46%	32649.3	999.90%

6. CONCLUSIONS

The present paper has targeted the 12 scheduled passenger airlines operating in India. Airlines of both service-type domestic and international are considered for evaluating their operational efficiencies. TE, PTE and SE scores are obtained in Table 2 using input-orientation CCR and BCC models. It is found that five fully efficient airlines are lying in the most productive scale size region, region where constant returns-to-scale property dominates. SE scores suggest that airlines are globally inefficient because of their inefficient management or operations. Therefore the airlines executives need to strategise their managerial policies in order to become an efficient airline.

It is also found that efficient airlines not present in the reference sets are holding a low rank according to ranking methods. Super-efficiency and cross-efficiency scores in Table 3 justify that Air India Express is the most efficient airline among all the efficient airlines. Ranks obtained from super-efficiency and cross-efficiency methods are independent.

Results in Tables 4 and 5 showed that Alliance Air and True Jet need to curb the expenditure on their personnel for their survival. As mentioned in section 5 that LF is the variable which needs no improvement according to CCR model but BCC model suggested that Alliance Air needs to increase its load factor by 10%. In this direction, future research work can be done to investigate the relation between the load factor and other output variables related to revenue. Also, sensitivity analysis can be done in relation to the effect on operational efficiency of airlines on deletion of any input and output variables under study. In addition airline companies should strategize their future policies in a holistic way which is propitious for both customers and personnel.

REFERENCES

[1] Andersen, P. and Petersen, N.C. (1993). A Procedure for Ranking Efficient Units in Data Envelopment Analysis. Management Science, 39(10), 1261–1264. doi:10.1287/mnsc.39.10.1261.

[2] Banker, R.D., Charnes, A. and Cooper, W.W. (1984). Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. Management Science, 30(9), 1078–1092. doi:10.1287/mnsc.30.9.1078.

[3] Charnes, A., Cooper, W. and Rhodes, E. (1978). Measuring the efficiency of decision making units. European Journal of Operational Research, 2(6), 429–444. doi:10.1016/0377-2217(78)90138-8.

[4] Cook, W.D., Liang, L., Zha, Y. and Zhu, J. (2009). A modified super-efficiency DEA model for infeasibility. Journal of the Operational Research Society, 60(2), 276–281. doi:10.1057/palgrave.jors.2602544.

[5] Cooper, W. and Tone, K. (1997). Measures of inefficiency in data envelopment analysis and stochastic frontier estimation. European Journal of Operational Research, 99(1), 72–88. doi:10.1016/s0377-2217(96)00384-0.

[6] Doyle, J. and Green, R. (1994). Efficiency and Cross-efficiency in DEA: Derivations, Meanings and Uses. Journal of the Operational Research Society, 45(5), 567–578. doi:10.1057/jors.1994.84

[7] Directorate General of Civil Aviation (11.01.2019). Retrieve from www.dgca.nic.in

[8] Fethi, M.D., Jackson, P.M. and Weyman-Jones, T.G. (2001). European Airlines: A Stochastic DEA Study of Efficiency with Market Liberalisation.

[9] Good, D.H., Röller, L. and Sickles, R.C. (1995). Airline efficiency differences between Europe and the US: Implications for the pace of EC integration and domestic regulation. European Journal of Operational Research, 80(3), 508–518. doi:10.1016/0377-2217(94)00134-x.

[10] Jain, R.K. and Natarajan, R. (2015). A DEA study of airlines in India. Asia Pacific Management Review, 20(4), 285–292. doi:10.1016/j.apmrv.2015.03.004.

[11] Joo, S. and L. Fowler, K. (2014). Exploring comparative efficiency and determinants of efficiency for major world airlines. Benchmarking: An International Journal, 21(4), 675–687. doi:10.1108/bij-09-2012-0054.

[12] Kottas, A.T. and Madas, M.A. (2018). Comparative efficiency analysis of major international airlines using Data Envelopment Analysis: Exploring effects of alliance membership and other operational efficiency determinants. Journal of Air Transport Management, 70, 1–17. doi:10.1016/j.jairtraman.2018.04.014.

[13] Land, K.C., Lovell, C.A. and Thore, S. (1993). Chance-constrained data envelopment analysis. Managerial and Decision Economics, 14(6), 541–554. doi:10.1002/mde.4090140607.

[14] Lee, B. and Worthington, A. (2010). The relative efficiency of international, domestic, and budget airlines: nonparametric evidence.

[15] Liang, L.M., Cook, W.D. and Zhu, J. (2008). DEA models for two-stage processes: Game approach and efficiency decomposition. Naval Research Logistics, 55(7), 643–653. doi:10.1002/nav.20308.

[16] Lu, W., Wang, W., Hung, S. and Lu, E. (2012). The effects of corporate governance on airline performance: Production and marketing efficiency perspectives. Transportation Research Part E: Logistics and Transportation Review, 48(2), 529–544. doi:10.1016/j.tre.2011.09.003.

[17] Michaelides, P., Belegri-Roboli, A. and Marinos, T. (2010). Evaluating the Technical Efficiency of Trolley Buses in Athens, Greece. Journal of Public Transportation, 13(4), 93–109. doi:10.5038/2375-0901.13.4.5.

[18] Schefczyk, M. (1993). Operational performance of airlines: An extension of traditional measurement paradigms. Strategic Management Journal, 14(4), 301–317. doi:10.1002/smj.4250140406.

[19] Scheraga, C.A. (2004). Operational efficiency versus financial mobility in the global airline industry: A data envelopment and Tobit analysis. Transportation Research Part A: Policy and Practice, 38(5), 383–404. doi:10.1016/j.tra.2003.12.003.

[20] Seiford, L.M. and Zhu, J. (1999). Infeasibility of Super-Efficiency Data Envelopment Analysis Models. INFOR: Information Systems and Operational Research, 37(2), 174–187. doi:10.1080/03155986.1999.11732379.

[21] Selvan, N.T., Arasu, B.S. and Arul, S.T., 2017. Benchmarking for Indian Airlines Industry in Contemporary Market Scenario. Global Journal of Management And Business Research.

[22] Sexton, T.R., Silkman, R.H. and Hogan, A.J. (1986). Data envelopment analysis: Critique and extensions. New Directions for Program Evaluation, 1986(32), 73–105. doi:10.1002/ev.1441.

[23] Tofallis, C. (1997). Input efficiency profiling: An application to airlines. Computers & Operations Research, 24(3), 253–258. doi:10.1016/s0305-0548(96)00067-6.

[24] Tone, K. (2001). A slacks-based measure of efficiency in data envelopment analysis. *European Journal of Operational Research*, 130(3), 498–509. doi:10.1016/s0377-2217(99)00407-5.

[25] Tone, K. (2002). A slacks-based measure of super-efficiency in data envelopment analysis. *European Journal of Operational Research*, 143(1), 32–41. doi:10.1016/s0377-2217(01)00324-1.

[26] Zhu, J. (2011). Airlines Performance via Two-Stage Network DEA Approach. *Journal of CENTRUM Cathedra: The Business and Economics Research Journal*, 4(2), 260–269. doi:10.7835/jcc-berj-2011-0063.

Big Data Applications in Real-World

Shalini Sharma[1], Naresh Kumar[2] and Kuldeep Singh Kaswan[3]

ABSTRACT: *This paper presents the application of big data in real world. Nowadays Big data is used in all spheres of personal and professional lives be it agriculture, banking, travelling, healthcare, financial institutions, food industry, power sector, managing smart cities and homes. The goal of big data is to analyse the vast amount of data at the disposal of varied sectors and help them in taking better decisions regarding management, security, sales etc. in order to perform better in this competitive world. Although Big data is used in various fields we usually consider only a few real world applications which impact the general population and provide support and comfort to commoners in tremendous ways. Using qualified and skilled analyst to analyze large amount of complex data collected from various resources like sensors, smart phones, digital devices, Internet of Things, customers feedback, social media, cloud computing and data mining, big data is helping companies to make better informed decisions to secure business growth.*

Keywords: Internet of Things, Big Data Analytics, Data Visualization, Integration, Cloud Computing, Machine Learning.

1. INTRODUCTION

Big data is a term used to refer to data sets that are too large or complex for traditional data-processing application software to adequately deal with.

The definition of Big Data, given by Gartner is, "Big Data is high-volume, and high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation" [1]. As the name suggests data that is big in amount or we can say that big data is collection of large amount of data collected from diverse fields in real

[1,2,3] School of Computing Science and Engineering, Galgotias University, Greater Noida, UP, India. E-mail: [1]shalini.sharma@rediffmail.com; [2]Naresh.dhull@gmail.com; [3]kaswankuldeep@gmail.com

world i.e. it can be from social networking sites, or medical data, data related to climate, or it can be your mobile data, transactional data etc.

The data collected could be of different variety and format, they might be structured, semi-structured or non-structured depending upon the areas from where it was collected. The data collected in such manner will be generated and logged at very high speed making it impossible for a single machine to store or process. Big data is something that can be used to analyse insights which can lead to better decisions and strategic business moves According to a study by Boyd & Crawford [2] it "rests on the interplay of:

(a) *Technology:* Maximizing computation power and algorithmic accuracy to gather, analyze, link and compare large data sets.
(b) *Analysis:* Drawing on large data sets to identify patterns in order to make economic, social, technical, and legal claims.
(c) *Mythology:* The widespread belief that large data setsoffer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy."

The data collected is of no use unless it is analysed and a useful outcome is produced for decision making. In today's world enormous amount of data is at the disposal of decision makers. Big data is formally defined as 4V's which refers to the datasets high in volume, variety, velocity and veracity, where veracity means availability and accountability.

IBM Scientists Mention that *Big Data* has Four-Dimensions: Volume, Velocity, Variety, and Veracity [7]. Because of these characteristics of big data it is difficult to handle them using traditional tools and techniques. Furthermore these data sets need to be studied and solutions must be provided in order to extract value and information from these datasets. Decision makers can gain valuable insights from solutions provided after analysing the varied and rapidly changing data using big data analytics.

2. BIG DATA ANALYTICS

Big Data gives us a new technology that provides insights available in real-time. Business, governments and even NGOs are benefited with the insights offered by big data analysis. Nowadays data is available in abundance and so is its access like search on Google or interaction on Facebook through status updates likes or dislikes pictures and videos. All such things are data or information of varying nature thus it can be analysed, and utilised for numerous applications.

By analysing social media data, user preferences can be known while shopping, destinations that they like to travel, type of hotels they prefer to stay at, mode of travel etc. The shop keepers can identify which items are on demand and at what time which thing is best selling. Similarly travel agents can monitor the data looked by customers for preferred destination and can customise their packages according to the prevalent demand. This technique is known as data analytics, and it is one of the biggest trends of 2018.

Now with everything smart be it watches, glasses, television, mobiles and even smart clothes the world is poised to become a huge data collection mechanism. Big Data analytics can help organizations to better understand the information contained within the data and can also help identify the data which is most important to the business and future business decisions. Analysts working with Big Data typically want the knowledge that comes from analyzing the data [6].

3. APPLICATIONS OF BIG DATA

Agriculture, Healthcare industries, educational sector, banking sector, insurance service, industrialization, transportation services all have the use of big data analysis and are tremendously growing.

3.1 Agriculture

India is an agricultural based country and plays an important role in the socio-economic fabric of India. Climate and economy are two major factors on which the production of crop is dependant. An important issue for agricultural is the accurate yield estimation for the various crops involved in the planning. Condition of climate, soil, fertilizers to be used, and commodity prices all these information helps a farmer to make decisions regarding farming. Accurate information regarding the crop yield and risks plays a significant role in the production scheduling of supply chain decision for companies that uses agricultural products as raw materials, food, animal feed, fertilizers, livestock, pesticides, seed and paper.

3.2 Business

Business houses are using big data to lead the competitive landscape by taking advantage of marketing product and services according to the customers need. In order to achieve the desired goal, customer feedback collected using mobile apps, smart phones, e-commerce sites or POS

systems is analysed by which businesses are in position to identify what works, what doesn't and why. Businesses that make strategies to improve productivity or performance depending upon the analysis done on customer experience further help boost revenue growth and customer loyalty. On the other hand if they choose to neglect the indication they not only lose their customers but also have to deal with more data savvy competitor.

To get the desired quality and optimized service process improvements, we need to focus on efficiency gains, optimising the workforce, increasing revenue, savings cost, improving operations and speeding time to manufacture company's product. Understanding Big Data and its advantages will eventually help in transforming a major business process by analysing the data using appropriate analytical tools.

3.3 Healthcare

Cost of healthcare is increasing with the advancement of technology. Big data is a great helping hand in the field of medicine and healthcare. Record of patient's history is required by every physician for accurate diagnosis whereas the after patients successful treatment his history will be stored in the database forever which can only be accessed by his physician or himself. Doctor whenever wants can view a patient's history stored in database either for reference purpose or for further treatment of the said patient. Since large number of medical equipments are data oriented so it is possible for a doctor to prescribes medicine without even visiting patient only by looking into patients history and data provided by remote devices (heart, temperature, BP monitoring watches) attached with the patient at a remote location.

For accurate diagnosis doctors rely on patients clinical record which means collection of vast amount of data from different patients and also from different sources like billing, Electronic medical Records (EMR), path-labs etc., which cannot be done with the old storing techniques of data. Since the data that is being collected is large in variety, from various sources and in different format therefore there is a need for Big Data approach to handle all this.

Thus with the use of big data analytics, the patients outcome can be improved while lowering the cost. Data generated by hospital information system (HIS) is growing at tremendous rate which marks the era of Big data in healthcare and is being used by the healthcare industry for workflow management, in taking good care of patients, in scientific

research and education in medical field. If we timely and appropriately utilize the characteristics of big medical data then it can bring huge benefits in terms of cost, improved quality and better productivity. Because of highly interdisciplinary distributed and complex characteristic of medical data new and efficient technologies like semantic web, cloud computing and data mining are becoming necessary to share, obtain and utilize the information and knowledge imparted by these medical Big data. The ultimate goal of analysis of Big data is to present the result to the user using some visualization tools, but traditional visualization tools such as bar graph, scatter plot and histogram can not reflect the complexity of the results of Big data analysis. Therefore for powerful visual explanation of big data analysis result network steam graph, 3D scatter plot and multi dimensional heat map has been introduced.

3.4 Education

In higher education data is captured and generated in large amount from various sources and in different format. It can be termed big as it contain volume, velocity and variety. The data so generated can be from student's interaction, learning activities and materials, syllabus, result, evaluation, administrative data and other type of data required for innovative teaching methods and procedures for quality improvement. Size and type of the educational data in higher education reflects the requirement for special techniques to be applied in order to gain better insight into student's domain. The techniques are used to improve learning approaches, student's performance and quality of education. Big data analytics also affect the decision making in administrative work and in better utilisation of organizational resources. With the help of Big data analysis of previous years result data and current continuous assessments data we can find out the weak students and can take the measurements in the form of extra coaching classes or guidance to prevent the student from failing in that particular subject. Major part of big data is applicable for investigation into probable improvement in the content of course curriculum.

Big data usage in Students Performance

Student's performance can be measured after analysing the data collected regarding students activities by placing various sensors inside the campus. For students to perform well an assessment of their personal life like time spent in eating, sleeping, studying, environment at home is required for sentiment analysis. Sentiment along with conduct plays a vital role in

shaping students academic life and affecting his grades. To get the desired result of improved student's grade and to reduced fail rate we need to combine the analysis of large student's data sets and IOT technology.

Efficient Teaching

Teachers are require to enhance learning experience by incorporating various innovative teaching methods, thus forcing them to keep pace with the current technologies, new tools and enabling them to be equipped and qualified enough to perform the analysis on vast students data for desired results. To determine student's requirement, learning curve and performance, instantaneous feedback regarding teaching methods, student's social and economic sentiments, effectiveness of course material will help a teacher to take corrective measures teaching methodologies. Analysis of student's data helps a teacher to identify the students which require special attention in academic field as well as personal field. Teacher can employ various methods to boost their morale through appreciation and rewards.

Educational institutions are yet to utilize the full potential of Big data because of lack of computational tools and human resources .Need for students privacy and security of highly confidential data also attributed to slow pace of education institutes in utilizing Big data techniques.

3.5 Banking and Fraud Detection

Comprehensive Analytics

It is required by banks and other financial institutions to build a strong fraud detection strategy.

- Banks detect suspicious patterns after analysing the bank data with the help of necessary analytical tools, notify the customer quickly about fraud incident and place themselves in better position for settlement.
- Analytics look for pattern in data that helps detect fraud. Customers usually have same pattern for deposits, withdraws, using debit and credit cards and personal loan accounts, which analytic algorithm can combine and look for any fraud indicators. Deviation from the usual pattern or withdrawing large amount must be confirmed from the user to avoid any fraud activity. Keeping track of number of bounced cheques, average balances and other customer characteristics for deep analysis helps a bank to detect any potential cheque fraud. Thus in banking sector Big data is implemented and hugely used to find

out any fraud activity. Big data detects the misuse of debit or credit cards, detects unusual patterns in customer statics, detect any credit, debit or other transactional hazard.

- As the technology advances so does the fraudulent attacks although the financial institutions are investing in preventing and capturing the fraud attacks but the nature and extent of fraud keep evolving and changing with the time. Eliminating the current fraud is not an ultimate practical goal of the financial sector, what they need is to change the way fraudulent attackers are dealt with by including real-time detection and behaviour analysis techniques to give new perspective to fraud detection and prevention.
- Analytics allow banks or financial institutions to derive patterns, keep track of any unusual activities and timely locate unusual transactions helpful in preventing any fraudulent activity. Financial institutions are finally waking up to the new ways of fighting fraud with the help of Big data analytics. Big data help financial institution to deal with fraud in different manner and possible get different results too.

Identify Suspicious Activities

Suspicious behaviour is monitored by banks to spot fraudulent activities by analysing real time data. For example, banks usually get notified about the use of card from a device first time, if bank receives multiple notification from a first time usage of device for credit card payment placed at different locations bank confirm about the same from the client and block the card and online transactions in case of fraud usage.

Banks can access data related to customer from various sources such as social sites, call centres conversation, on-line transaction logs which help them in identifying unusual activities. For example -suppose a customer who is credit card holder of a particular bank is on vacation, thus after analysing the data from social networking site, online banking transactions and credit card transaction any attempt on credit card or debit card transaction during the time he is travelling in plane must be blocked

Identify Possible Fraud Attempts

Insurance companies access vast amount of data from various resources like social media, medical record, past claims, phone records and criminal records to examine the claim when received. If something suspicious is being detected the claim is flagged and forwarded for further investigation.

3.6 Traffic Management

Government is planning to incorporate Big data solutions in implementing co-ordinated traffic signals and other variables to improve traffic flow, congestion and fuel consumption and to reduce accidents and passengers safety using new age algorithms and real-time information. Growing focus on the use of big data analytics and development of intelligent networks result in reduced road block, traffic congestion, pollution and assist in managing smooth traffic with fewer safety hazards.

Use of GPS and advanced sensors are designed to help in traffic management and forward alerts to traffic officials for a possible roadblock or any traffic violation. Sensors are usually installed in the vehicles like trucks, ships, planes and provide real time data regarding driver's capabilities and traffic.

GPS signals are used to predict the flow and condition of traffic like estimated delay due to congestion, alternate route suggestion, estimated time to reach the destination etc.

Transport industry handles bulks of data. It has number of equipment's and numerous control models of applications. If anything goes wrong or accident happens, this will put the system in inefficient state and the entire system will shut down.

3.7 Travelling Sector

Predictive analysis is nothing but applying appropriate statistical model to vast amount of data, collected from different sources, to find hidden patterns for prediction and gain insight from the information at our disposal.

- This predictive analysis is used by the travel agencies, airlines, airports etc. to predict the customer behaviour and unlock possibilities for traveller using predictive model.
- Tourism portals, Social networks along with tourist's collaboration are becoming main reference point while planning a trip. We search everything while planning a trip: the type of destination, staying in a hotel, which flight is best to choose etc.
- Predictive analysis helps to find a match between users need and requirements to product and services available in the market. Recommender systems are also used to recommend types of destinations one can visit on his own. Tourism industry now a day's take advantage of technology provided by Big data to continue

expanding and improving services. Today most airlines make use of the Big data to access information about a customer's preference for seat, food and type of disease or constraint they might be suffering in order to serve them better. Because of the maximum adaptability to customers need and improvement in services depending upon the feedback provided Big data is the best ally of Tourism companies.

In today's time, travellers leave digital footprints on the web, social media, post reviews of every hotel stay, every airline reservation and every rental car reservation during their trip also travel companies have lots of data regarding each step of journey planned for customer from flight plans to check-ins. This information can be used by the travelling companies to provide better services so that customers feel appreciated and choose recommend travelling agency which results in more income and profits.

Big data is very powerful and has a very strong positive effect in travel and hospitality industry, consumers expected to see these positive developments in travel industry too.

3.8 Food Industry

It is very difficult for food industry to personalize the food according to the desire of customers as their tastes might vary depending upon their background and age**.** So it's a Big challenge for food industry to utilize data to better serve and understand the need of customers. Use of mobile phones and social media for food safety is an example how Big data is changing the food industry. Restaurants, retailers and food producers everyone is using big data to better understand customer's requirements and unravel market trends related to food. To better understand what customers needs, Big data is helping one of the world's largest business sector, by using sophisticated methods and analytical. Restaurants owners now realizing the importance of Big data and analytics in competitive market, they are optimizing their menu by including revised prices, adding new dishes, modifying the old once according to data collected from customer.

4. SUMMARIZATION

Big Data technology has made insights available in real-time. Agriculture, Healthcare, Education, Financial institutions, Business, Government and even NGOs can benefit by utilizing the insights offered by big data analysis. Big data technology is the new challenge it needs new architecture, techniques, algorithms for its implementation. Also technical

skills are required to analyze the data. So experts are needed to execute Data Mining tools and techniques. Big data presents more opportunities, creates much interest for research and reference in technical progress. Major concern is ensuring big data security and integrity. When collecting data, all measures for security practices must be followed; keeping in mind the complexity and huge size of Big Data.

Table 1: Comparative Analysis of Research Done in Different Fields Using Big Data

S.No.	*Domain*	*Author*	*Brief*	*References*
1.	Agriculture	Jharna Majumdar, Sneha Naraseeyappa and Shilpa Ankalaki	Analysis is done to obtain best yield of wheat product by using various data mining techniques like PAM, CLARA and DBSCAN V to obtain optimum climatic conditions (best temperature, worst temperature and rainfall) for higher production. Out of the three the DBSCAN gives the best analyses of clustering quality metrics.	[18]
2.	Agriculture	Sukhpal Singh Gill, Inderveer Chana, Rajkumar Buyya	Author proposed a Cloud-based automatic information system (AaaS) for agriculture services. The system is used to collect and manage various types of agriculture-related data based on different domains through different sources and classify them using K-NN (k-Nearest Neighbour) classification mechanism in usable clusters of agriculture data.	[26]
3.	Agriculture	Andreas Kamilaris, Andreas Kartakoullis, and Francesc X	Emphasis is given on smart farming which is nothing but use of combination of technologies (IOT, cloud computing, remote sensing) to increase the yield by handling the agricultural challenges in terms of productivity, food security, environmental affect and sustainability. The survey paper provide us with a comparison table informing the reader about how big data is collected from various sources and how its characteristics are being used, which type of	[28]

S.No.	*Domain*	*Author*	*Brief*	*References*
			applications in agriculture are currently using big data and what techniques are employed for analysis. Various problems have been identified along with certain barriers and solutions for these problems were addressed with justification for barriers.	
4.	Agriculture	Iuliia Protopop and Aleksan Shanoyanb	This paper provides the utilization of big data applications and ICT adoption for the potential growth of agriculture in the developing countries. The case studies suggest that with the advancement of technologies (like mobile phones and cloud) provide an infrastructure for data collection and aggregation for development of big data applications.	[27]
5.	Agriculture	Alfons Weersink, Evan Fraser, David Pannell, Emily Duncan and Sarah Rotz	Paper emphasise that at present use of GPS technologies enable us to produce data more rapidly than we can analyze it, since generating data didn't require any specific skills set whereas adoption of variable technologies requires both new skills and decision-making models.	[29]
6.	Agriculture	Bellon-Maurel, Pascal Neveu, Alexandre Termier and Frédérick Garcia	Several barriers exist that are keeping stakeholders from adopting usage of ICT and big data in the agricultural sector. Beyond these barriers there are also certain risks involved from the development of technology in agriculture, such as farmers losing their autonomy, as the control is centralized in the hands of some global firms.	[30]
7.	Business	O.Liu, W.K. Chong, K.L. Man, and C.O. Chan	This paper discussed the application of big data analytics in various business situations and challenges. These business applications not only changed our everyday life but also bring problems related to security and privacy.	[7]

S.No.	*Domain*	*Author*	*Brief*	*References*
8.	Business	Ya Chuan YAO, Qiang HAN, Hong Ying YANG, Xu Yao ZHANG	This paper analysed the different parameters involving large amount of data in production of industry and design a micro cloud based distributed processing system and studies a storage scheme based on cloud computing.	[31]
9.	Banking	Elaheh Radmehr, Mohammad Bazmara	To achieve better and flawless decision making, depending upon the sheer volume of data, available through social or economical aspect of a financial organization, a bank requires strong and effective decision making tools. Business intelligence tools provide the past, present and future perspective in achieving better resource planning and customer relations.	[32]
10.	Education	Christos Vaitsis, Vasilis Hervatis and Nabil Zary	The goal of this chapter is to familiarise the user with the concept of educational data, scientific analytical techniques for data manipulation and how they can be utilized in health education system to exploit big data	[33]
11.	Education	ARCHANA RAO PN, KISHORE BAGLODI	This paper emphasise the use of big data in three main sections of education students, teachers and Institution. Every aspect of the education sector can be harnessed with suitable data mining technologies to predict and understand the relevant sentiments [9]	[20]
12.	Medical	Jing-Song Li, Yi-Fan Zhang and Yu Tian	This paper target the use of data generated from hospital information system (HIS). This data hold great value to patient care and treatment, the workflow management, scientific research, and education in healthcare industry. To improve the predictability of CP this study shows that combining traditional statistical and advanced semantic methods	[21]

S.No.	Domain	Author	Brief	References
			enables timely decision with final goal of improving healthcare quality	
13.	Transportation	Gang Zeng	Architecture of transportation system is proposed based on big data platform, challenges discussed and advantages of big data in transportation were outlined. At last the key technology in ITS is discussed emphasising on the advantages in the future.	[23]
14.	Transportation	Gavin Kemp, Genoveva Vargas-Solar, Catarina Ferreira da Silva, Parisa	This paper propose a service oriented architecture based on the usage of big data and cloud computing in transport applications, hosting several services to be composed to produce a DSS	[34]

This table reveals big data is used everywhere as agriculture, business, banking, education, medical etc. A number of architecture prototypes have been proposed by different researchers to improve big data process.

5. CONCLUSION

This paper includes the applications of big data in such fields of real world which influence the life of a common man for good. There is a growing use of Big data in basic utilities sector like power and water which requires to be explored further to provide better affordable solutions.

Moreover researcher can also focus on the use of big data in improving basic amenities of public and improving the standard of living in developing countries by focusing the use of Big data in fighting diseases. Also the technologies like mobile networking, data mining, multimedia, cloud computing, experts system can be enhanced by the use of big data.

REFERENCES

[1] Dana Boyd (2012). CRITICAL QUESTIONS FOR BIG DATA, Accessed 30 November 15, <http://www.tandfonline.com/doi/pdf/10.1080/1369118X.2012.678878>

[2] Gartner (2015). Technology Research, Accessed 30 November 15, < http://www.gartner.com/technology/home.jsp>

[3] Katal, A. (2013). Big Data: Issues, Challenges, Tools and Good Practices.

[4] Big Data Analytics what it is and why it matters, SAS, accessed 14 May 2018, <https://www.sas.com/en_us/insights/analytics/big-data-analytics.html>

[5] Vangie Beal. (2018), Big Data Analytics, accessed 10 May 2018. <https://www.webopedia.com/TERM/B/big_data_analytics.html>

[6] IBM Big Data and Analytics Hub. (2014). The Four V's of Big Data, Accessed 02 December 15. <http://www.ibmbigdatahub.com/infographic/four-vs-big-data>

[7] Liu, O., Chong, W.K., Man, K.L. and Chan, C.O. (2016). The Application of Big Data Analytics in Business World. In *Proceedings of the International Multi Conference of Engineers and Computer Scientists* Vol II, IMECS.

[8] Gil Press (2016). Top 10 Hot Big Data Technologies, accessed 11 May 2018, <https://www.forbes.com/sites/gilpress/2016/03/14/top-10-hot-big-data technologies/#36896a265d7b>

[9] Waller, M.A. and Fawcett, S.E. (2013). Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34(2), 77–84.

[10] Fan, J., Han, F. and Liu, H. (2014). Challenges of big data analysis. *National Science Review*, 1(2), 293–314.

[11] Joseph, R.C. and Johnson, N.A. (2013). Big data and transformational government. *IT Professional*, 15(6), 43–48.

[12] Bihani, P. and Patil, S.T. (2014). A comparative study of data analysis techniques. *International Journal of Emerging Trends and Technology in Computer Science*, 3(2), 95–101.

[13] Rehman, M.H., Chang, V., Batool, A. and Teh, Y.W. (2016). Big data reduction framework for value creation in sustainable enterprises. *International Journal of Information Management.*

[14] Banerjee, A., Bandyopadhyay, T. and Acharya, P. (2013). Data analytics: hyped up aspirations or true potential. Vikalpa. *The Journal for Decision Makers*, 38(4), 1–11.

[15] Chung, W. (2014). BizPro: Extracting and categorizing business intelligence factors from textual news articles. *International Journal of Information Management*,34(2), 272–284

[16] Hirschberg, J., Hjalmarsson, A. and Elhadad, N. (2010). You're as sick as you sound: Using computational approaches for modelling speaker state to gauge illness and recovery. In A. Neustein (Ed.), *Advances in speech recognition (pp. 305–322).*United States: Springer

[17] Patil, H.A. (2010). Cry baby: Using spectrographic analysis to assess neonatal health status from an infant's cry. In A. Neustein (Ed.), *Advances in speech recognition* (pp. 323–348). United States: Springer.

[18] Majumdar, J., Naraseeyappa, S. and Ankalaki, S. (2017). Analysis of agriculture data using data mining techniques: application of big data. *Journal of Big data.*

[19] Vaitsis, C., Hervatis, V. and Zary, N. (2016). Introduction to Big Data in Education and Its Contribution to the Quality Improvement Processes, chapter 3, INTECH.

[20] Rao, A., Kishore, P.N. and Baglodi (Feb. 2018). Role of Big Data in Education Sector: A Review. *International Journal of Advances in Science Engineering and Technology*, ISSN(p): 2321 –8991, ISSN(e): 2321-9009Vol-6.

[21] Li, J.S., Zhang, Y. F., and Tian, Y. Medical Big Data Analysis in Hospital Information System. chapter4, INTECH.

[22] Yilin, T., Caofang, G. and Rui, W. (2018). Big Data in Medical Research and Application. *Web of Conferences* 176, 03017 MATEC.

[23] Zeng, G. (Jan. – Feb. 2015). Application of Big Data in Intelligent Traffic System. *Journal of Computer Engineering.* (pp 01–04). IOSR.

[24] Rahaman, A., Rajesh, S. and Rani, G. (2018). Challenging tools on Research Issues in Big Data Analytics. *International Journal of Engineering Development and Research.* IJEDR.

[25] Naganathan, V. (May-2018). Comparative Analysis of Big Data, Big Data Analytics: Challenges and Trends. *International Research Journal of Engineering and Technology.*

[26] Gill, S.S., Chana, I. and Buyya, R. (December-2017). IoT Based Agriculture as a Cloud and Big Data Service: The Beginning of Digital India**.** *Journal of Organizational and End User Computing Volume 29, Issue 4.*

[27] Protopopa, I. and Shanoyanb, A. (2016). Big Data and Smallholder Farmers: Big Data Applications in the Agri-Food Supply Chain in Developing Countries. *International Food and Agribusiness Management Review Special Issue, Volume 19, Issue A.*

[28] Kamilaris, A., Kartakoullis, A. and Francesc X. A Review on the Practice of Big Data Analysis in Agriculture. *Prenafeta-Boldú GIRO Joint Research Unit IRTA-UPC*, Barcelona, Spain.

[29] Weersink, A. Fraser, E., Pannell, D., Duncan, E. and Rotz, S. (2018). Opportunities and Challenges for Big Data in Agricultural and Environmental Analysis. *Annual Review of Resource Economics*.

[30] Maurel, V.B., Neveu, P., Termier, A. and Garcia, F. Big data in Agriculture. *Institut Convergences Agriculture Numérique #DigitAg.*

[31] Yao, Y.C., Han, Q., Yang, Y.Y. and Zhang, X.Y. (2018). The application of big data in production and environment. *First International Conference on Advanced Algorithms and Control Engineering. IOP Conf. Series: Journal of Physics*.

[32] Radmehr, E. and Bazmara, M. (Jan-2017). A Survey on Business Intelligence Solutions in Banking Industry and Big Data Applications. IJMEC.

[33] Vaitsis, C., Hervatis, V. and Zary, N. Introduction to Big Data in Education and Its Contribution to the Quality Improvement Processes. http://dx.doi.org/10.5772/63896.

[34] Kemp, G., Vargas-Solar, G., Ferreira da Silva, C., Ghodous, P., Collet, C., Amaya, P.O. (2016). Cloud big data application for transport. HAL .<https://hal.archives-ouvertes.fr/hal-01374885>

Adaptive Control and Hybrid Projective Combination Synchronization of Chaos Generated by Generalized Lotka-Volterra Biological Systems

Ayub Khan[1] and Harindri Chaudhary[2]

ABSTRACT: *In this manuscript, we study hybrid projective combination synchronization (HPCS) between chaotic generalized three species Lotka-Volterra (GLV) biological system using adaptive control method. Lotka and Volterra introduced the well-known examples of basic biological models of two species predator and prey model. This system explains the interaction between two species predator-prey model and it contains a system of ordinary differential equations. Though it is a very famous model, it has many drawbacks, for instance, it avoids numerous essential characteristics such as interplay among another species of the similar ecosystem, connection with the environment and so on. Samardzija and Greller studied the dynamics and the chaotic analysis of GLV for the first time in the year 1988. Ever since the area of biological control in various biological systems existing in nature has been an active field for researchers and biologists. Here, we consider two predators and one prey population present in the biological system. Initially, this paper exhibits the phase portraits of GLV system, which is a 3-D chaotic system. The adaptive biological control law for attaining global and asymptotic hybrid projective chaos synchronization of the state variables of the system with uncertain parameters has been derived using Lyapunov stability theory (LST). Also, it is noted that both complete synchronization and anti-synchronization coexist in hybrid projective combination synchronization. Furthermore, numerical simulations are performed to show the efficiency and accuracy of the suggested scheme using MATLAB. Significantly, the analytical results are in*

[1,2] Department of Mathematics, Jamia Millia Islamia, New Delhi, india.
E-mail: [1]akhan12@jmi.ac.in; [2]harindri20dbc@gmail.com

excellent agreement with computational results. The considered approach has many applications in the field of secure communication and image encryption.

Keywords: Chaotic System, Hybrid Projective Synchronization, Combination Synchronization Lotka-Volterra System, Adaptive Control.

1. INTRODUCTION

Chaos theory is defined as the qualitative analysis of extremely complex nonlinear dynamical systems found in natural environment possessing inherent properties such as infinite recurrence, boundedness and high sensitivity to initial conditions. Chaos theory has a variety of applications in numerous fields of engineering and applied sciences, for instance, biomedical engineering [1], neural networks [2], weather models [3], ecological models [4], jerk systems [5], encryption [6], etc.

E.N. Lorenz [7] firstly discovered the most famous 3-D chaotic system in 1963 while studying a weather model for atmospheric convention. Consequently, Rossler [8] found an algebraically much simplified 3-D chaotic system other than the Lorenz system in 1976. These classical systems were ensued by the origination of several 3-D chaotic systems, for instance, Chen system, Sprott system, Lu-Chen system, Cai system, etc.

Generalized Lotka-Voltera (GLV) three-species model is the most prominent among all the existing biological models. Arnedo *et al*. [9] have demonstrated that three species can possess chaotic behavior in an ecosystem. Samardzija and Greller [10] studied GLV three species biological model in detail.

Pecora and carroll [11] firstly introduced the synchronization of chaotic dynamical systems in 1990 by applying master-slave configuration. Chaos synchronization is described as a procedure of adjusting identical or non-identical chaotic systems in such a way that both depict the same behavior owing to coupling to achieve stability. Till date, various types of synchronization techniques for chaotic systems have been developed such as complete synchronization [12], anti-synchronization [13], hybrid synchronization [14], hybrid projective synchronization [15], lag synchronization [16], phase synchronization [17] etc.

Up to now, numerous schemes have been introduced in control theory to stabilize the chaos including the active control [18], adaptive control [19], backstepping design [20], feedback control [21],sliding mode control [22],

etc. In the available literature, for an integer order dynamical system to be chaotic, the dimension of the state equations should be at least three and at least one Lyapunov exponent must be positive.

Chaos synchronization and control among chaotic systems using adaptive control scheme was firstly advocated by Hubler in 1989. [23]. Mainieri and Rehacek [24] in 1999 introduced the concept of projective synchronization for synchronizing the chaotic systems. In [25], synchronization of a modified Chua's circuit system using adaptive control method has been studied. Moreover, projective synchronization and chaos in secure communication are studied in [26]. Also, in [27], adaptive backstepping technique for synchronizing chaotic systems has been investigated. In [28] and [29], ACM is discussed to synchronize the generalized three species Lotka-Volterra biological system.

Considering the above mentioned discussions, this paper aims to investigate a hybrid projective combination synchronization (HPCS) between three identical GLV systems by using ACM. Adaptive control technique is very important to estimate the parameters between master and slave systems. Also, based on LST, we examime in detail an adaptive control law together with an estimated parameter update laws.

The description of the present paper is as follows: Section 2 deals with the preliminaries containing some notations and essential terminology to be used in the upcoming sections. Section 3 contains the basic structured properties of the considered GLV system. Section 4 investigates the ACM along with a parameter estimation update laws to stabilize asymptotically and globally the given GLV system. Section 5 contains the numerical simulations illustrating the effectiveness and feasibility of the considered HPCS approach. Lastly, Section 6 concludes the paper.

2. PRELIMINARIES

In this section, we describe the methodology to achieve combination synchronization according to master-slave configuration among three identical integer order chaotic systems with two master systems and one slave system which are requisites for the upcoming sections.

Two master systems and one slave system are considered as:

$$\dot{w}_1 = g_1(w_1)\eta_1 + G_1(w_1) \quad \ldots (1)$$

$$\dot{w}_2 = g_2(w_2)\eta_2 + G_2(w_2) \quad \ldots (2)$$

$$\dot{z}_1 = h_1(z_1)\theta_1 + H_1(z_1) + u_1 \quad \ldots (3)$$

where $w_1 = (w_{11}, w_{12}, \ldots, w_{1n})^T \in R^n$, $w_2 = (w_{21}, w_{22}, \ldots, w_{2n})^T \in R^n$, $z_1 = (z_{11}, z_{12}, \ldots, z_{1n})^T \in R^n$ are the state vectors of master and slave systems (1), (2) and (3) respectively, $G_1, G_2, H_1 : R^n \to R^n$ are three nonlinear continuous functions, $\eta_1 = (\eta_{11}, \eta_{12}, \ldots, \eta_{1p_1})^T$ is a $p_1 \times 1$ unknown parameter vector of the first master system (1), $\eta_2 = (\eta_{21}, \eta_{22}, \ldots, \eta_{2p_2})^T$ is a $p_2 \times 1$ unknown parameter vector of the second master system (2), $\theta_1 = (\theta_{11}, \theta_{12}, \ldots, \theta_{1p_3})^T$ is a $p_3 \times 1$ unknown parameter vector of the slave system (3), $g_1 : R^n \to R^{n \times p_1}$, $g_2 : R^n \to R^{n \times p_2}$, $h_1 : R^n \to R^{n \times p_3}$, $u_1 = (u_{11}, u_{12}, \ldots, u_{1n}) \in R^n$ is the appropriate controller which is to be designed.

Definition 1: If there exist three constant matrices $P, Q, R \in R^n \times R^n$ and $R \neq 0$ such that:

$$lim_{t \to \infty} \| E \| = lim_{t \to \infty} \| (Pw_1 + Qw_2 - Rz_1) \| = 0$$

the combination of two master systems (1) and (2) is said to perform combination synchronization with one slave system (3) and $\|.\|$ denotes matrix norm.

Remark 2.1: The constant matrices P, Q and R are called the scaling matrices. Moreover,P, Q and R can be extended as matrices of functions of state variables w_1, w_2 and z_1.

Remark 2.2: The problem of combination synchronization will be converted into the chaos control issue for $P = Q = 0$.

Remark 2.3: If $R = -I$ and $P = Q = -\beta I$, for $\beta = 1$ it turns into combination completes synchronization and for $\beta = -1$ it will be reduced to combination anti-synchronization. Combination of complete and anti- synchronization makes hybrid projective combination synchronization (HPCS). Thus the HPCS error takes the form :

$$E = z_1 - \beta(w_1 + w_2), \quad \ldots (4)$$

where $\beta = \text{diag}(\beta_1, \beta_2, \ldots, \beta_n)$.

Remark 2.4: Definition 1 displays that the combination of master systems and slave systems may be expanded to more systems. Further, master systems and slave systems of the combination synchronization scheme can be identical or non-identical.

3. DESCRIPTION OF 3-D GLV CHAOTIC SYSTEM

Introduced by Samardzija *et al*. [10] in 1988, the considered chaotic system famously known as Generalized Lotka-Volterra (GLV) three species biological system is described as:

$$\begin{cases} \dot{x}_{11} = & x_{11} - x_{11}y_{11} + n_1 x_{11}^2 - l_1 x_{11}^2 z_{11} \\ \dot{y}_{11} = & -y_{11} + x_{11}y_{11} \\ \dot{z}_{11} = & m_1 z_{11} + l_1 x_{11}^2 z_{11} \end{cases} \quad \ldots (5)$$

where $(x_{11}, y_{11}, z_{11})^T \in R^3$ is the state vector and l_1,m_1 and n_1 are positive parameters. Also, in (5), x_{11} represents the prey population and y_{11}, z_{11} denote the predator populations. For $l_1 = 2.9851$, $m_1 = 3$ and $m_1 = 3$ and $n_1 = 2$,the system (5) depicts chaotic behaviour. Also, Figure 1(a-d) display the phase plots of (5). Moreover, the detailed theoretical study and numerical simulation results for (5) can be found in [10].

4. ILLUSRATIVE EXAMPLE

In this section, we discuss HPCS scheme to obtain the laws in order to estimate parameters together with adaptive laws in such a way that all the state variables x_{11}, y_{11} and z_{11} tend to equilibrium points as t approaches infinity.

Here, the two master systems and the corresponding slave system with controllers can be defined as:

$$\begin{cases} \dot{x}_{11} = & x_{11} - x_{11}y_{11} + n_1 x_{11}^2 - l_1 x_{11}^2 z_{11} \\ \dot{y}_{11} = & -y_{11} + x_{11}y_{11} \\ \dot{z}_{11} = & m_1 z_{11} + l_1 x_{11}^2 z_{11} \end{cases} \quad \ldots (6)$$

$$\begin{cases} \dot{x}_{21} = & x_{21} - x_{21}y_{21} + n_1 x_{21}^2 - l_1 x_{21}^2 z_{21} \\ \dot{y}_{21} = & -y_{21} + x_{21}y_{21} \\ \dot{z}_{21} = & m_1 z_{21} + l_1 x_{21}^2 z_{21} \end{cases} \quad \ldots (7)$$

$$\begin{cases} \dot{x}_{31} = & x_{31} - x_{31}y_{31} + n_1 x_{31}^2 - l_1 x_{31}^2 z_{31} + u_{11} \\ \dot{y}_{31} = & -y_{31} + x_{31}y_{31} + u_{12} \\ \dot{z}_{31} = & m_1 z_{31} + l_1 x_{31}^2 z_{31} + u_{13}, \end{cases} \quad \ldots (8)$$

where u_{11},u_{12} and u_{13} are adaptive controllers to be designed in such a manner that HPCS between three identical chaotic systems will be obtained.

Define now the state errors as

$$\begin{cases} E_{11} = & x_{31} - \beta_1(x_{11} + x_{21}) \\ E_{12} = & y_{31} - \beta_2(y_{11} + y_{21}) \\ E_{13} = & z_{31} - \beta_3(z_{11} + z_{21}) \end{cases} \quad \ldots (9)$$

The main goal of this paper is to introduce controllers $u_{1i}, (i = 1,2,3)$ so that the state errors defined in (9) satisfy

$$lim_{t\to\infty} E_{1i}(t) = 0 \; for \; (i = 1,2,3).$$

The consequent error dynamics is transformed as:

$$\begin{cases} \dot{E}_{11} = & E_{11} - x_{31}y_{31} + n_1x_{31}^2 - l_1x_{31}^2z_{31} - \beta_1(x_{21}y_{21} + n_1x_{21}^2 \\ & -l_1x_{21}^2z_{21} - x_{11}y_{11} + n_1x_{31}^2 - l_1x_{11}^2z_{11}) + u_{11} \\ \dot{E}_{12} = & -E_{12} + x_{31}y_{31} - \beta_2(x_{11}y_{11} + x_{21}y_{21}) + u_{12} \\ \dot{E}_{13} = & -m_1E_{13} + l_1x_{31}^2z_{31} - \beta_3(l_1x_{21}^2z_{21} + l_1x_{11}^2z_{11}) + u_{13} \end{cases} \quad \ldots (10)$$

Now, we define the adaptive controllers as:

$$\begin{cases} u_{11} = & -E_{11} + x_{31}y_{31} - \hat{n}_1x_{31}^2 + \hat{l}_1x_{31}^2z_{31} + \beta_1(x_{21}y_{21} + \hat{n}_1x_{21}^2 \\ & -\hat{l}_1x_{21}^2z_{21} - x_{11}y_{11} + \hat{n}_1x_{31}^2 - \hat{l}_1x_{11}^2z11) - K_1E_{11} \\ u_{12} = & E_{12} - x_{31}y_{31} + \beta_2(x_{11}y_{11} + x_{21}y_{21}) - K_2E_{12} \\ u_{13} = & \hat{m}_1E_{13} - \hat{l}_1x_{31}^2z_{31} + \beta_3(\hat{l}_1x_{21}^2z_{21} + \hat{l}_1x_{11}^2z_{11}) - K_3E_{13} \end{cases} \quad \ldots (11)$$

where K_1, K_2 and K_3 are positive gain constants

By substituting the values of controllers (11) in error dynamics (10), we get

$$\begin{cases} \dot{E}_{11} = & (n_1 - \hat{n}_1)x_{31}^2 - (l_1 - \hat{l}_1)x_{31}^2z_{31} - \beta_1\,[(n_1 - \hat{n}_1)x_{21}^2 \\ & -(l_1 - \hat{l}_1)x_{21}^2z_{21} + (n_1 - \hat{n}_1)x_{11}^2 - (l_1 - \hat{l}_1)x_{11}^2z_{11}] - K_1E_{11} \\ \dot{E}_{12} = & -K_2E_{12} \\ \dot{E}_{13} = & -(m_1 - \hat{m}_1)z_{31} + (l_1 - \hat{l}_1)x_{31}^2z_{31} + \beta_3\,[(m_1 - \hat{m}_1)z_{11} \\ & -(l_1 - \hat{l}_1)x_{21}^2z_{21} + (m_1 - \hat{m}_1)z_{11} - (l_1 - \hat{l}_1)x_{11}^2z_{31}] - K_3E_{13} \end{cases} \quad \ldots (12)$$

where $\hat{l}_1$, $\hat{m}_1$ and $\hat{n}_1$ are estimated values of unknown parameter l_1, m_1 and n_1 respectively.

Now defining parameter estimation error as:

$$\tilde{l}_1 = l_1 - \hat{l}_1, \tilde{m}_1 = m_1 - \hat{m}_1, \tilde{n}_1 = n_1 - \hat{n}_1, \quad \ldots (13)$$

Applying (13), the error dynamics (12) simplifies as:

$$\begin{cases} \dot{E}_{11} = & \tilde{n}_1(x_{31}^2 - \beta_1 x_{21}^2 _ \beta_1 x_{11}^2) - \tilde{l}_1(x_{31}^2 z_{31} - \beta_1 x_{21}^2 z_{21} - \beta_1 y_{11} z_{11}) - K_1 E_{11} \\ \dot{E}_{12} = & -K_2 E_{12} \\ \dot{E}_{13} = & -\tilde{m}_1 E_{13} + \tilde{l}_1(x_{31}^2 z_{31} - \beta_3 x_{21}^2 z_{21} - \beta_3 x_{11}^2 z_{11}] - K_3 E_{13} \end{cases}$$

The derivative of the parameter estimation error (13), we get

$$\dot{\tilde{l}}_1 = -\dot{\tilde{l}}_1, \dot{\tilde{m}}_1 = -\dot{\tilde{m}}_1, \dot{\tilde{n}}_1 = -\dot{\tilde{n}}_1 \quad \ldots (14)$$

The Lyapunov function is constructed as

$$V(E(t)) = \frac{1}{2}\left[E_{11}^2 + E_{12}^2 + E_{13}^2 + \tilde{l}_1{}^2 + \tilde{m}_1{}^2 + \tilde{n}_1{}^2\right] \quad \ldots (15)$$

which implies that $V(E(t))$ is positive definite .

Derivative of the Lyapunov function $V(E(t))$ is given by

$$\dot{V}(E(t)) = E_{11}\dot{E}_{11} + E_{12}\dot{E}_{12} + E_{13}\dot{E}_{13} - \tilde{l}_1\dot{\tilde{l}}_1 - \tilde{m}_1\dot{\tilde{m}}_1 - \tilde{n}_1\dot{\tilde{n}}_1 \quad \ldots (16)$$

Keeping (16) in view, we describe the parameter estimates laws by the following rule:

$$\begin{cases} \dot{\tilde{l}}_1 = & -(x_{31}^2 z_{31} - \beta_1 x_{21}^2 z_{21} - \beta_1 x_{11}^2 z_{11})E_{11} + (x_{31}^2 z_{31} - \beta_3 x_{21}^2 z_{21} \\ & -\beta_3 x_{11}^2 z_{11})E_{13} + K_4 \tilde{l}_1 \\ \dot{\tilde{m}}_1 = & -E_{13} + K_5 \tilde{m}_1 \\ \dot{\tilde{n}}_1 = & x_{31}^2 - \beta_1 x_{11}^2 - \beta_1 x_{21}^2 + K_6 \tilde{n}_1, \end{cases} \quad \ldots (17)$$

where K_4, K_5 and K_6 are positive gain constants.

Theorem 1: The chaotic systems (6)-(8) are globally and asymptotically hybrid projective combination synchronized for each initial states $(x_{11}(0), y_{11}(0), z_{11}(0)) \in R^3$ by the adaptive controller (11) and the parameter update law (17).

Proof: The Lyapunov function $V(E(t))$ as defined in (15) is a positive definite function in R^6. On simplifying equations (16) and (17),we have

$$\dot{V}(E(t)) = -K_1 e_{11}^2 - K_2 e_{12}^2 - K_3 e_{13}^2 - K_4 \tilde{l}_1{}^2 - K_5 \tilde{m}_1{}^2 - K_6 \tilde{n}_1{}^2 < 0$$

which establishes that $\dot{V}(E(t))$ is negative definite.

Hence, by using LST, we deduce that the HPCS errror $E(t) \to 0$ asymptotically as $t \to \infty$ for each initial conditions $E(0) \in R^3$. This ends the proof.

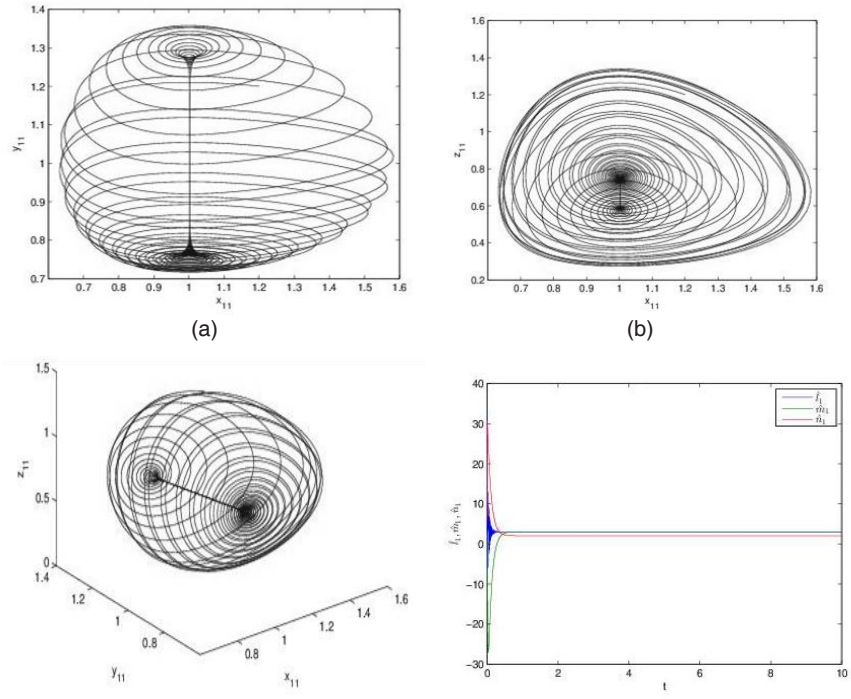


Fig. 1: Phase Plots of 3-D GLV Chaotic System in (a) $x_{11} - y_{11}$ Plane, (b) $y_{11} - z_{11}$ Plane, (c) $x_{11} -$ y11-z11 Space, (d) Parameter Estimation

5. NUMERICAL SIMULATION

In this section, we present some simulation experiments to illustrate the effectiveness and feasibility of the proposed HPCS scheme using ACM. To achieve this, we use the fourth-order Runge-Kutta method to solve systems of ordinary differential equations. The parameters of the given system are taken as $l_1 = 2.9851$, $m_1 = 3$ and $n_1 = 2$ to show that the system behaves chaotically in the absence of control inputs. The initial states of the master and slave systems are $(x_{11}(0) = 27.5, y_{11}(0) = 23.1, z_{11}(0) = 11.4$, $(x_{21}(0) = 1.2, y_{21}(0) = 1.2, z_{21}(0) = 1.2$ and $(x_{31}(0) = 2.9, y_{31}(0) = 12.8, z_{31}(0) = 20.3$ respectively.

We achieve hybrid projective synchronization between master and slave systems by choosing the scaling matrix β with $\beta_1 = 2, \beta_2 = -4, \beta_3 = 3$. Here, the control gains are taken as $K_i = 6$ for $i = 1, 2, 3$. The numerical simulations are exhibited in Fig. 2(a-c) which display the state trajectories of master and slave systems. The synchronization errors $(e_{11}, e_{12}, e_{13}) =$

$(-54.5,110,-17.5)$ approaches zero as t approaches infinity in Figure 2(d). Moreover, Figure 1(d) depicts that estimated values $(\hat{l}_1, \hat{m}_1, \hat{n}_1$ of unknown parameters converge respectively to their original values asymptotically with time. Therefore, the proposed HPCS scheme between master and slave system is verified computationally.

6. CONCLUSION

In this paper, we have investigated the proposed HPCS scheme between identical chaotic GLV system via adaptive control technique. By designing proper control inputs based on LST, the considered HPCS technique is achieved. The particular cases of anti-synchronization, complete synchronization, and chaos control problem are also discussed. The effectivity and feasibility of the theoretical results are justified in simulations by using MATLAB. Remarkably, the theoretical analysis and the numerical results both are in complete agreement. In fact, the considered HPCS scheme has potential and advantages as it has various applications in encryption and secure communication. In our investigation, the time taken by the synchronization error converging to zero is less in comparison with other related published work. We noticed that our discussed technique is basic yet theoretically rigorous. Moreover, we understand that the investigated HPCS scheme may be generalized by using other control techniques.

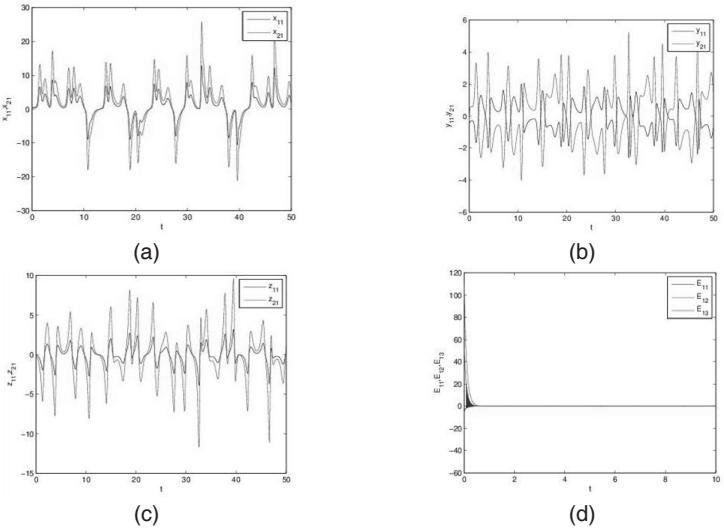


Fig. 2: HPCS Trajectories of GLV System (a) between $x_{11}(t) - x_{21}(t)$, (b) between $y_{11}(t) - y_{21}(t)$, (c) between $z_{11}(t) - z_{21}(t)$, (d) HPCS Synchronization Error

REFERENCES

[1] Provata, A., Katsaloulis, P. and Verganelakis, D.A. (2016). Dynamics of chaotic maps for modelling the multifractal spectrum of human brain diffusion tensor images. In *Chaos, Solitons & Fractals*, 45(2):174–180.

[2] Bouallegue, K. (2017). A new class of neural networks and its applications. *In Neurocomputing*, 249:28–47.

[3] Russell, F.P., Düben, P.D., Niu, X., Luk, W. and Palmer, T.N. (2017). Exploiting the chaotic behaviour of atmospheric models with reconfigurable architectures. In *Computer Physics Communications*, 221:160–173.

[4] Sahoo, B. and Poria, S. (2014). The chaos and control of a food chain model supplying additional food to top-predator. In *Chaos, Solitons & Fractals*, 58:52–64.

[5] Wang, X., Vaidyanathan, S., Volos, C., Pham, V.T. and Kapitaniak, T. (2017). Dynamics, circuit realization, control and synchronization of a hyperchaotic hyperjerk system with coexisting attractors. In *Nonlinear Dynamics*, 89(3):1673–1687.

[6] Wu, G.C., Baleanu, D. and Lin., Z.X. (2016). Image encryption technique based on fractional chaotic time series. In *Journal of Vibration and Control*, 22(8):2092–2099.

[7] Lorenz, E.N. (1963). Deterministic nonperiodic flow. In *Journal of the atmospheric sciences*, 20(2):130–141.

[8] Rössler,O.E. (1976). An equation for continuous chaos. In *Physics Letters A*, 57(5):397–398.

[9] Arneodo, A., Coullet, P. and Tresser, C. (1980). Occurence of strange attractors in three-dimensional volterra equations. In *Physics Letters A*, 79(4):259–263.

[10] Samardzija, N. and Greller, L.D. (1988). Explosive route to chaos through a fractal torus in a generalized lotka-volterra model. In *Bulletin of Mathematical Biology*, 50(5):465–491.

[11] Pecora, L.M. and Carroll, T.L. (1990). Synchronization in chaotic systems. *In Physical review letters*, 64(8):821.

[12] Singh, A.K., Yadav, V.K. and Das, S. (2017). Synchronization between fractional order complex chaotic systems. In *International Journal of Dynamics and Control*, 5(3):756–770.

[13] Li, G.H. and Zhou,S.P. (2007). Anti-synchronization in different chaotic systems. *In Chaos, Solitons & Fractals*, 32(2):516–520.

[14] Sudheer, K.S. and Sabir, M. (2009). Hybrid synchronization of hyperchaotic lu system. In *Pramana*, 73(4):781.

[15] Zhenyuan X.uaLiuxiaoGuoac Manfeng Huac, Yongqing Yang (2008). Hybrid projective synchronization in a chaotic complex nonlinear system. In *Mathematics and Computers in Simulation*, 79(1):449–457.

[16] Li, C. and Liao, X. (2004). Complete and lag synchronization of hyperchaotic systems using small impulses. In *Chaos, Solitons & Fractals*, 22(4):857–867.

[17] Ma, J., Mi, L., Zhou, P., Xu, Y. and Hayat, T. (2017). Phase synchronization between two neurons induced by coupling of electromagnetic field. In *Applied Mathematics and Computation*, 307:321–328.

[18] Delavari, H. and Mohadeszadeh, M. (2018). Hybrid complex projective synchronization of complex chaotic systems using active control technique with nonlinearity in the control input. In *Journal of Control Engineering and Applied Informatics*, 20(1):67–74.

[19] Khan, A. and Bhat, M.A. (2017). Hyper-chaotic analysis and adaptive multi-switching synchronization of a novel asymmetric non-linear dynamical system. *In International Journal of Dynamics and Control*, 5(4):1211–1221.

[20] Rasappan, S. and Vaidyanathan, S. (2012). Synchronization of hyperchaotic liu system via backstepping control with recursive feedback. In *International Conference on Eco-friendly Computing and Communication Systems*, pages 212–221. Springer.

[21] Chen, M. and Han, Z. (2003). Controlling and synchronizing chaotic genesio system via nonlinear feedback control. In *Chaos, Solitons & Fractals*, 17(4):709–716.

[22] Vaidyanathan, S. and Sampath, S. (2012). Anti-synchronization of four-wing chaotic systems via sliding mode control. In *International Journal of Automation and Computing*, 9(3):274–279.

[23] Hubler, A.W. (1989). Adaptive control of chaotic system. In *Helv Phys Acta*, 62:343–346.

[24] Mainieri, R. and Rehacek, J. (1999). Projective synchronization in three-dimensional chaotic systems. In *Physical Review Letters*, 82(15):3042.

[25] Yassen, M.T. (2003). Adaptive control and synchronization of a modified Chua's circuit system. In *Applied Mathematics and Computation*, 135(1):113–128.

[26] Li, Z. and Xu, D. (2004). A secure communication scheme using projective chaos synchronization. In *Chaos, Solitons & Fractals*, 22(2):477–481.

[27] Li, S.Y., Yang, C.H., Lin, C.T., Ko, L.W. and Chiu, T.T. (2012). Adaptive synchronization of chaotic systems with unknown parameters via new backstepping strategy. *In Nonlinear Dynamics*, 70(3):2129–2143.

[28] Vaidyanathan, S. (2016). Hybrid synchronization of the generalized lotka-volterra three-species biological systems via adaptive control. In *Int J PharmTech Res*, 9(1):179–192.

[29] Vaidyanathan, S. (2015). Adaptive biological control of generalized lotka-volterra three-species biological system. In *International Journal of PharmTech Research*, 8(4): 622–631.

E-Leaning:
Emerging Trend in Learning

Diksha Grover[1] and Shweta Paradkar[2]

ABSTRACT: *In a fast developing nation like India where the majority of the population comprises of young adults, a strong education system is must for the overall growth and development of the nation. Unfortunately, our educational planners and policy makers have not succeed in providing quality education and the required infrastructure to the remote locations, thus depriving the youth of the country the basic education. Thus our educational planners are in the view that E-learning can overcome this problem through effective utilization of the resources in a cost effective manner. E-learning has a larger outreach and thus can have a larger audience. Today, integration of digital technology in all areas of businesses and technology is changing how enterprises offering corporate training. As digital transformation has made the enterprises to think how employees should be trained. Today the enterprises are gradually moving from the traditional classroom-based training approach to technology-enabled eLearning instead. The percentage of on line learning is increasing day by day as E-Learning is becoming very popular at different places of work. Whether E-Learning is the solution for all the businesses so in this regard we have discussed that enterprises providing corporate training should consider the advantages and disadvantages of eLearning when seeking out a solution to meet the challenges of business world. This paper review one case study that provides a direct comparison of direct classroom teaching and teaching through E-Learning environment. The study was conducted to know whether there is any major difference in students learning and their performance when learning is through E-learning based approach and through classroom lecture based learning approach. In addition*

[1] Assistant Professor, Kalindi College, University of Delhi, India.
E-mail: dik.grover@gmail.com
[2] Assistant Professor, Jesus and Mary College University of Delhi, India.
E-mail: shweta.paradkar@rediffmail.com

to this a survey was conducted in this study which allows the students to give their feedback and suggestions on both the learning approaches.

Keywords: Digital Transformation, Infrastructure, Integration, Simulation, Tablets.

1. INTRODUCTION TO E-LEARNING

E-Learning a platform to provide skills and knowledge with the help of technology enablers i.e. computer, smart devices and internet networks. E-Learning consists of applications and processes that includes internet based learning, computer based learning and training, virtual learning prospects and digital collaboration. E-Learning Content can be self paced or instructor controlled and it can be in the form of text, audio, image, streaming video and audio. E-Learning can be used by a wide range of people who have difficulty in attending conventional classroom training due to many reasons such as geographically dispersed with limited time and resources, restricted from participating in classroom sessions because of cultural and religious beliefs, located in conflict and post conflict areas.

1.1 Techniques and "Tools of E-Learning"

1. **Smart Class Room:** Smart classroom an E-learning tool, aims at developing the students learning ability and in order to grow academically teacher can use different methods of teaching and learning are necessary to inculcate deeper level of growth and conceptual development. It is a step to the future of education. Teachers can choose variety of technologies including slide presentation, Audio and video conferencing, Application sharing, Shared whiteboard etc.
2. **Use of Technology in Education in Rural and Isolated Communities:** The increased availability of low cost video cameras those in mobile phones helps teachers to take short videos of their peers and then jointly review and discuss pedagogical approaches and specially difficult topics to teach in informal way as a part of their professional development. Lack of resources (including textbooks and other teaching materials), promoting more use of mobile phones helping teachers in small but meaningful ways by providing access to education content. Generally the places where Internet connectivity is unreliable innovative approaches to caching and distributing digital content can enable off-line access to vast numbers of online resources enabling the students to access and read on their small devices.

E-Shalas provide education to children in rural areas, mostly in areas where power supply doesn't exist and literacy levels are low. High school students sit in a smart classroom and learn their lessons on an entirely latest solar-powered LED projector. The syllabus of e-Shalas is available in both Kannada and English, with a voice-over making the lessons self-explanatory. The visual instruction encapsulate the content knowledge and can be accessed via devices such as Android mini-PCs, tablets and mobile phones. The teachers are given a tablet loaded with the same content, so they are well-versed with the lessons before they run it in the classroom. This unique projector has an android mini-computer, speakers, power supply panel and remote mouse fitted to it, which makes it energy-efficient. As against the normal LED projector's power consumption of 250-watt, this projector consumes only 25 watts. The solar panel on the building terrace is a single 75-watt power panel and the entire system costs less than INR 1,00,000 (one-time investment).

Fig. 1: e-Shala Solar Powered Classes in Addagal Rural School (Karnataka)

3. **Computer Based Training Tools:** In today's digital world, we need CBT Tools to upload or create learning materials for effective students learning and can also provide corporate training to your employees by providing training material or by creating your own training site. For example **iSpring Learn** is a cloud-based LMS. In iSpring Learn, you can upload an unlimited number of training materials (presentations, text documents, video and audio files, as well as SCORMcourses).

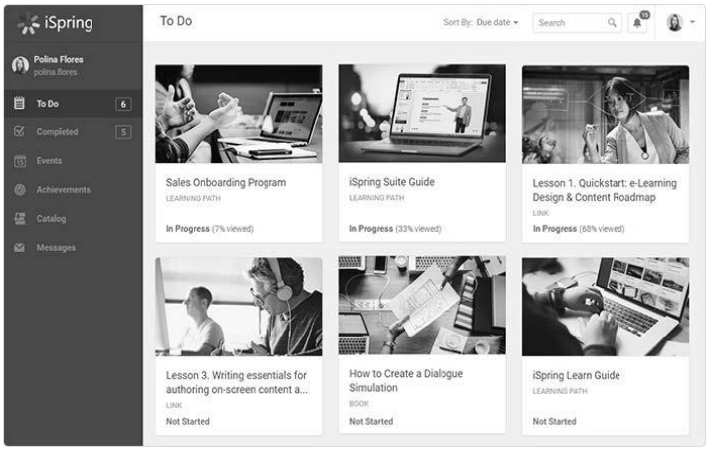


Fig. 2: Ispring (a cloud based LMS)

Learn Cube is a **virtual classroom learning tool** that offers live video and Audio support, various classroom communication tools, class scheduling, and media sharing. During online classes, teachers can interact with learners via video and text chat, use a whiteboard for explaining the learning material, and share content including PDFs, audio files, videos, and measure their performance using online quizzes and tests.

2. E-LEARNING—A DIGITAL TRAINING PLATFORM FOR CORPORATES

Integration of digital technology in all areas of businesses and technology is also changing how enterprises offering corporate training. As digital transformation has made the enterprises to think how employees should be trained and organizations have recognized the training an employee is much easier and less expensive than replacing them. Today the enterprises are gradually moving from the traditional classroom-based training approach to technology-enabled eLearning instead. The percentage of on line learning is increasing day by day as E-Learning is becoming very popular at different places of work.

- The main advantage of providing training to the employs using **E-Learning** as compared to traditional methods of training it can be delivered fast and employee can learn anywhere and at anytime regardless of geography and time zones.
- **Training through E-learning** is less expensive as compared to the traditional tainting that must be developed in house in a classroom

style format where a dedicated staff, room and devices are required whereas E-Learning can be instituted immediately.

- **E-Learning** is tailor-made as it ensures that each employee learns the skills that should matched with the business needs of the organization.
- **E-Learning** can easily be scale as we have one employee to provide training, a dozen or even a hundred, eLearning will work for all of them.
- In some situations **E-Learning** is not a complete solution as some of us learn better when we can discuss our knowledge and ideas with others in these situations one can go for blended learning –mix of online and instructor-led courses.
- **Training to employees through E-Learning** fails to offer the hands-on experience that helps people to retain new knowledge and skills, so one should look for the E-Learning techniques that offer hands-on projects to learn new skills and virtual labs for software development classes.

There are a number of **Web Based Authoring Software's** that helps us to provide online training to employees and to create own training site. One of the example of such tools T**ovuti** (https://www.tovuti.io) is web-based training software that allows you to buildyour own online training site and add thelearning content in the form of videos, power points, PDF's, SCORM files etc. This authoring tool offers more than50 different types of training activities and assignments that can be used to builde-courses.

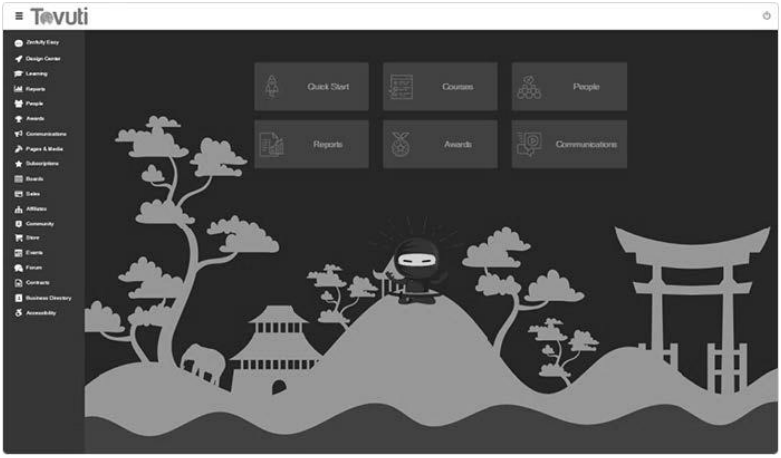


Fig. 3: Tovuti (a web based authoring tool)

3. E-LEARNING AND RURAL INDIA

India has witnessed the education divide in terms of equality and accessibility for decades now. It is expected that the population of the country will touch 1.5 B by 2030 and 1.7B by 2050; hence, the existing physical infrastructure will fall short of meeting the learning needs of this ever-growing population. Being enormously diverse nation, along with infrastructure being the major hurdle, language and content also pose problems. It is very difficult to provide quality teachers in every remote part of the country or even in the semi urban and urban areas.

Over 60% of students up to the age of ten in the rural India, lack basic reading skills. There has been a considerable improvement in the Pupil Teacher Ratio in the past decade, yet the single teacher schools in rural parts accounting 8.8 percent of the total schools in India, is the major concern of the education system. These are the main factors contributing towards degrading quality of education and high school dropout rates by the age of fourteen.

Thus, the importance of e-learning cannot be over emphasized providing best learning resources to the student's right at the comfort of their home and convenience. India provides one of the cheapest internet facilities across the world making it even more convenient for the users to access these resources rather than spending hundreds of rupees on the books."E-Kranti", an initiative undertaken by the government with collaborations various telecom companies for providing basic internet services in the most remote locations of the country under the Digital India campaign. The barriers faced by the teachers in the rural areas can be listed as limited training in the use of digital tools, no prior experience in using latest technologies or tools. This problem can be overcome by training the trainers, initiated by several government and non-government organizations (NGO's) on using various digital tools such as LCD screens, smart boards etc, to teach different concerts through these learning aids. A single teacher can address a larger audience located across remote areas along with overcoming the problem of shortage of teachers.

The union budget of 2018–19 has allocated a sum of Rs. 456 cr towards the development and deployment of the infrastructure across the rural locations in order to bring about a sea change in the style of learning. The NGO's play a very significant role in spreading e-learning across the villages and rural area in the country. 'PRATHAM" is one such program in collaboration with Vodafone, who have inducted digital classroom called "Learn, out of the box" to enrich and enhance the quality of

education. The teachers have the freedom to design the content and curriculum as per the need of the students depending upon the area and location under the following pointers:

- The content should have a sense of purpose that is achievable.
- Due consideration given to time and space constraints, yet involve a complete participation.
- Discovering individual talents.
- Learning outcome closely linked with the practical application.

Fig. 4: A School in Amravati District Using Alexa for E-learning

4. E-LEARNING AND WOMEN

Women have always been considered the weaker sex in the society through generation across the globe. They had to face restriction, cultural and societal pressures and were always considered a handicap. However, being one of the fastest growing economies in the world, socio-economic development of women is a key pillar that holds the development of the entire nation. Empowerment of women being with providing them basic education so that they understand their fundamental rights and duties. In rural India, women are not given the opportunities to complete their primary education, where there is still a gender divide when it comes to empower women. For women who are not allowed to go to school, or drop out at the primary level, or are not able to complete their education due to either family, financial pressure or distance being hindrance in their progress, e-learning can offer a life changing solution, create social and legal awareness and overall upliftment of the society, confront the patriarchal forces. E-learning offers a viable solution to the barriers related to the women education. E-learning can:

- Help raise the self esteem.
- Help complete the education at any level.
- Give the liberty to study at their own pace, time and place.
- Increase career opportunities, earn income, become financially independent.
- Help women look after their families and yet have a career.
- Help women contribute to society and have an analytical approach.

Thus, e-learning serves as a very powerful tool for women to empower themselves as well as society in general.

5. E-LEARNING AND EMPLOYABILITY

Employment opportunities are one of the major concerns of students pursuing online courses and they still have a long way to go as compared to regular campus based courses or degrees. But the future prospects for online courses are on an upward trend. Currently, India is only second to USA in the number of e-learners and is expected to grow leaps and bounds by the year 2025. The online education market stands USD 247 million that will grow to about USD 1.96 billion by the year 2021. The students mainly opt for online courses for pursuing PG courses, mostly in tier 1 cities, whereas in tier 2cities they go for online course that provide lucrative job opportunities.

Several online courses offer different programs to improve skill set of an individual for promotion and appraisal chances. The working professional can enroll in such courses in the comfort of their homes and improve upon their knowledge base. An added degree or a certificate course can do wonders to the resume of a student or professional. Many universities in India offer many such courses that everyone can take advantage.

Some online course offer preparatory programs for clearing competitive examinations specially in the tier 2and 3 cities, where exam specific mobile app helps prepare for these examination in lack of quality teachers through mock tests, videos, counseling, and webinar. MOOC's (Massive Open Online Courses) has been revolutionary in e-learning globally and offers many lectures even helpful for the educators .

E-learning has created massive impact on the way education can reach all irrespective of age, sex, caste and creed. Factors that make e-learning a preferred choice:

- E-learning comes with low cost as compared to regular courses. Even the lectures recorded once can be used any number of times anywhere.

- It offers complete flexibility in terms of choice of place, time and pace, which suits the busy schedule of an individual.
- Learn on the move. We can access contents from anywhere and do not have dependency of going and sitting in a classroom.
- The whole concept of e leaning is to make education more available and interesting.

However, there are areas where e-learning can have few disadvantages:

- Computer proficiency is a very important prerequisite of e-learning and some user may find it very tedious affair as not everyone is comfortable in front of a machine ad prefer traditional classrooms.
- Dependency on good internet connection, which can be a major hindrance sometimes.
- Learners may lose interest or motivation since there are no fixed schedules.
- Lack of proper assessment and evaluation, as students might not be honest while answering.

6. A REVIEW OF CASE STUDIES OF E-LEARNING PROJECTS CARRIED OUT IN THE FIELDS OF BANKING, FINANCE AND INSURANCE SECTORS

E-Learning has been successfully integrated into the corporate world and various companies for educating, informing and spreading awareness among both employees and customers. A learning management system forms the basis for training, educating, assessing and delivering content.

We review the case study of Enhance System Private Limited, a company based in Delhi that specializes in providing e solutions to various banking and financial firms for over 12 years. In the review we discuss how Enhance System has helped ICICI bank, one of the largest banks in India, develop customized content, e-learning software and training management systems.

6.1 Enhance Assists the Bank in Training the Employees to Better Understand Their Product and Sell Them

The Enhance System developed a module that provided the required information about the products being offered by the bank, also effectively train its employees to understand the products thoroughly and hence be able to sort out the customer queries. The features of each product were explained through the intense use of graphics and quizzes. The executives were exposed to interactive real life scenarios between the bank and

customer. A feedback mechanism was provided to have a comprehensive leaning of the module. This module is currently live and the employees are trained using the same.

6.2 Enhance Provides E-solutions with Design Mechanism for Instructional Strategies for Its Banking Clients

For managing the instructional design strategies, Enhance Systems partitioned each objective into smaller topics, which then underwent treatment on levels of interactivity from being empirical to simple knowledge, creating real life scenarios with the use of graphics. Later in the stage role plays, case studies, and stories were inducted. Modules were successfully accepted by the bank.

6.3 Enhance Develops Simulations for Managing Sales for ICICI Bank

For the purpose of absorption through campuses across India, ICICI bank felt a requirement of providing information on banking and finance through e-learning. Enhance provided a LMS to prepare students in banking field for absorption. For this Enhance integrated its own LMS with other LMs already into the system and with success. It was the first time that such concept was introduced in educational institutes. The LMS was hosted on temporary servers and was able to handle service calls from students. It resulted in ICICI bank getting favorable response from the colleges.

6.4 Enhance Uses Simulation Method for Counseling for ICICI Bank

ICICI bank is one of the biggest recruiters in India and their young recruits face the task of leading the teams very early in their careers and face the heat. For this they needed counseling. Enhance developed a simulations for the users to go through real life situations. A counselee would be there in front and data would be supplied from the background on which the users willinteract with the simulation and answer accordingly. Additional support through a mentor was provided that could be called upon at any stage for guidance. An audit is provided to user at the end of the session on his performance with a constructive feedback.

6.5 Enhance Distributes Learning Material on CD's

Enhance distributes course content on CD's which were also made available via LMS of the Enhance. This helped the employees to study

these courses at their leisure if the internet facilities were down. These could be viewed offline as well as updated versions using LMS.

From the above case study, the importance e-learning is shown across various applications in which e-learning can be helpful for acquiring additional skill.

7. CONCLUSION

E-learning has made a very strong impact on the learning and teaching in general. Looking at the current trends in India it is very safe to say that the markets are expected to grow considerably in the next five years. Many universities and institutes have adopted e-learning to increase the trainers and students access to information. It is improving the literacy rate across the country at a very fast pace. It has help to address and overcome a majority of problems related to higher education and bridge the gap between the rural and urban parts of the country. E-Learning has helped the employees in corporate sector to learn and master new skills in order to meet the challenges in changing technology at such a rapid pace. E-learning has raised the level of literacy and economic development in countries where technical education is costly, opportunities are limited and economic disparities exist.

REFERENCES

[1] Mukherjee, D. eLearning a Potential Learning Solution for Rural India—A Critical Analysis.

[2] https://www.researchgate.net/publication/235685738_Innovative_Elearning_in_Rural_Areas_A_Review_Network_Promoting_e-Learning_for_Rural_Development_e-Ruralnet.

[3] Singh, A. and Singh, L.B. E-Learning for Employability Skills: Students Perspective.

[4] https://www.enhancelearning.co.in/Images/Financeebook.pdf

[5] https://www.simplilearn.com/pros-and-cons-of-elearning-as-a-corporate-training-solution-article

[6] https://targetstudy.com/articles/advantages-and-disadvantages-of-e-learning.html

[7] LetsEndorse _ Solar-powered e-classrooms delivering digital education in rural areas.pdf

Contributors

Dr. Abhishek K. Singh, Academic Expert and Media Researcher, Ex-Faculty, Ministry of Higher Education, Sultanate of Oman.

Ms. Amanpreet Kaur, Student, BCA 2nd Year, Institute of Innovation in Technology and Management, Affiliated to G.G.S.I.P. University, New Delhi.

Ms. Anu Sharma, Student, Department of Computer Science, Kalindi College, University of Delhi, India.

Ms. Anu Yadav, PhD CSE research scholar, Department of computer science and engineering IGDTUW, Delhi, India.

Ms. Arokia Ramya Terrance, Assistant Professor, Department of Computer Science, Kalindi College, University of Delhi, India.

Prof. Ayub Khan, Department of Mathematics, Jamia Millia Islamia, New Delhi, India.

Ms. Bharti Seth, Research Scholar, Indira Gandhi Delhi Technical University for Women, Delhi, India.

Prof. Bhupender Kumar Som, Professor, JIMS Rohini, Sector–5, New Delhi–110085.

Dr. Diksha Grover, Assistant Professor, Kalindi College, University of Delhi.

Prof. Ela Kumar, Professor and HOD, Department of Computer Science and Engineering IGDTUW, Delhi, India.

Ms. Geeta Arneja, Guest Faculty, Department of Computer Science, Rajdhani College, University of Delhi, Delhi.

Ms. Harindri Chaudhary, Department of Mathematics, Jamia Millia Islamia, NewDelhi.

Dr. Harmeet Kaur, Department of Computer Science, Hansraj College, University of Delhi, Delhi–110007, India.

Ms. Kanika Tyagi, Student, Department of Computer Science, Kalindi College, University of Delhi, India.

Ms. Kirti Bhardwaj, Student, JIMS, Rohini, Delhi.

Dr. Kuldeep Singh Kaswan, School of Computing Science and Engineering, Galgotias University, Greater Noida.

Prof. L.N. Das, Department of Applied Mathematics, Delhi Technological University, Delhi–110042, India.

Prof. Madhav Govind, Professor, Centre for Studies in Science Policy, JNU.

Prof. Naresh Kumar, School of Computing Science and Engineering, Galgotias University, Greater Noida.

Ms. Nikita Shokeen, Student, Department of Computer Science, Kalindi College, University of Delhi, India.

Ms. Nitya Saini, Student, Department of Computer Science, Kalindi College, University of Delhi, India.

Dr. Punita Saxena, Associate Professor, Shaheed Rajguru College of Applied Sciences for Women, University of Delhi, India.

Ms. Rubina, Research Scholar – Centre for Studies in Science Policy, JNU.

Mr. Sanyam Gupta, Department of Applied Mathematics, Delhi Technological University, Delhi, INDIA 110042.

Dr. Shalini Arora, Associate Professor, Indira Gandhi Delhi Technical University for Women, Delhi, India.

Ms. Shalini Sharma, Assistant Professor, Department of Computer Science, Kalindi College, University of Delhi, India.

Mr. Sharique Hassan Manazir, Research Scholar, Centre for Studies in Science Policy, JNU.

Ms. Shivani Kandwal, IITM, Delhi.

Ms. Shubhangi Gupta, Student, JIMS, Rohini, Delhi.

Dr. Shweta Paradkar, Assistant Professor, Jesus and Mary College, University of Delhi.

Dr. Suman Madan, Assistant Professor, JIMS, Rohini, Delhi.

Mr. Sunny Seth, Assistant Professor, JIMS Rohini, Sector-5, New Delhi–110085.

Dr. Sushma Malik, Assistant Professor (IT), Institute of Innovation in Technology and Management, Affiliated to G.G.S.I.P. University, New Delhi.

Dr. Tripti Lamba, Associate Professor, IITM, Delhi.

Ms. Uzma Nigar, Department of mathematics, Jamia millia Islamia, New Delhi, India.

Ms. Yukta Pandey, Student, BCA 2nd Year, Institute of Innovation in Technology and Management, Affiliated to G.G.S.I.P. University, New Delhi.

Author Index

Arneja, Geeta 38
Arora, Shalini 141

Bhardwaj, Kirti 64

Chaudhary, Harindri 174

Das, L.N. 78

Govind, Madhav 48, 90
Grover, Diksha 185
Gupta, Sanyam 78
Gupta, Shubhangi 64

Kandwal, Shivani 127
Kaswan, Kuldeep Singh 158
Kaur, Amanpreet 98
Kaur, Harmeet 38
Khan, Ayub 107
Khan, Ayub 174
Kumar, Ela 1
Kumar, Naresh 158

Lamba, Tripti 127

Madan, Suman 64
Malik, Sushma 98
Manazir, Sharique Hassan 48

Nigar, Uzma 107

Pandey, Yukta 98
Paradkar, Shweta 185

Rubina 90

Saini, Nitya 15
Saxena, Punita 141
Seth, Bharti 141
Seth, Sunny 119
Sharma, Anu 15
Sharma, Shalini 15, 158
Shokeen, Nikita 15
Singh, Abhishek K. 28
Som, Bhupender Kumar 119

Terrance, Arokia Ramya 15
Tyagi, Kanika 15

Yadav, Anu 1

View publication stats

The Journal of Supercomputing
https://doi.org/10.1007/s11227-019-03026-x

An ANFIS-based compatibility scorecard for IoT integration in websites

Akshi Kumar[1] · **Anshika Arora**[1]

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract

Cyber-physical systems and Internet of Things (IoT) form two different levels of the vertical digital integration. Integration of websites with IoT-connected devices has compelled creation of new web design and development strategies where websites are designed keeping in mind the permutations of smart devices. The design should be seamless across different devices and the website design company or web designer should be well informed and aware of the different considerations for design with IoT interactions. In this work, we expound the effectiveness of IoT integration in website design. To realize an IoT-powered IT ecosystem as an essential technology for improving customer experience, a strength–weakness–opportunity–threat analysis is done. Further, with an intent to apprehend the integration support that an existing GUI front end may provide to a smart device, an ANFIS model is proposed to determine the compatibility of an e-commerce website for integration with IoT devices. A dataset of 600 e-commerce websites from .com domain is used to train and test the learning model. Seven features (page loading speed, broken links, browser compatibility, resolution, total size, privacy and security, and interface and typography) which impact the compatibility of IoT integration in websites have been used. Evaluation criteria for assigning score to each feature has been identified. Finally, the compatibility score, the $\text{IoTScore}_{\text{site}}$ which evaluates the websites' integration capabilities and support to IoT devices is generated by adding all the feature scores. The preliminary results generated using the prediction model clearly determine the worthiness of website for IoT integration.

Keywords IoT integration · Website design · Cyber-physical systems · ANFIS · Machine learning

✉ Akshi Kumar
akshikumar@dce.ac.in

[1] Department of Computer Science and Engineering, Delhi Technological University, Delhi, India

Published online: 18 October 2019

1 Introduction

An IoT is a network of objects embedded with electronics, software, network connectivity and sensors [1]. It is about connecting numerous heterogeneous cyber-physical devices or things that sense, monitor, analyse and adapt stochastic events gathered without human origination to create value from autonomic agents that process data, to create granular knowledge and wisdom for both machines or humans to make smarter decisions and actions [2]. It merges the virtual and physical worlds, creating a smart environment such that when these devices share and analyse data, they have the capability to transform our lives in innumerable ways. Figure 1 explains how IoT and cyber-physical systems (CPS) are connected.

A generic IoT system involves three key components:

1. *Embedded systems* Consists of heterogeneous sensors responsible for providing information about the physical entity.
2. *Middleware* Is responsible for data aggregation, acquisition control and network connectivity.
3. *Cloud services* Provides comprehensive storage, computation, analytics, application hosting and management mechanisms.

IoT is a biggest frontier which can improve our lives in many aspects. This cutting-edge technology not only affects the consumers but also the providers of IoT solutions. With the ubiquitous internet connectivity and bandwidth in abundance along with reduced cost of wearables and smart devices, the impact of the Internet of Things (IoT) and technology in our lives is manifold and profound. IoT devices are

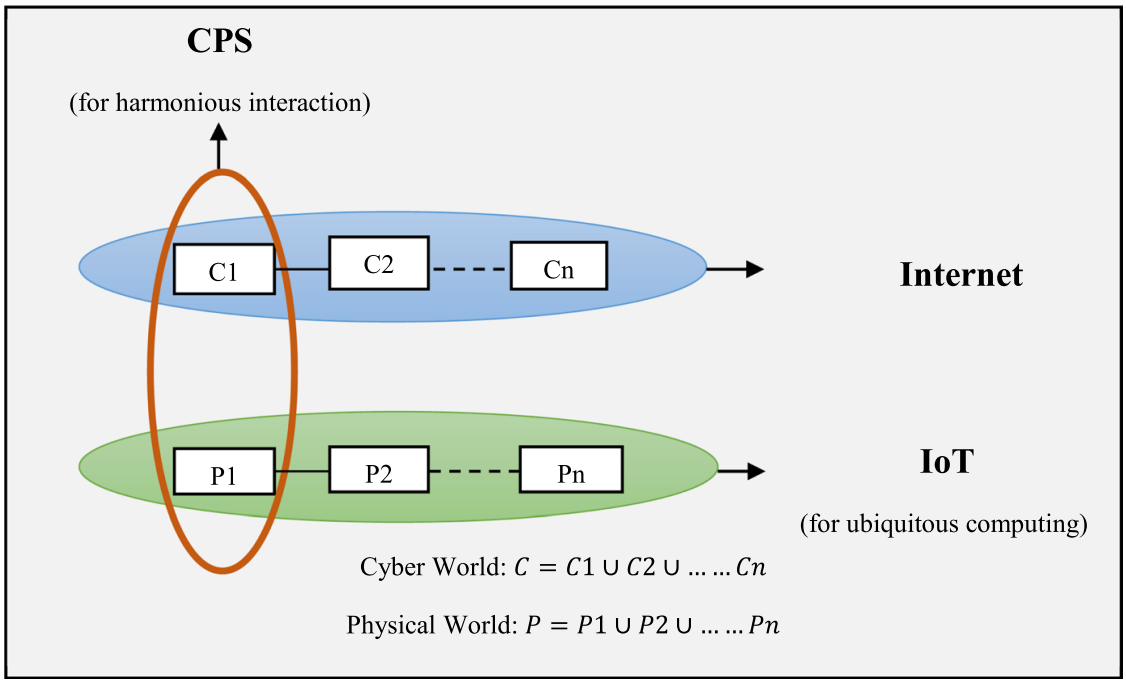


Fig. 1 Interconnection between IoT and CPS

becoming a part of the mainstream electronics culture [3]. According to Statista,[1] the IoT market is expected to reach $457 billion by 2020. This is a phenomenal increase from just $194 billion in 2017.

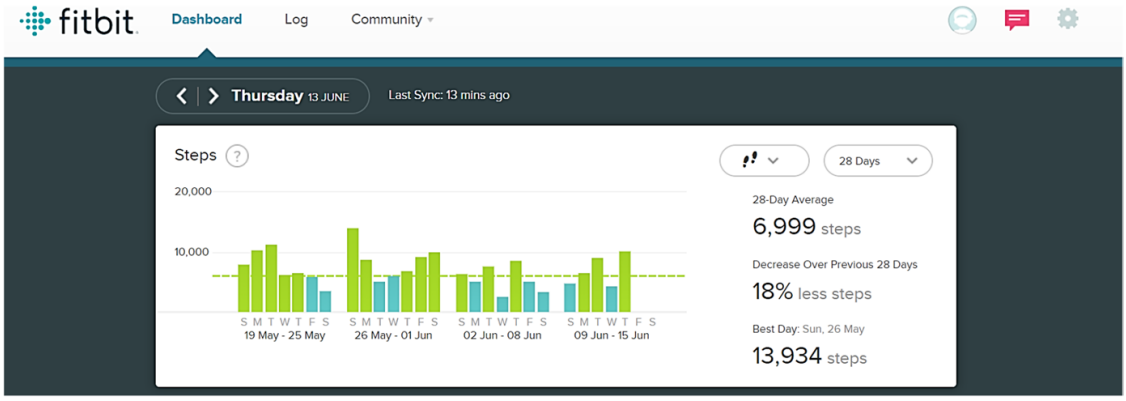
As the IoT solutions define the new paradigm for customer-oriented digital experience, they are inevitably getting more complex and dynamic. The network-connected cyber-physical devices can communicate via standard protocols and share data in different environments via the intermediate gateway nodes like modems, routers, switches and cellular base stations. To build a single communication platform which acts seamlessly for devices and applications to talk to each other, regardless of how they are physically connected is a challenging task [4]. Developing a completely new platform of communication where all devices are digitally augmented is expensive and time-consuming. A better solution is to reuse the established World Wide Web ecosystem and leverage the capabilities of its readily available application layer protocol, thus typifying the concept of Web of Things (WoT). WoT complements the existing IoT ecosystem, reduces costs and creates value by combining multiple devices and information services [5, 6]. More specifically, while IoT is all about creating a network of things, objects, people, applications and systems, WoT makes their integration to web. Integration aids capturing data from smart devices and making the data move into business applications with the intent of automating processes, supporting real-time monitoring and applying analytics for understandings. Recently there have been many web-enabled IoT applications such as smart homes, smart cities, smart industry, smart agriculture and smart healthcare which play a significant role in our lives. Though organizations and businesses are eager to develop applications that support these devices, exploiting IoT device interfaces to integrate applications with those devices can be complex. Further, deriving meaningful information is another crucial aspect.

Web-enabled IoT application development is a lot more intricate than a customary web development. This is primarily owing to the voluminous data, dynamic user interfaces, intense communications and security and reliability problems pertaining to the IoT applications [7]. These peculiarities influence the current generation of website design and development. Web design can be broken into two key categories [8]:

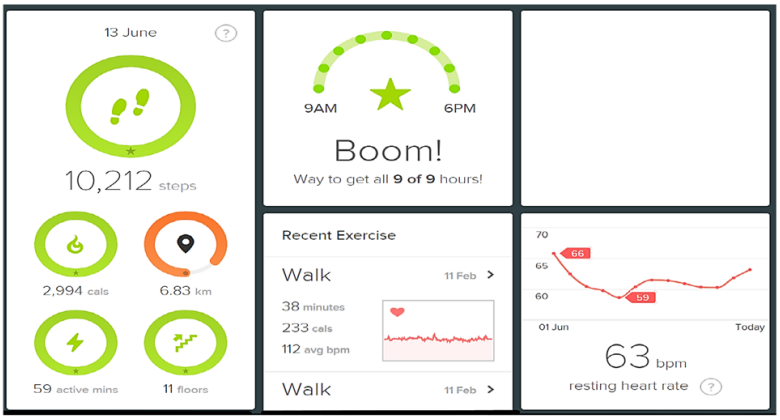
1. Front end: Graphical User Interface (GUI) which can be related to User Experience (UX) [9].
2. Back end: Database storage, data manipulation and data processing [10].

The front-end interface is used by the users to connect and communicate with IoT devices like cameras, sensors, etc. For example, the wearable technology, fitbit uses a bluetooth connectivity to track user's physical activity including the number of steps taken, distance traversed and calories burned, etc. [11] and integrates with software that logs this personal activity data (Fig. 2). A clean interface is needed for an easy communication with smart devices. The design should be tractable and

[1] www.Statista.com.



(a)



(b)

Fig. 2 fitbit dashboard on "fitbit.com"; **a** user's physical activity tracked, **b** personal activity log

must adapt itself according to various IoT-connected devices. Concurrently, numerous websites may also require the multifaceted capabilities to communicate with the back-end databases which store personalized data obtained from IoT devices [12, 13].

The design should be seamless across different devices and the website design company or web designer should be aware and well-informed of the different considerations for design with IoT interactions. Thus, many challenges ascend for web designers in terms of both back-end and front-end design to make a website worth for IoT integration. Some design considerations required to successfully develop website designs with IoT interactions include [14]:

- *Strong back-end* It is important that communication between users and IoT devices is very clear. This can be attained establishing a robust back-end design architecture. Each of the devices has several capabilities, data and commands that operate just according to the inclinations of users. A fine back-end receives and transmits data, and it helps the users in communicating with the device in a seamless way.
- *Intuitive user interface* For an IoT application, it is vital to have a clean, fast and an interactive layout. A web designer must work on latest web design practices but ensure that design inclusions are meaningful and support a rational

workflow. Also, the major concern is that the new IoT-based design should be mobile compatible so that smart users can benefit from it. User interface further requires to work on two aspects as given next:

1. *Easier front-end interaction* JavaScript is the most efficient language which has the potential to make a smart device interactive. It is an event-driven language and has capability of executing commands simultaneously with the interaction with smart devices and capturing data in the database. Moreover, it has the capability of multi-tasking where multiple tasks can run at the same time. JavaScript must be targeted by all web developers as it has an extensive set of plugins and libraries.
2. *Mobile user-experience* Mobile interfaces are the need of today. Simple designs are better for understanding and are easier to load on mobile. Web developers are expected to produce designs that are capable of adapting itself across many smartphone screens irrespective of the feature and size for the smart device.

- *Speed measurability* Conventional page requests are directly analysed by the web server. However, in the case of contemporary IoT-integrated websites, there are two hardware components, namely the IoT device and web server using an internet cloud service, thus increasing the chances of slow user response and more latency. Thus, design strategies for slow connections must be adopted. Simple and lighter themes make the pages load faster. Following are some of the steps that can improve page load speed:

  1. Create accelerated mobile pages (AMP) pages that are capable of loading faster on mobile devices.
  2. Use browser caching.
  3. Minimize the number of CSS and JS files.
  4. Reduce the number of redirects.
  5. Minimize the number of images.
  6. Hotlinking of images should be disabled.
  7. Reduce the server load with the use of CDN.
  8. Fix the broken links.
  9. Use caching plugins and clean the database.

- *Power management* Most of the IoT devices are battery-powered and wireless. For this, fully fledged power management is essential. Heavy programmes running in the back end, do drain the battery and this leads to less user communication. This is why the new layout must be designed for minimizing excess power use.
- *Security and privacy* One common drawback of IoT or the associated solutions is that they offer a way to hackers. If one could unlock a web-connected front door, a hacker potentially could do the same and help himself to the contents of the home. Though device designers are responsible for privacy and

security, but there is a need of a third-party web designer who must make security the main design consideration.
- *Testing considerations* Testing of IoT integrated website is complicated than the conventional websites.

IoT has made a significant impact on e-commerce businesses. IoT-enabled devices exchange data with each other through the internet, helping retail and e-commerce businesses to carry on their operations efficiently. As more devices get connected and gain smart features, more data are gathered, and consumer experience is improved. With the penetration of IoT, the e-commerce industry is expected to see a growth in revenue, a better supply chain management and inventory management and an improved customer service. This integration of websites with IoT devices has fostered the need to upgrade web design and development strategies where websites are designed keeping in mind the permutations of smart devices. Flexible, clean and user-friendly interface is the need of the hour where the information has to be displayed using lesser content.

Motivated by this, in this work, we expound the effectiveness of IoT integration in website design and propose an adaptive neuro-fuzzy inference system (ANFIS) to determine the compatibility of an e-commerce website with IoT devices. The ANFIS combines the learning capabilities of neural networks with the abilities of fuzzy logic to model uncertainty in expressiveness. It incorporates the benefits of adaptive control technique, artificial neural network and the fuzzy inference system [15]. The intent is to apprehend the integration support that an existing GUI front-end may provide to a smart device using a prediction model. A dataset of 600 e-commerce websites from .com domain is used to train and test the learning model. Seven features (page loading speed, broken links, browser compatibility, resolution, total size, privacy and security, and interface and typography) which impact the compatibility of IoT integration in website are used. These features are derived from the design considerations for successful IoT interactions with website [14]. Evaluation criteria for assigning score to each feature is identified. The sum of all the feature scores generates the compatibility score, the $\text{IoTScore}_{\text{site}}$. This compatibility score is finally used to evaluate the websites integration capabilities and support to IoT devices. Thus, the primary contributions of this research are:

1. A strength–weakness–opportunity–threat (SWOT) analysis to demonstrate the effectiveness of IoT integration in website design.
2. A novel ANFIS model to determine the compatibility of an e-commerce website for integration with IoT devices.

The proposed prediction model will facilitate to realize an IoT-powered IT ecosystem as an essential technology for improving customer experience. The paper is organized as follows: Review of related literature work is given in Sect. 2 followed by Sect. 3 which expounds the SWOT analysis of IoT integration in website design. Section 4 describes the concepts used in proposed model and Sect. 5

presents the proposed ANFIS model. Section 6 discusses the results and finally, the conclusion and future scope of the research is presented in Sect. 7.

2 Related work

Various studies have been reported in the domain of IoT. Gubbi et al. [2] presented an implementation of IoT based on cloud vision and also discussed the application domains and technologies that were likely to drive IoT research in future. They also presented an implementation using Aneka, which considered the interaction of private and public clouds. Lee et al. [16] presented five major IoT technologies which were important for the deployment of IoT-based products and services. The authors discussed three categories of IoT for enterprise applications used to enhance customer value and also described five technical and managerial challenges. Gigli et al. [17] introduced four categories of services which can be used to build an IoT application. Many applications of IoT have been reported in the field of healthcare and medical science [13, 18], agriculture [19], across fog and cloud [20], etc.

Software architectures originated from Web service research mainly for Web service interoperation to Web resources management for IoT have been proposed. Castellani et al. [21] worked on web services for IoT and concentrated on the communication technology for IoT devices and its implementation adopting the Representational State Transfer (REST) approach. The REST relied on GET and POST methods of HTTP and data were exchanged using standard XML. Xu et al. [13] proposed an IoT-based model for emergency medical services. They described the data collection, data integration and data interoperation methods in case of medical emergencies. They proposed a global data accessing method (UDA-IoT) to handle heterogeneous IoT data in medical service using RESTful architecture. Doukas et al. [18] discussed how IoT features web applications which can be accessed globally, were scalable, and were able to provide communication interfaces to peripheral applications. They discussed how appropriate interfaces enable the data dissemination to external applications. They also discussed how web apps provided the indispensable data real-time monitoring and management. The authors also explained that device and location independence enabled users to operate the systems via web browser and illustrated various websites which managed IoT sensors data. Yu et al. [22] argued that interactivity of a website is essential in IoT computer-mediated communications as it is considered as one of the key advantages of the medium. They also stated that interactivity is a vital part of the IoT service and influences customers' online experience in a positive way.

Researchers have proposed models which facilitate evaluation and assessment of websites based on various design features of the website. Kumar and Gupta [9] discussed the shift from software quality models to web-based quality models eventually proposing a π model where the horizontal line of the π signified the backbone of quality models with quality assessment parameters common to both kind of software, whereas the two vertical pillars of the π depicted the quality attributes specific to the software type. Loiacono et al. [23] proposed WebQual model which evaluated website quality based on a number of dimensions,

and the model was evaluated on three samples. Hartmann et al. [24] focused on user experience exploring some of the contextual features that impact user experience. Contextual evidence facilitates digging out User-centric information [25]. Sobecki et al. [26] demonstrated website usability assessment models and discuss about website interface quality. They also present models for organizing data in frames further to be used in an expert system. Schubert and Selz [27, 28] developed Web Assessment Tool considering the three phases of a transaction on a website, namely information, agreement and settlement and named this tool as WA. Extended Web Assessment Model (EWAM) was an extended model of WA which contained fundamentals of TAM and social influence proposed by Schubert et al. [29]. In our previous work [30], we identified website quality features which were necessary to predict the quality of website and introduced a website quality model using filter-wrapper-based feature selection algorithms.

None of the above-mentioned website quality models consider integration with IoT to evaluate website quality. This research is the first reported study to effectively capture and measure the IoT integration in websites qualitatively. We propose a prediction model to calculate an IoT integration compatibility scorecard for e-commerce websites. The study substantiates the need of IoT integration to collect data from smart devices to be utilized by the business applications to automate processes, support real-time monitoring and apply analytics for insights. Thus, the proposed model is superlative as compared to other website assessment models [23, 24, 26–29]. We propose an Adaptive Neuro-Fuzzy Inference System (ANFIS)-based model to validate the compatibility of IoT within a website. ANFIS has previously been applied in other areas of predictive learning which include image processing, face detection systems, etc. [31–33].

3 SWOT analysis

A SWOT analysis framework is generally used by business organizations as a strategic planning technique for analysing relevant internal and external factors making an impact over people, process, product, projects, initiatives, objectives and outcome of a system. It is a fundamental model used to evaluate the organizations internal strengths and weaknesses, as well as external opportunities and threats. It also acts as a vigorous mechanism for decision making [34].

Integrating websites with IoT-connected devices brings the opportunity to adopt new front-end design strategies of websites. Sophisticated, flexible and a clean user interface is imperative. To examine how IoT integration with websites has affected the website design, a SWOT analysis is done. The strengths, weaknesses, opportunities and threats of IoT integration in website design are as follows:

- *Strengths* Strengths determine the internal, positive attributes of integrating websites with IoT devices.

- Innovative front-end web designs allow many IoT devices to display web data searched via browsers on the device's screen.
- Interactive user interface help users to interact with the websites on various IoT-connected devices.
- New front-end web designs are mobile compatible that are used across various smartphone screens.
- New designs are being used which adapt themselves according to the varied screen sizes of IoT-connected devices having different display types.
- New designs are capable of integrating with multiple web browsers which may be used to search web data on most of the IoT devices.

- *Weaknesses* Weaknesses are negative factors that detract from the strengths discussed.
  - A single design may not be easy to adopt across different devices that are connected through IoT.
  - New website designs require the complex capability to communicate with back-end databases that are able to store personalized data obtained from IoT devices.
  - Due to communication and data transfer taking place through cloud slow user responses are experienced.
  - Heavy programmes running in the back end, use enough battery which leads to less user communication.
  - Testing IoT-integrated websites are complicated than conventional websites.
- *Opportunities* Opportunities are external factors that are likely to contribute to the success of the proposed model.
  - Avoid loss of productivity and loss of product.
  - Improved customer experience and engagement.
  - Increased competitiveness.
  - New revenue streams.
- *Threats* Threats are external factors that cannot be controlled.
  - Designs for extremely small screen sizes are potentially difficult to be made.
  - As small screen sizes require web content without images which may not give desired e-satisfaction to users of websites.
  - Multi-device interaction mostly implies recurring authentications, gateway processes that differs with devices.
  - Security and privacy may be breached when websites are integrated with IoT-connected devices.

Quite clearly, creating a standardized setting for normal and seamless interactions between a user and a system is challenging. Adopting an IoT solution requires cross-functional coordination between business, operations and security verticals. Thus, based on the identified need to create efficient user experience for IoT products, in this research, we assess the effectiveness of IoT integration in website design and propose an ANFIS to determine the compatibility of an e-commerce website with

IoT devices. The ANFIS is trained using key features derived from the core design components supporting IoT integration with websites, namely interactive interface, speed measurability, privacy and security and power management. The following sections discuss the details.

4 Background concepts

Sections 4.1–4.3 discuss the concepts and techniques used in this work.

4.1 Artificial neural networks (ANN)

ANNs imitates the human nervous system where number of nodes in ANN represents the neurons in brain [35]. The nodes are connected to each other by links which have weights associated with them. The nodes take data as input to perform basic operations on it and then pass the output to other nodes through connecting links. This output is known as activation value. The learning in ANNs takes place by updating weights of links whenever needed. Whether to alter the weight or not depends on the output predictions. In order to get accurate outputs, activation functions are applied on inputs. The most commonly used activation function in ANN is sigmoid function [36]. The way of modification of connection weights can be categorized as follows [37]:

- *Supervised learning* The ANN is given a dataset. The output of ANN is compared with the desired output present in the dataset. Based on the difference between predicted and actual values, the weights are updated. This process repeats until ANN gives the desired output.
- *Unsupervised learning* This way of training is independent. During training, ANN clusters the same kind of input data into clusters. For a new input data, ANN categorizes it into one of the clusters. No feedback on results is provided in this case.
- *Reinforcement learning* It is similar to supervised learning as in this case ANN gets feedback for its outputs. The difference is that here the feedback is evaluative instead of instructive. The weights are adjusted on the basis of these feedbacks.

There are two types of ANN topologies: feedback (feedback loops present) and feed-forward (unidirectional information flow no feedback loops).

4.2 Fuzzy inference system (FIS)

'A fuzzy Inference System abbreviated as FIS is a system that uses fuzzy set theory to map inputs to outputs'. If intended to do fuzzy classification, the inputs are features and the outputs are the classes [38]. The process of fuzzy inference involves *membership functions, fuzzy logic operators* and *IF–THEN rules* [39, 40].

4.2.1 Membership functions

Membership functions characterize fuzziness or the degree of truth in fuzzy logic. It represents a curve that describes the mapping between the points in the input space and membership values which are also called degree of membership. The value of degree of membership lies between the range of 0 and 1. The Fuzzy Logic Toolbox comprises of 11 types of built-in membership functions. These functions are built from some basic functions as:

1. piecewise linear functions,
2. quadratic and cubic polynomial curve,
3. sigmoid curve,
4. Gaussian distribution functions.

4.2.2 Fuzzy logic operators

Fuzzy logical reasoning is related to standard Boolean logic. Standard logical operations will hold when the fuzzy values are kept at boundaries like 0 (completely false) and 1 (completely true). In fuzzy logic, AND operator is implemented with minimum, OR operator is implemented using maximum and NOT as negation, i.e.

$$A \text{ AND } B = \text{Min}(A, B) \tag{1}$$

$$A \text{ OR } B = \text{Max}(A, B) \tag{2}$$

$$\text{NOT } A = 1 - A \tag{3}$$

4.2.3 IF–THEN rules

Usually the knowledge involved in fuzzy reasoning is expressed as rules. A fuzzy IF–THEN rule is of the form as given in Eq. 4.

$$\text{IF Input}_1 \text{ is } A_1 \text{ AND Input}_2 \text{ is } A_2 \ldots \text{AND Input}_n \text{ is } A_n \text{ THEN } Z = B \tag{4}$$

where Input_i and Z are fuzzy variables and A_i and B are fuzzy terms (fuzzy sets). The 'IF' part is known as antecedent or premise, whereas the 'THEN' part is termed as a consequence or conclusion. The inputs are connected via fuzzy logical connectives such as 'AND' and 'OR'.

4.3 Adaptive neuro-fuzzy inference system (ANFIS)

ANFIS combines the best features of neural network and fuzzy inference system [41]. It is defined as a multilayer feed-forward architecture where each neuron executes its function for arriving signals. It maps input features to input membership functions. Mapping is done between the input membership functions and the if–then rules. These rules are further mapped to a set of output features. Finally, the output features are mapped to output membership functions which are further

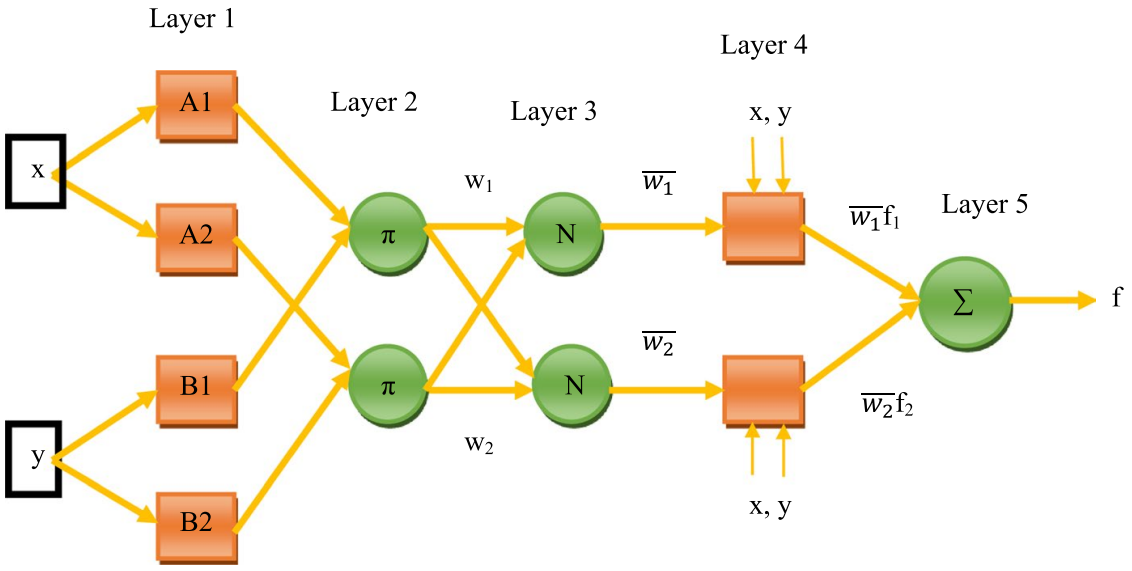


Fig. 3 ANFIS architecture

mapped to the decision which is associated with the output. ANFIS is an effective soft computing technique with easy implementation and easy incorporation of both numeric and linguistic knowledge for solving a problem. Figure 3 shows the architecture of an ANFIS.

Each layer in ANFIS contributes in calculating the parameters with the help of the respective functions present in them. The working of each layer is briefly described as follows:

1. *Layer 1* Every node in layer 1 is an adaptive node. In this layer, the output of each node is fuzzy membership grade of the inputs. Here, the parameters are known as premise parameters. During the process of learning, different parameters are used for the specification and tuning of membership functions. The membership function can be represented in Eq. 5.

$$\mu_{Ai}(x) = \frac{1}{1 + \left|\frac{x - c_i}{a_i}\right|^{2b}} \tag{5}$$

   where $\mu_{Ai}(x)$ lies between 0 and 1; a_i, b, c_i are parameter set.
2. *Layer 2* Each node is a fixed node in this layer. Product of input signals is the output of the nodes in this layer which can be written as given in Eq. 6.

$$O_{2,i} = w = \mu_{Ai}(x) \cdot \mu_{Bi}(x) \tag{6}$$

3. *Layer 3* All nodes in this layer are fixed nodes. They calculate the ratio of the ith rule's firing strength relative to the sum of all rule's firing strengths. The normalized firing strength is the output which can be written as given in Eq. 7.

$$O_{2,i} = \bar{w} \times w_1 / (w_1 + w_2) \tag{7}$$

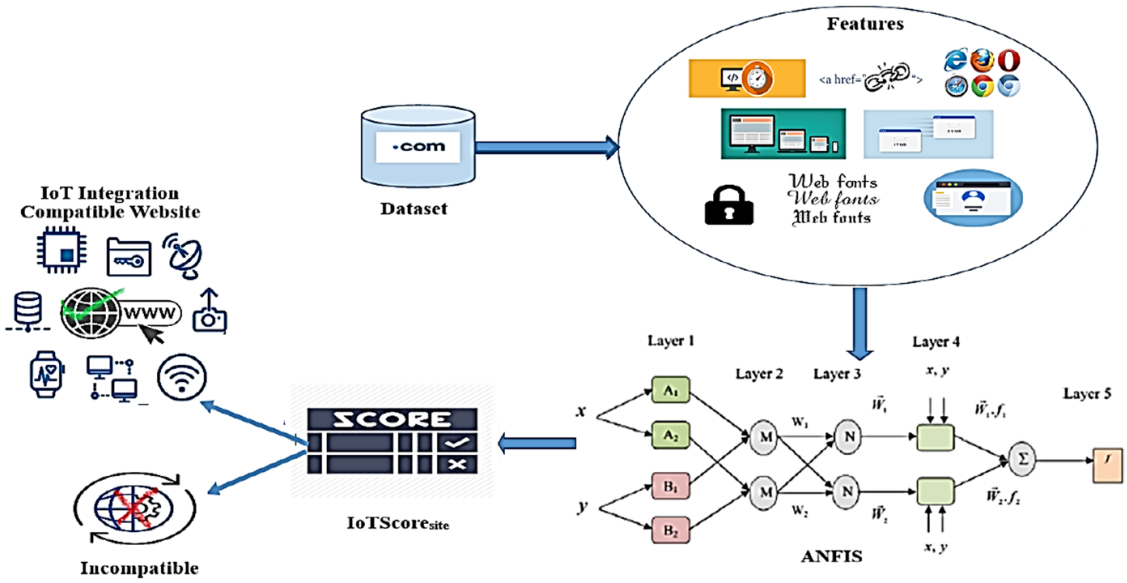


Fig. 4 Graphical illustration of the proposed model

4. *Layer 4* Each node in layer 4 is an adaptive node. The output of these nodes can be defined as given in Eq. 8.

$$O_{4,i} = \overline{w}_i f_i = \overline{w}_i (p_i x + q_i y + r_i) \tag{8}$$

where $\{p_i, q_i, r_i\}$ is the parameter referred to as consequent parameters.

5. *Layer 5* In this layer, all nodes are fixed nodes. Their job is to sum all arriving signals and output that as total output. The total output can be defined as given in Eq. 9.

$$O_{5,i} = \sum \overline{w}_i f_i \tag{9}$$

ANFIS identifies the rules on its own and tunes the membership function parameters accordingly. In order to develop an efficient model, some choices should be made carefully, for example, setting the optimization methods, types of output MFs and the number of epochs.

5 ANFIS-based compatibility model for IoT-integrated websites

Interacting with things via websites demands distinctive design considerations to eventually enhance the user experience. It is thus imperative to investigate the worthiness of a website for IoT integration. In this work, we propose an ANFIS model for evaluating the compatibility of website design and IoT device integration. Various e-commerce websites have been collected (with the domain.com) and seven website design features have been identified which typify the compatibility of website with IoT devices. An ANFIS model has been trained with these seven features and tested for the testing data. The ANFIS output as $\text{IoTScore}_{\text{site}}$ has been observed for the sample websites. Figure 4 depicts the architecture of the proposed model.

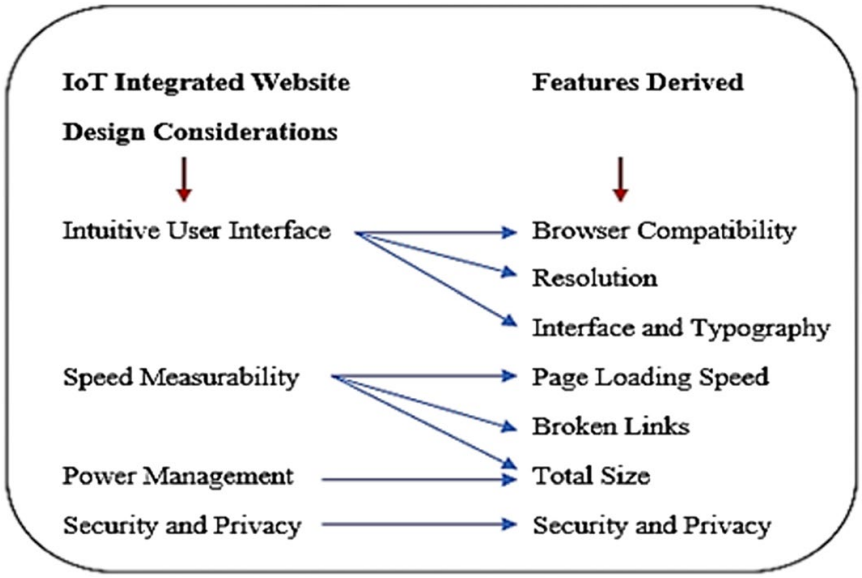


Fig. 5 Derived features for IoT integrated website design

5.1 Dataset preparation

A top-level domain recognizes a certain element regarding the associated website type, such as its objective (business, government and education), its owner or the geographical area from which it belongs. Examples of some of the popular top-level domains include .com, .org, .net, .gov, .biz and .edu. Since this work is to propose a model for the evaluation of business websites compatibility with IoT integrated devices we consider only the.com websites. In this work, 600 e-commerce websites (with the *.com* domain) have been evaluated for compatibility with IoT devices using an adaptive neuro-fuzzy inference system. Also, we have identified seven features of website design which impact the compatibility of website with IoT devices and are imperative to predict the IoTScore of a website. These features are as follows:

- Page loading speed.
- Broken links.
- Browser compatibility.
- Resolution.
- Total size.
- Privacy and security.
- Interface and typography.

The features have been derived from the design considerations for successful IoT interactions with website [14]. Figure 5 illustrates the mapping between the aforementioned website design considerations for IoT-integrated websites and website design features used in this study.

The following sub-sections describe these features and their evaluation criteria.

Table 1 Evaluation criteria based on percentage of broken links

Percentage of broken links	Score
0	10
0–0.010	9
0.010–0.020	8
0.020–0.030	7
0.030–0.040	6
0.040–0.050	5
0.050–0.080	4
0.080–0.100	3
0.100–0.200	2
0.200–0.500	1
> 0.500	0

5.1.1 Page loading speed

Page loading speed is defined as the 'speed by which a web page loads' [30]. It is inversely proportional to page loading time which is, the time taken by a web page to load completely.

Evaluation criteria An online tool 'GTMETRIX[2]' has been used to determine the loading speed. It gives the percentage value for a page's speed. In this study we have converted the percentage to score out of 10. For example: 81–90% has been scored as 9.

5.1.2 Broken links

Broken Links are 'links that don't work and send visitors to a webpage that no longer exists' [42]. It occurs due to variety of reasons such as:

- Non-availability of webpage.
- Movement of webpage without a redirect being added.
- Change in the URL structure of website.

Evaluation criteria Another online tool 'DEADLINKCHECKER'[3] has been used in this study which crawls through the website to identify presence of broken links. The evaluation criteria are based on the percentage of broken links found, and the score have been assigned as given in Table 1.

[2] www.gtmetrix.com.

[3] www.deadlinkchecker.com/.

Table 2 Scoring criteria based on total size of a website

Web page size (in MB)	Score
> 50 MB	0
25–50 MB	1
15–25 MB	2
13–15 MB	3
11–13 MB	4
9–11 MB	5
7–9 MB	6
5–7 MB	7
3–5 MB	8
1–3 MB	9
< 1 MB (in KB)	10

5.1.3 Browser compatibility

Browser compatibility is defined as the 'capability or flexibility of a website to function on different web browsers [30].

Evaluation criteria The websites have been browsed via various browsers, namely Google Chrome, Internet Explorer, Mozilla Firefox and Microsoft Edge. Websites show varied behaviours with these browsers. Some websites are not compatible with all the browsers. Scores in the range of 1–10 have been given to websites assigning 2.5 for compatibility with each browser.

5.1.4 Resolution

Resolution is defined as 'the number of pixels present on a display screen where a pixel represents an individual point of colour. The resolution is generally expressed in terms of the number of pixels on the horizontal axis and the number on the vertical axis'. The resolution feature is used to evaluate the appearance of website with different screen sizes.

Evaluation criteria The website either showed same or varied behaviour with different screen sizes. Scores in the range of 1 to 10 have been given in accordance to the ability of websites to adapt with different screen sizes, where a value of '1' corresponds to the websites which are unable to adjust effectively to different screen sizes and a value of 10 indicates the website is able to adapt itself to various screen sizes. Websites which are able to adjust with different screen sizes have been given higher scores and websites which show distorted content when evaluated on small screen sizes have been given lower scores.

5.1.5 Total size

Web page size is an integral feature which determines the speed of page loading as the bigger a page is, the longer it takes to load the required resources to display it. Total Size is the size of web page in MB.

Evaluation criteria For evaluating the size of webpage, an online tool *'GTMETRIX'* has been used in this study. Table 2 gives the scoring criteria of websites based on its total size.

5.1.6 Privacy and security

Privacy and security determine the steps taken while website design to secure the information and prevent websites from hackers and malwares.

Evaluation criteria For assessment of website privacy and security, an online tool '*Qualys SSL Labs*[4]' has been used. It grades the website as A+, A, B+, B, C and D based on Cipher Suits and Strength, SSL/TLS version, Key Exchange, Handshake simulation, Protocol Support and Details, BEAST and much more. A score of 10 has been given to the websites evaluated with A+ grade, 8 to A grade websites, 6 to websites with B+ grade, 4 and 2 to C and D grade websites, respectively.

5.1.7 Interface and typography

Intuitive Interface is defined as the extent of ease by which a website can be used and understood. Typography & Font is based on font size, font style and readability of website's content [30].

Evaluation criteria In this work interface of a website is evaluated using crowded sourcing and websites are scored in the range of 1–10 depending on its ease of usage. Also, typography and font have been evaluated and weighted average of interface score and font score has been used to determine overall score of this feature. Highest score has been given to 12-point Arial font because of its legibility and sharpness. It is also considered as the most preferred font by websites users [43]. Lower to it is scored 12-point Times New Roman as it is considered as the fastest to read font by users [43]. The other fonts lagging behind in the queue are 12-point anti-aliased Arial, 10-point Arial, 12-point anti-aliased Times New Roman and 10-point times New Roman which have been scored with decreasing values. Other fonts sizes and styles have been scored as per their readability.

Intuitive interface and typography score has been calculated using Eq. 10:

$$\frac{\alpha.\mathrm{IS} + \beta.\mathrm{TFS}}{100} \tag{10}$$

where IS is the Interface Score attained by crowd sourcing, TFS is the Typography and Font Score, α and β are the coefficients where $\alpha = 100$ and $\beta = 10$..

[4] www.ssllabs.com.

Table 3 $\text{IoTScore}_{\text{Site}}$ evaluation

Sum of features score	$\text{IoTScore}_{\text{site}}$	Compatibility
≥ 50	2	Website design is *highly compatible* with IoT devices
≥ 30	1	Website design has *medium compatibility* with IoT devices
< 30	0	Website design is *not compatible* with IoT devices

The websites in the dataset are labelled with the $\text{IoTScore}_{\text{site}}$ based on the sum of all the feature scores. This cumulative score, $\text{IoTScore}_{\text{Site}}$ finally characterizes the compatibility score of the website to IoT devices. The generated scorecard depicts the integration support that these websites have. The compatibility thresholds for integration support are as given in Table 3.

5.2 ANFIS implementation

The ANFIS implementation contains two processes: that is, the ANFIS system training and the model testing. The following sub-sections describe the details.

5.2.1 ANFIS system training

A typical ANFIS training process has been depicted in Fig. 6 [44]. The first step is to obtain training and checking data sets. Training data set must have an output vector which contains the output for every set of inputs. The premise parameters for MFs are set by using training data set. A threshold error value is determined. Least squares method is used to find consequent parameters. The premise parameters are updated using the gradient decent method if error is greater than threshold. Once the error comes out to be lesser than the threshold value, the process terminates. The role of checking data set is to compare the predicted values to actual values [44]. Hybrid learning is used by ANFIS for learning while training. Hybrid learning is an integration of the gradient descent and the least squares method. The first step of ANFIS training in MATLAB is to create training data set. The data set must be in the form of a matrix where the output should be in the last column. The number of columns depends on the numner of inputs and the rows denote the existing data points. There are 2 commands in the MATLAB to generate FIS using the training data set, namely GENFIS1 and GENFIS2. However, these two commands differ in a way that GENFIS1 produces grid partitioning of the input space to generate the FIS and thus it is more likely to have the problem of the ``curse of dimensionality'', while GENFIS2 uses subtractive clustering (SUBCLUST) to produce scattering partition. Due to this reason in this study, we have used the GENFIS2 command for generating the FIS. After setting the number of epochs matlab system training begins.

Once the training process terminates, we get the final membership functions and training. ANFIS functions well with one training data set, but if checking data set is also provided, then the chances are more that system will understand the model

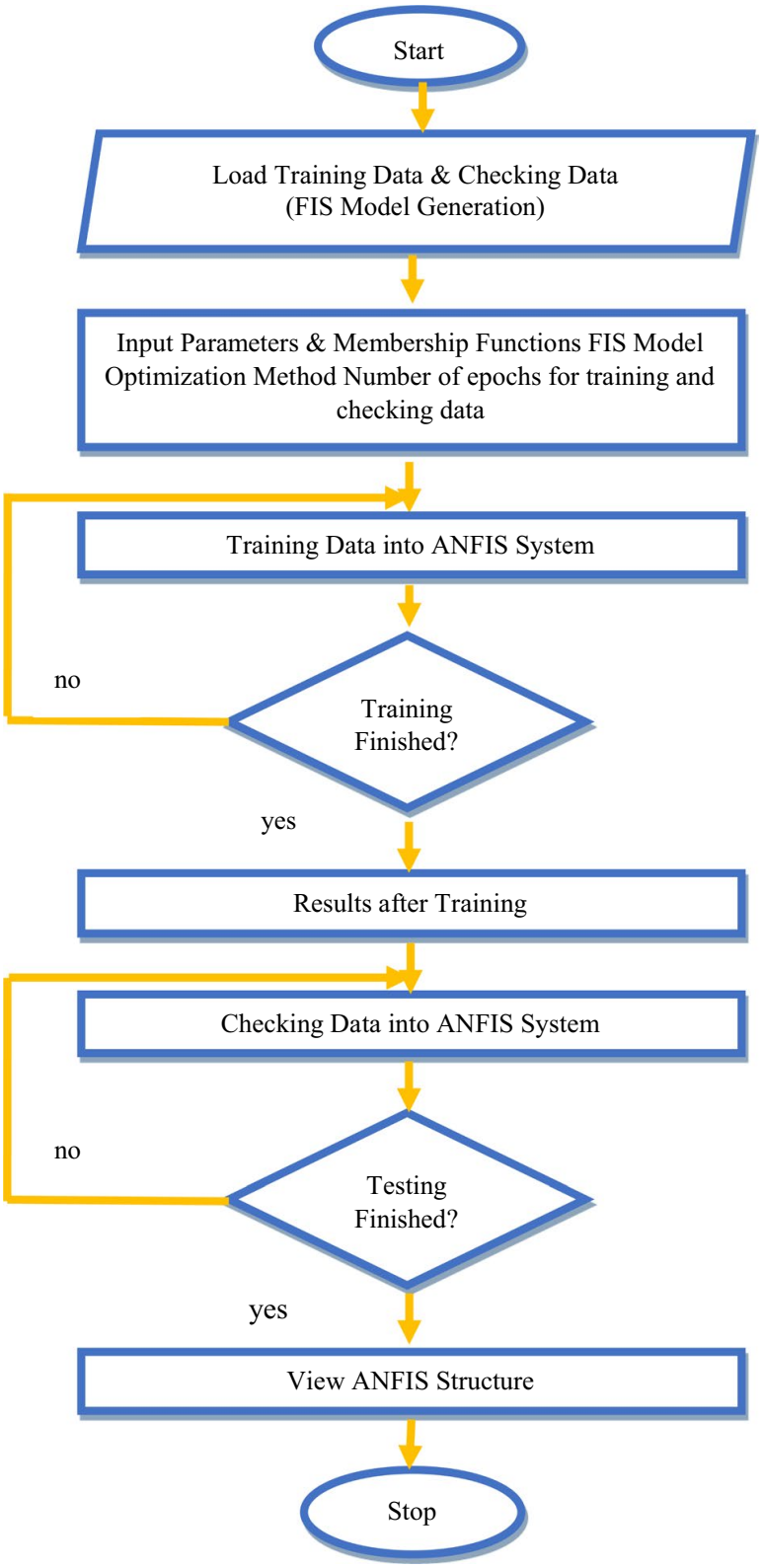


Fig. 6 ANFIS training process

more accurately. Using evalfis function, system performance can be evaluated. That is, the trained model is tested for different sets of inputs in order check its performance. Here, we used the dataset created in the previous section.

Table 4 Parameter setting

Parameter	Value
Range of influence	0.5
Accept ratio	0.5
Reject ratio	0.15
Squash factor	1.25

Table 5 Generated FIS

No. of inputs	7
No. of outputs	1
No. of MFs	38 38 38 38 38 38 38

5.2.2 ANFIS model testing

ANFIS model testing is done to validate the model which is a process by which the input vectors on which the FIS model was not trained are given to the trained FIS. It is done to check the ability of the FIS model to predict the output values corresponding to the given input vectors and to check its generalization capability [45]. Validation of the model is done after the FIS is trained using a testing data which is different from the data set used for training of FIS.

6 Results and discussion

In this research, ANFIS model is implemented to evaluate the IoT compatibility of website designs. Section 6.1 describes the ANFIS model training process. Section 6.2 describes the result of model testing. Section 6.3 gives the ANFIS output for a sample website based on the rules generated.

6.1 Training ANFIS

The parameters selected to generate the FIS are given in Table 4.

The ANFIS model has 7 inputs and 1 output. In this study, we choose the sub-clustering method to generate the FIS. The sub-clustering method generates a FIS with following details as given in Table 5.

Figure 7 illustrates the FIS generated with 7 inputs represented as initial black discs, 38 clusters for each input represented as white discs, 38 rules as blue discs, 38 output MFs as the next white discs, one aggregated output and one final output.

The Epoch 1 error was observed to be 0.2048. We trained the FIS with epoch 100 and epoch 150. It was observed that the training error was decreased as the

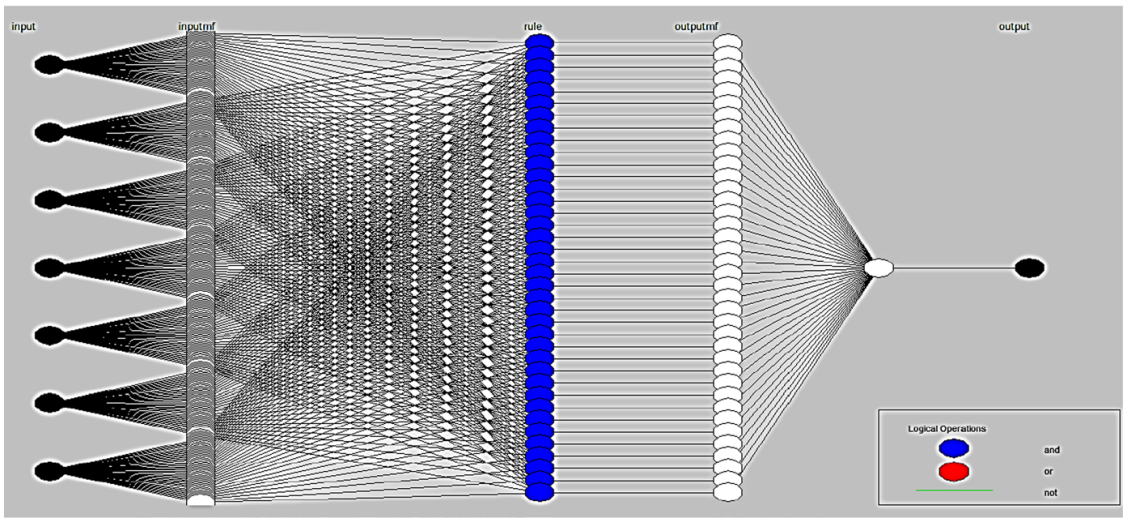


Fig. 7 Structure of generated FIS

Table 6 Training and checking error with varying epochs

Epoch	Training error	Checking error
100	0.1467	0.6001
150	0.1414	0.6731

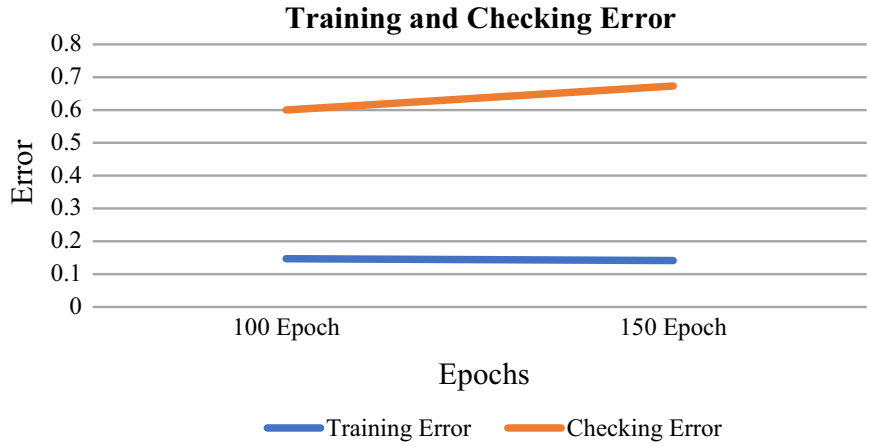


Fig. 8 Training and checking error w.r.t epochs

number of epochs increased but the checking error increased. Table 6 represents the training and checking error with epoch 100 and epoch 150.

Figure 8 shows the trends of training and checking error with respect to number of epochs.

As the checking error increases with no. of epochs we chose 100 epochs for training the FIS and build the final model. Figure 9 represents the graph for training data with which the FIS was trained and Fig. 10 represents graph for checking data.

The total number of parameters and fuzzy rules for ANFIS were 836 and 38, respectively, as shown in Table 7.

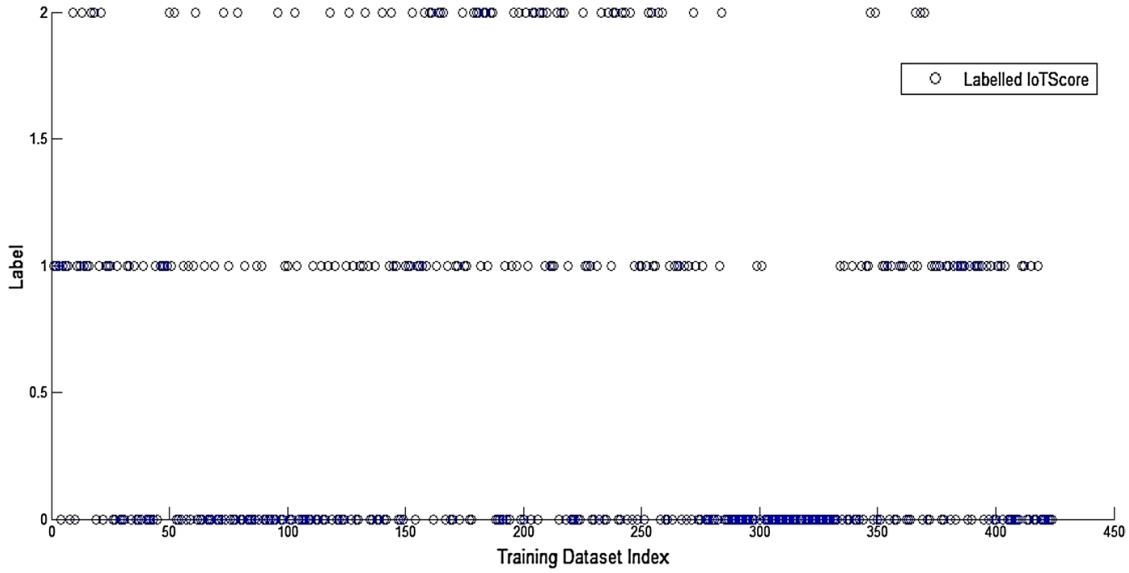


Fig. 9 Training data distribution

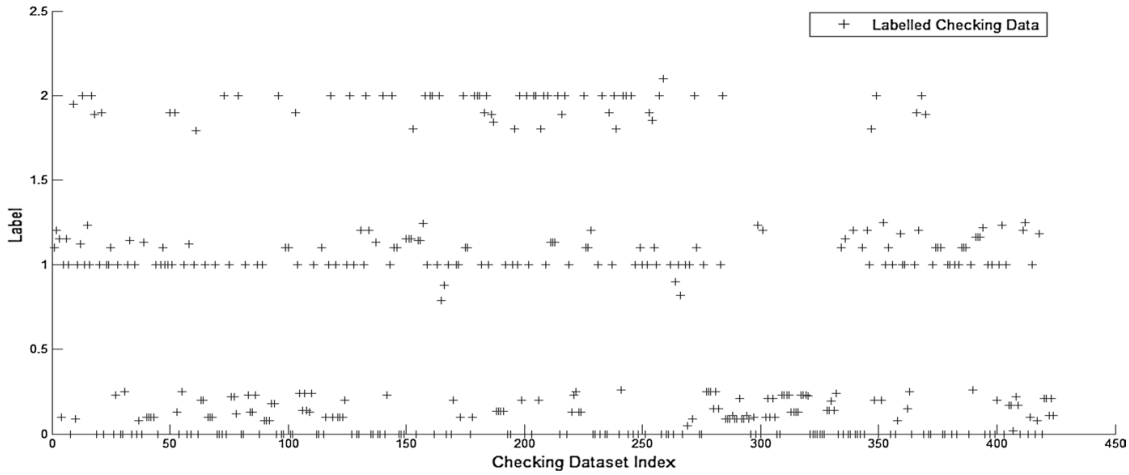


Fig. 10 Checking data distribution

Table 7 ANFIS information

ANFIS item	Value
Number of nodes	618
Number of non-linear parameters	532
Number of linear parameters	304
Number of total parameters	836
Number of training data pairs	424
Number of checking data pairs	424
Number of fuzzy rules	38

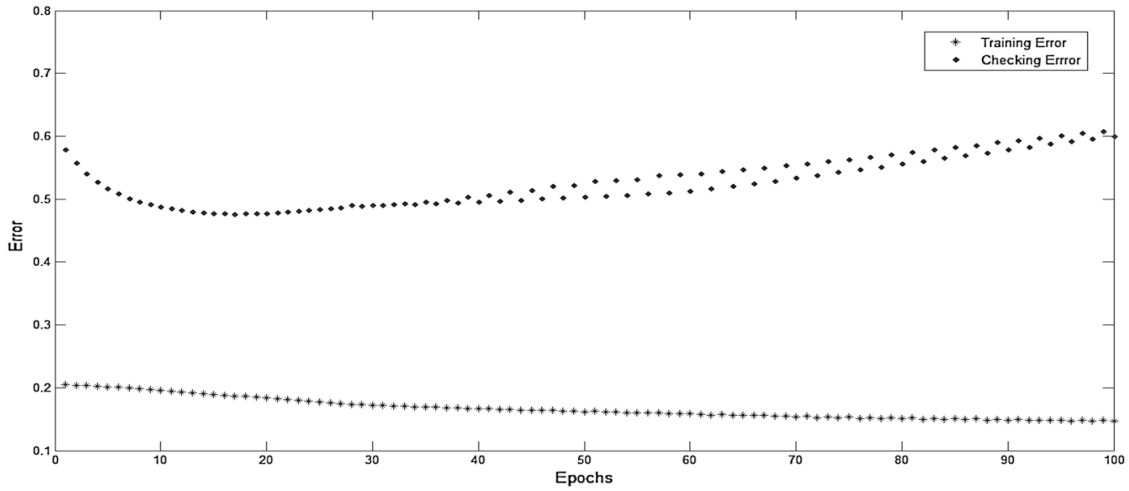


Fig. 11 Training and checking error

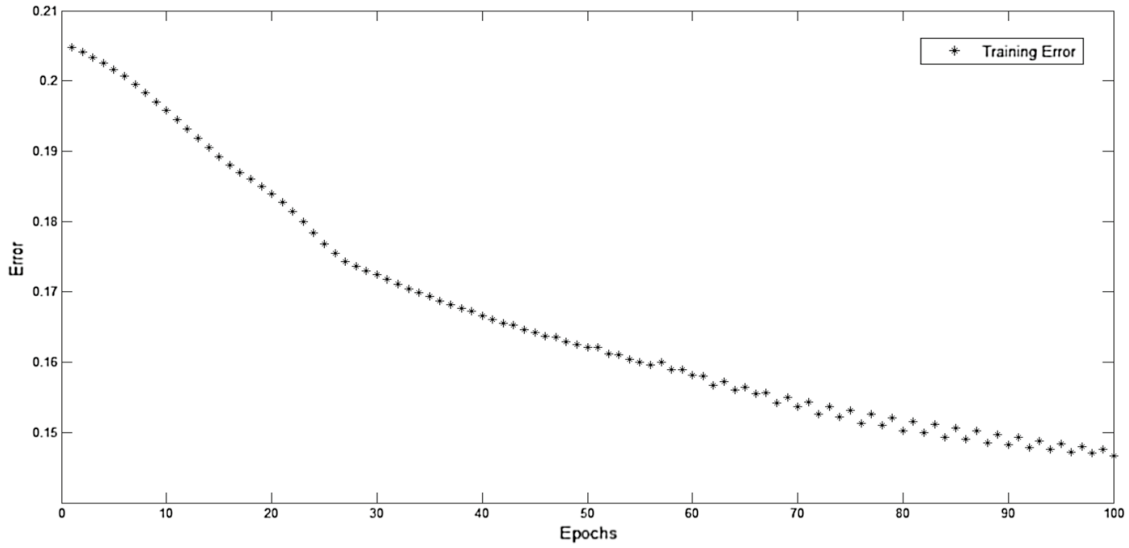


Fig. 12 Training error w.r.t epochs

The error tolerance was set to zero and model optimization was chosen to be *hybrid*. The hybrid optimization method is a combination of least squares and back propagation gradient descent method [46]. After training of ANFIS, the training data errors and checking errors were plotted against the number of epochs as shown in Fig. 11.

The checking error rises up after a certain point till where it decreases which is the point of model overfitting. ANFIS automatically selects its parameters allied with the minimum checking error (just prior to this jump point).

At epoch 100, the training error and the checking error were observed to be 0.1467 and 0.6001, respectively. Figure 12 represents the fall of training error as the number of epochs increases.

Figure 13 shows the result of testing the checking data against the trained FIS. The average testing error in this case came out to be 0.6001.

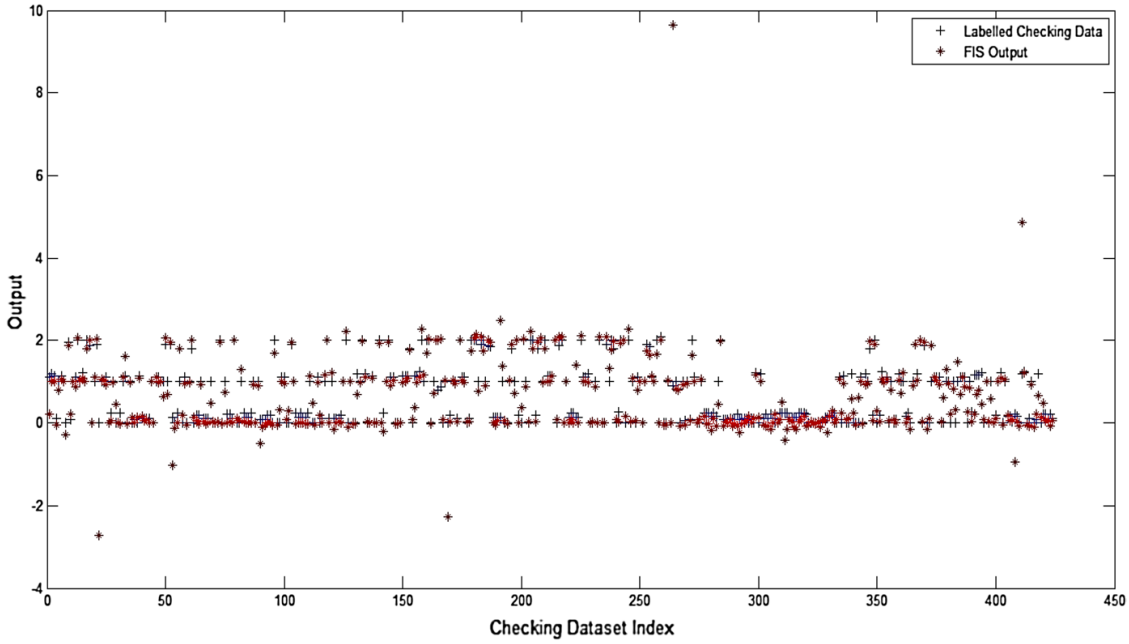


Fig. 13 Testing checking data against the trained FIS

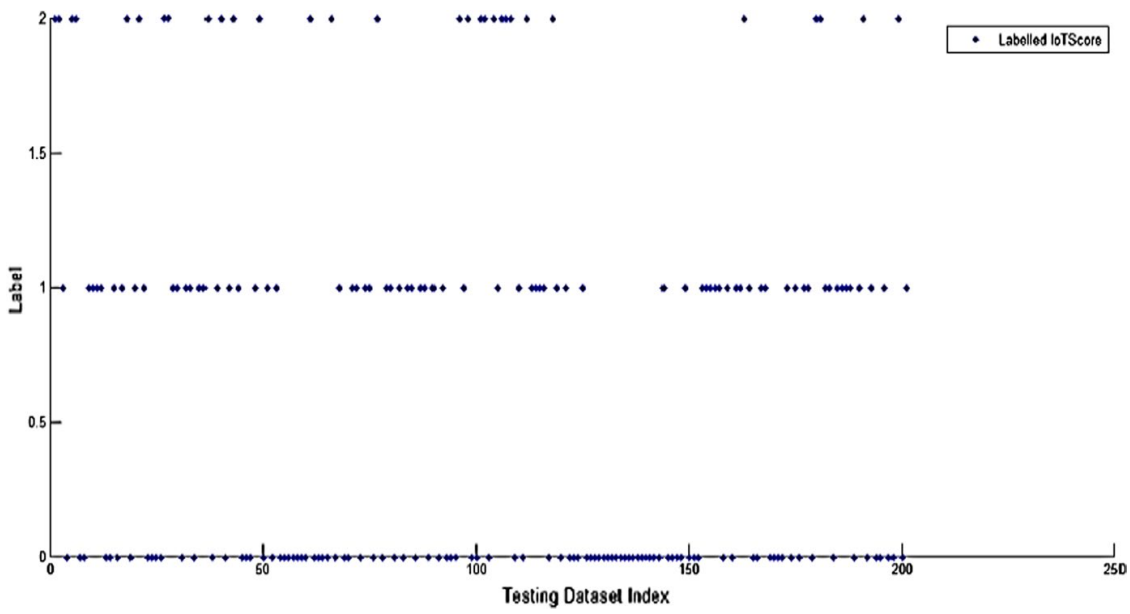


Fig. 14 Test data distribution

6.2 Testing the model

Figure 14 represents testing data points graphically. The number of test data pairs is observed to be 201.

Figure 15 shows the result of testing the model against training data itself. The average testing error in this case came out to be 0.1467.

The performance of the proposed model is shown in Fig. 16 which shows the plot of actual and predicted values of website quality for testing data. The average testing error came out to be 5.0941.

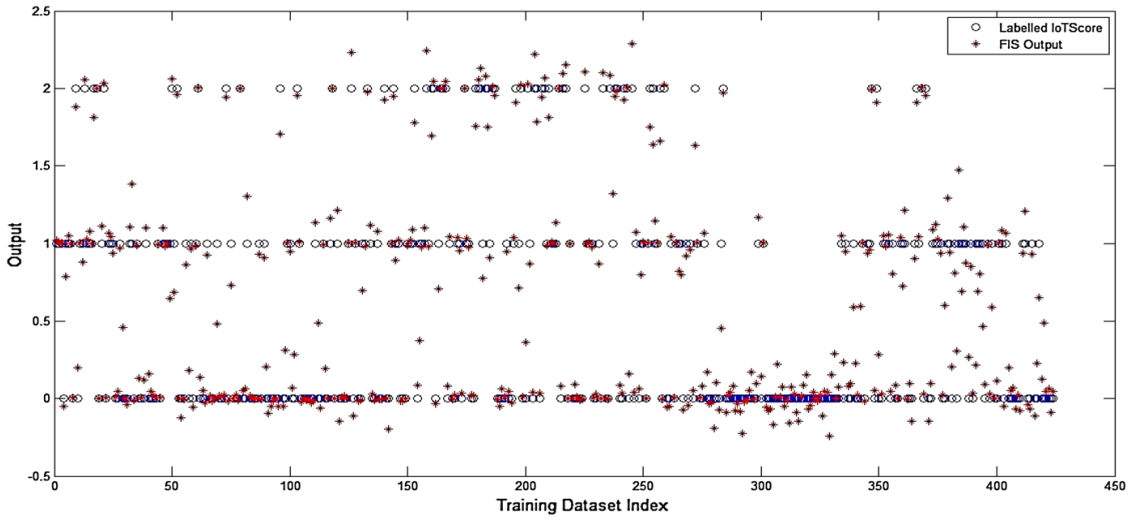


Fig. 15 Training data points and FIS output

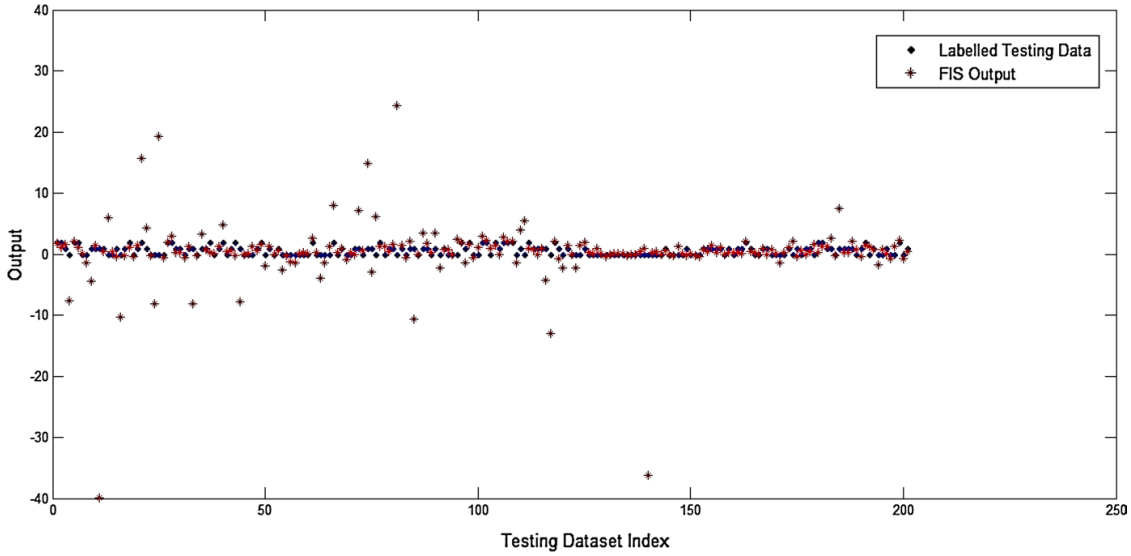


Fig. 16 Testing data points and FIS output

Table 8 Summary of ANFIS model

Epoch	Training error	Average testing error
100	0.1467	5.0941

Table 8 represents the summary of training and testing ANFIS model for evaluating $\text{IotScore}_{\text{site}}$ for the training and testing data.

6.3 ANFIS output

The example of IoT compatibility has been illustrated below for a sample website. Inputs for the sample website are given in Table 9.

Table 9 Feature input values

Input no.	Input value
Input 1	6
Input 2	5
Input 3	7
Input 4	9
Input 5	5
Input 6	5
Input 7	9.5

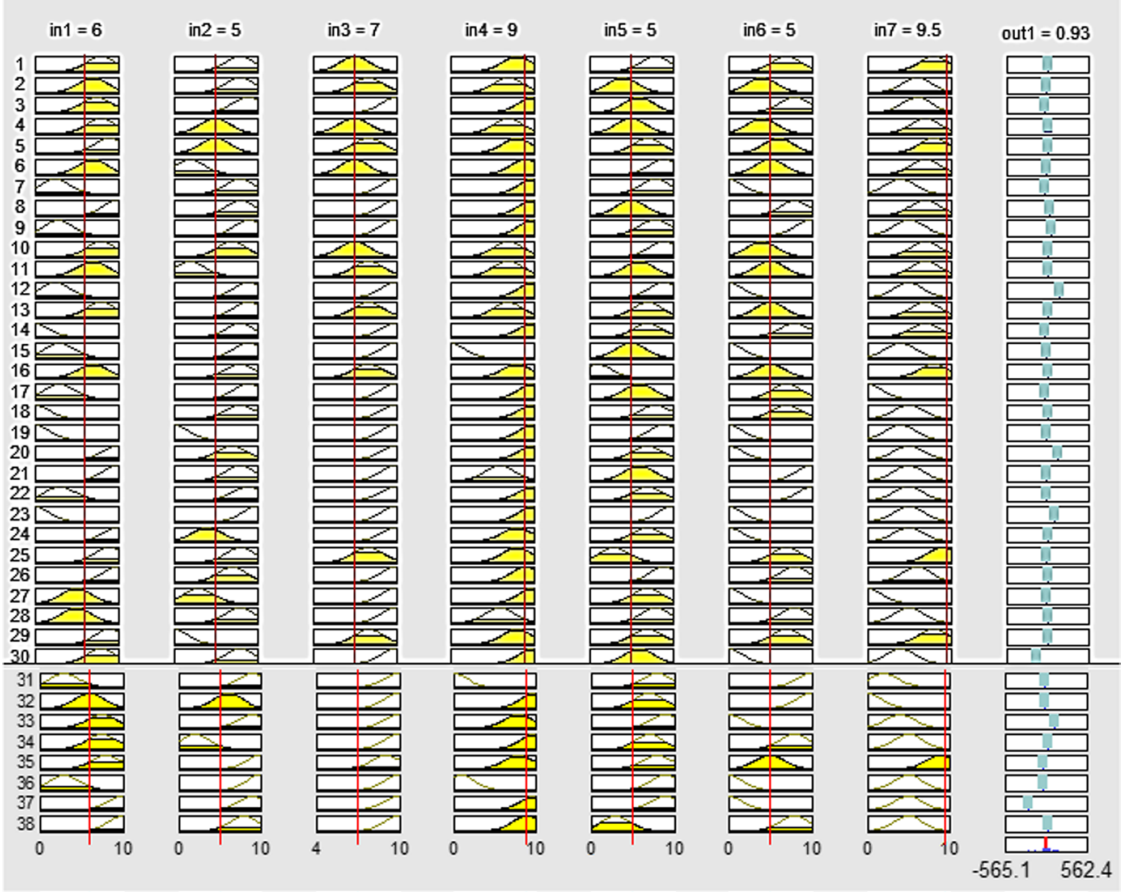


Fig. 17 Rule base and evaluation of $\text{IoTScore}_{\text{site}}$ for sample website

Figure 17 illustrates the rule base and evaluation of IoT scorecard for the sample website with the above input values.

The output $\text{IoTScore}_{\text{site}}$ for the sample website is 0.93. The value of $\text{IoTScore}_{\text{site}}$ lies between 0 and 1 and closer to 1 which implies that the website has medium compatibility when integrated with IoT devices.

On evaluating all 600 websites, it was observed that more than 60% websites show low compatibility with IoT integrated devices with the value of $\text{IoTSite}_{\text{score}}$ closer to 0. Only 10% websites show a value closer to 2 and can be considered as IoT compatible websites. This creates an alarming situation for e-commerce websites

and indicates that current business websites require major design changes and need to adopt new development strategies.

7 Conclusion

It is evident that at the heart of IoT systems are the User Interface (UI) which comprises the visual elements of a page and the User Experience (UX) which focuses on site navigation, search engine optimization and research and analytics. Web designers are the architects of these elements and are responsible to build sites that communicate with back-end databases that have collected information from IoT devices and create front-end interfaces, allowing users to interact with numerous devices that frequently have different screen sizes or run on different platforms. This study evaluates the websites integration capabilities and support to IoT devices and depicts that most of the samples in the show low or medium IoT integration compatibility instigating new challenges for web developers in terms of both back- and front-end design to make a website worth for IoT integration. The proposed ANFIS generates a compatibility score $\text{IoTScore}_{\text{Site}}$ to comprehend the interconnected development, new devices and web development methodologies which create new business opportunities.

IoT is *not* the end product and customers do not buy IoT. It is not simply an iPhone, a networking application or a wearable device. It requires a technology-aware approach to product experience which includes upgrades, adapting to other products, personalization and big data [47–51]. As a future direction, ensemble machine learning methods [52], deep learning models such as recurrent neural networks and convolution neural networks can be implemented to evaluate the compatibility of website design with IoT integrated devices. Also, compatibility studies with respect to integrating IoT to applications, cloud service, big data and semantic web technologies need investigation.

References

1. Atzori L, Iera A, Morabito G (2010) The Internet of Things: a survey. Comput Netw 54(15):2787–2805
2. Gubbi J, Buyya R, Marusic S, Palaniswami M (2013) Internet of Things (IoT): a vision, architectural elements, and future directions. Future Gener Comput Syst 29(7):1645–1660
3. Sachdeva N, Dhir R, Kumar A (2016) Empirical analysis of machine learning techniques for context aware recommender systems in the environment of IoT. In: Proceedings of the International Conference on Advances in Information Communication Technology and Computing. ACM, p 39
4. Bonomi F, Milito R, Zhu J, Addepalli S (2012) Fog computing and its role in the internet of things. In: Proceedings of the First Edition of the MCC Workshop on Mobile Cloud Computing. ACM, pp 13–16
5. Jabbar S, Khan M, Silva BN, Han K (2018) A REST-based industrial web of things' framework for smart warehousing. J Supercomput 74(9):4419–4433
6. Guinard D, Trifa V (2009) Towards the web of things: web mashups for embedded devices. In: Workshop on Mashups, Enterprise Mashups and Lightweight Composition on the Web (MEM 2009), in proceedings of WWW (International World Wide Web Conferences), Madrid, Spain, vol 15

7. Wang KH, Chen CM, Fang W, Wu TY (2018) On the security of a new ultra-lightweight authentication protocol in IoT environment for RFID tags. J Supercomput 74(1):65–70
8. Duyne DKV, Landay J, Hong JI (2002) The design of sites: patterns, principles, and processes for crafting a customer-centered Web experience. Addison-Wesley Longman Publishing Co. Inc, Boston
9. Kumar A, Gupta D (2018) Paradigm shift from conventional software quality models to web based quality models. Int J Hybrid Intell Syst 14(3):167–179
10. Abrahami Y, Shalev ZS, Marcus A, Ohana T, Kaufman A, Blumenfeld UA, Weiner S, Nagar S, Geva A, Wixcom Ltd (2019) Custom back-end functionality in an online website building environment. U.S. Patent Application 10/209,966
11. Haghi M, Thurow K, Stoll R (2017) Wearable devices in medical internet of things: scientific research and commercially available devices. Healthc Inform Res 23(1):4–15
12. Dong L, FutureWei Technologies Inc (2019) User oriented IoT data discovery and retrieval in ICN networks. U.S. Patent Application 10/193,805
13. Xu B, Da Xu L, Cai H, Xie C, Hu J, Bu F (2014) Ubiquitous data accessing method in IoT-based information system for emergency medical services. IEEE Trans Ind Inf 10(2):1578–1586
14. https://readwrite.com/2019/04/26/how-website-design-integrates-with-the-internet-of-things-iot/. Retreived 17 June 2019
15. Jang JS (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23(3):665–685
16. Lee I, Lee K (2015) The Internet of Things (IoT): applications, investments, and challenges for enterprises. Bus Horiz 58(4):431–440
17. Gigli M, Koo SG (2011) Internet of Things: services and applications categorization. Adv Internet Things 1(2):27–31
18. Doukas C, Maglogiannis I (2012) Bringing IoT and cloud computing towards pervasive healthcare. In: 2012 Sixth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing. IEEE, pp 922–926
19. Zhao JC, Zhang JF, Feng Y, Guo JX (2010) The study and application of the IoT technology in agriculture. In: 2010 3rd International Conference on Computer Science and Information Technology, vol 2, pp 462–465. IEEE
20. Giang NK, Blackstock M, Lea R, Leung VC (2015) Developing IoT applications in the fog: a distributed dataflow approach. In: 2015 5th International Conference on the Internet of Things (IOT). IEEE, pp 155–162
21. Castellani AP, Gheda M, Bui N, Rossi M, Zorzi M (2011) Web Services for the Internet of Things through CoAP and EXI. In: 2011 IEEE International Conference on Communications Workshops (ICC). IEEE, pp 1–6
22. Yu X, Roy SK, Quazi A, Nguyen B, Han Y (2017) Internet entrepreneurship and "the sharing of information" in an Internet-of-Things context: the role of interactivity, stickiness, e-satisfaction and word-of-mouth in online SMEs' websites. Internet Res 27(1):74–96
23. Loiacono ET, Watson RT, Goodhue DL (2002) WebQual: a measure of website quality. Mark Theory Appl 13(3):432–438
24. Hartmann J, De Angeli A, Sutcliffe A (2008) Framing the user experience: information biases on website quality judgement. In: Proceedings of SIGCHI Conference of Human Factors in Computing Systems, April 2008. ACM, pp 855–864
25. Bhatia MPS, Kumar A (2007) Contextual proximity based term-weighting for improved web information retrieval. In: International Conference on Knowledge Science, Engineering and Management. Springer, pp 267–278
26. Sobecki J, Żatuchin D (2009) Knowledge and data processing in a process of website quality evaluation. In: New challenges in computational collective intelligence. Springer, Berlin, pp 51–61
27. Schubert P, Selz D (1999) Web assessment-measuring the effectiveness of electronic commerce sites going beyond traditional marketing paradigms. In: Proceedings of the 32nd Annual Hawaii International Conference on Systems Sciences, HICSS-32. Abstracts and CD-ROM of Full Papers, Track 5, pp 10
28. Selz D, Schubert P (1998) Web assessment—a model for the evaluation and the assessment of successful electronic commerce applications. In: Proceedings of Thirty-First Hawaii International Conference on System Sciences, January 1998, vol 4. IEEE, pp 222–231

29. Schubert P, Dettling W (2002) Extended Web assessment method (EWAM)-evaluation of e-commerce applications from the customer's viewpoint. In: Hawaii International Conference on Systems Science, Hawaii, 2002. IEEE, p 10
30. Kumar A, Arora A (2019) A filter-wrapper based feature selection for optimized website quality prediction. In: 2019 Amity International Conference on Artificial Intelligence (AICAI). IEEE, pp 284–291
31. Tavallali P, Yazdi M, Khosravi MR (2017) An efficient training procedure for Viola–Jones face detector. In: 2017 International Conference on Computational Science and Computational Intelligence (CSCI). IEEE, pp 828–831
32. Tavallali P, Yazdi M (2015) Robust skin detector based on AdaBoost and statistical luminance features. In: 2015 International Congress on Technology, Communication and Knowledge (ICTCK). IEEE, pp 98–103
33. Carreira-Perpinán MA, Tavallali P (2018) Alternating optimization of decision trees, with application to learning sparse oblique trees. In: Advancesin Neural Information Processing Systems, (NeurIPS 2018), vol 31, pp 1211–1221
34. Kumar A, Dabas V, Hooda P (2018) Text classification algorithms for mining unstructured data: a SWOT analysis. Int J Inf Technol. https://doi.org/10.1007/s41870-017-0072-1
35. Yao X (1999) Evolving artificial neural networks. Proc IEEE 87(9):1423–1447
36. Cybenko G (1989) Approximation by superpositions of a sigmoidal function. Math Control Signals Syst 2(4):303–314
37. Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117
38. https://www.sciencedirect.com/topics/engineering/fuzzy-inference-system. Retreived 14 May 2019
39. Nedeljkovic I (2004) Image classification based on fuzzy logic. Int Arch Photogramm Remote Sens Spat Inf Sci 34(30):3–7
40. Kumar A, Sharma A (2019) Systematic literature review of fuzzy logic based text summarization. Iranian J Fuzzy Syst 16(5):45–59
41. Sugeno M (1985) An introductory survey of fuzzy control. Inf Sci 36(1–2):59–83
42. Barnes S, Vidgen R (2000) WebQual: an exploration of website quality. In: ECIS 2000 Proceedings, p 74
43. Bernard M (2000) Mills M (2000) So, what size and type of font should I use on my website. Usability News 2(2):1–5
44. Al-Hmouz A, Shen J, Al-Hmouz R, Yan J (2012) Modeling and simulation of an adaptive neuro-fuzzy inference system (ANFIS) for mobile learning. IEEE Trans Learn Technol 5:226–237
45. Pathak J, Vidyarthi N, Summers SL (2005) A fuzzy-based algorithm for auditors to detect elements of fraud in settled insurance claims. Manag Audit J 20(6):632–644
46. Raja P, Pahat B (2016) A review of training methods of ANFIS for applications in business and economics. Int J u-and e-Service Sci Technol 9(7):165–172
47. Hashemi SY, Aliee FS (2018) Dynamic and comprehensive trust model for IoT and its integration into RPL. J Supercomput 75:3555–3584
48. Barnaghi P, Sheth A, Singh V, Hauswirth M (2015) Physical-cyber-social computing: looking back, looking forward. IEEE Internet Comput 19(3):7–11
49. Atzori L, Iera A, Morabito G (2014) From "smart objects" to "social objects": the next evolutionary step of the internet of things. IEEE Commun Mag 52(1):97–105
50. Hahanov V (2018) Cyber physical computing for IoT-driven services. Springer, Berlin
51. Sheth A (2016) Internet of things to smart IoT through semantic, cognitive, and perceptual computing. IEEE Intell Syst 31(2):108–112
52. Tavallali P, Yazdi M, Khosravi MR (2019) Robust cascaded skin detector based on AdaBoost. Multimed Tools Appl 78(2):2599–2620

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Journal of Computational Electronics
https://doi.org/10.1007/s10825-019-01400-9

Analytical modeling for static and dynamic response of organic pseudo all-*p* inverter circuits

Neha Mishra[1] · **Poornima Mittal**[2] · **Brijesh Kumar**[1]

© Springer Science+Business Media, LLC, part of Springer Nature 2019

Abstract
This paper presents the performance analysis of an all-*p*-organic pseudo-inverter circuit using dual gate organic thin film transistors. The proposed inverter design has shown significantly high performance in terms of noise margin, gain and propagation delay, leading to the design of more robust digital circuits that, too, exhibit augmented performance. The parameters of the all-*p*-organic pseudo-inverter are compared with those of zero-V_{gs} load logic (ZVLL) and dynamic load logic based conventional inverters, and a substantial improvement is found for the novel combination of a dual gate flexible TFT with a pseudo-design. Performance parameters were deeply analyzed, and we observed that the noise margin is improved by 42.8% as compared to ZVLL based conventional inverters. A bootstrap technique was implemented to further improve the performance and reduce the threshold voltage drop. The performance parameters were analyzed mathematically and compared with simulated values. The static as well as dynamic characteristics of organic pseudo-all-*p* inverter, with and without bootstrap technique, were observed. Static power consumption of the organic pseudo-all-*p* inverter was estimated. In this way, improvement of noise margin by the bootstrap circuit of the organic pseudo-inverter is characterized. The reduced threshold voltage applied to design of low power circuits enables longer power backup for various applications.

Keywords Bootstrapping · Dual gate · Noise margin · Organic pseudo-inverter · Organic thin film transistor · Power consumption

1 Introduction

The field of flexible electronics has achieved remarkable developments in recent years. Organic semiconductors (OSCs) are inexpensive substitutes for low cost large–area electronics. However, in conventional semiconductors, production cost and flexibility are obvious constraints. Universal logic gates and inverters are essential for applications of organic thin film transistors (OTFTs), but only a few logic designs have been reported to date [1–6]. The performance of an organic inverter is affected by the process of fabrication. Uniformity in deposition of the layers is highly desirable for better performance of organic inverters. The different reasons for degradation in the transistor performance include its geometry defects, oxide breakdown, electro migration, hot carrier injection DIBL and ESD. These defects cause deflation in the performance of the transistor and the circuit. The external disturbances mentioned above include noise, overvoltage and overcurrent that affect the circuit. It will be possible in the future to use disposable organic circuits, for example by printing them on paper, cloth; glass etc., because of their economically viable fabrication processes at low temperature [3–7]. Of two different structures based on the position of the gate, i.e. single gate and dual gate, the dual gate configuration is superior in its performance as compared to single gate. The presence of two gates provides a higher charge injection area, hence better charge accumulation over the single gate thin film transistor. Dual gate organic transistors (DG-OTFTs) operate on the principle of accumulation rather than inversion, unlike MOSFET devices [1, 4]. The charge transport in the transistors occurs due to a hopping mechanism, which explains the movement of charge particles (holes and electrons) from HOMO and LUMO levels. For conduction to occur, the holes (or electrons) must

✉ Poornima Mittal
poornimamittal@dtu.ac.in; bkece@mmmut.ac.in

[1] Department of Electronics and Communication Engineering, Madan Mohan Malaviya University of Technology, Gorakhpur, UP 273010, India

[2] Department of Electronics and Communication Engineering, Delhi Technological University, Delhi 273010, India

Published online: 04 October 2019

have an energy level which is higher than the band gap energy (E_g). In the organic semiconductor, proper overlapping of orbitals is also important; the commonly preferred overlapping among orbitals is sp^2 hybridized orbitals [3].

This paper comprises six sections. Section 1 is the introduction, Sect. 2 includes the structures and electrical properties of the single and dual gate based organic transistors. In Sect. 3, dual gate based inverters, along with new organic pseudo-all-p structure, are presented. Sections 4 and 5 give the performance improvement techniques and power consumption of the inverter along with mathematical extraction of parameters. Finally, Sect. 6 highlights the important outcomes, limitations and challenges of dual gate-based inverters.

2 Dual gate versus single organic thin film transistors

The single gate OTFT comprises a gate (preferably at the bottom) passivated by high-k dielectric material. A commonly used gate material is aluminum (Al), which exhibits a work function somewhere in the range of 3.5 to 3.9. In production, the OSC layer is placed over the dielectric [3, 7]. Depending upon the relative position of source and drain electrodes with respect to the OSC, devices are described as bottom contact or top contact devices. In bottom contact structures, the source and drain are sandwiched between the OSC and the dielectric; in top contact devices, the source and drain are positioned above the OSC layer, as shown in Fig. 1. Bottom contact-based devices can be fabricated using various techniques such as spin coating, solution processing and sample printing [3].

Single gate-based devices and circuits are less performant than dual gate-based circuits [1], due to low mobility and presence of a single channel that reduces the charge injection area. In contrast, the charge injection area gets doubled in a dual gate transistor in the presence of two separate channels at top and bottom. Another reason for the reduced performance of the single gate design is its higher threshold voltage as compared to the dual gate design. The presence of two channels in DG-OTFTs provides significant reduction in threshold voltage. Higher threshold voltage has a negative effect on OTFT performance, giving rise to structural defects in the device. These defects affect the parameters such as trans-conductance, mobility and sub-threshold slope (SS). Sub-threshold slope should be as low as possible. Patterning of layers in the device is an important concern which is used to estimate structural disorders. If the grain particles are not aligned uniformly, the device will consume a higher amount of voltage, raising the threshold voltage of the device and thereby increasing the sub-threshold slope. In describing device performance, there are numerous

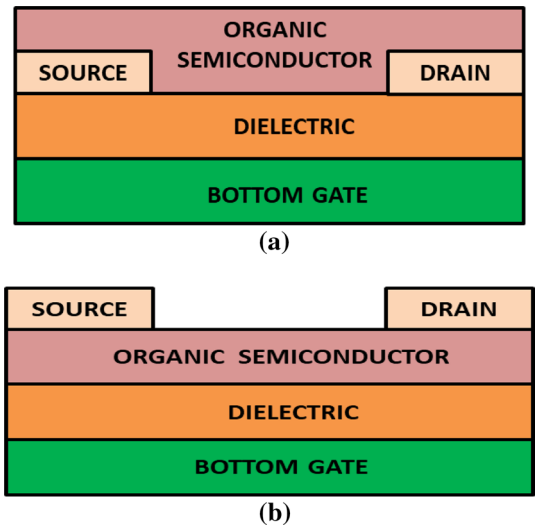


Fig. 1 Single gate based **a** bottom contact and **b** top contact OTFT structures

parameters that play a significant role in predicting OTFT device behaviour. These include current I_{ON}/I_{OFF} ratio, sub-threshold slope (SS), trans-conductance (g_m), threshold voltage (V_T) and mobility (μ).

2.1 Simulation setup and parameters extraction of DG-OTFT

The dual gate based organic thin film transistor was analyzed using a Silvaco Atlas 2-D simulator. The structure of the device and its input and output characteristics are shown in Figs. 2 and 3, respectively. Here, the channel was formed within the pentacene; a p-type material. The dimensions of the device are summarized in Table 1. The mobility was analyzed using a Poole–Frenkel mobility model, which is given as [8],

$$\mu = \left(\frac{\Delta}{kT} - 1\right) - \left(\frac{\beta}{kT} - \gamma\right), \tag{1}$$

where μ is the field dependent mobility, μ_o is the zero field mobility and E is the electric field. Activation energy and Poole–Frenkel constant (for a p-type) are specified by Δ and β, respectively. The values of Δ and β are given by 1.792×10^{-2} eV and 7.758×10^{-5} eV, respectively [3]. The iterations were carried out using the Newton–Raphson (N–R) algorithm. The analysis was performed on both the steady and transient inputs. The N–R algorithm was used under the mixed mode simulation for steady state analysis and transient state analysis. In this work, the Atlas device was used with full and modified Newton algorithms for steady state simulation.

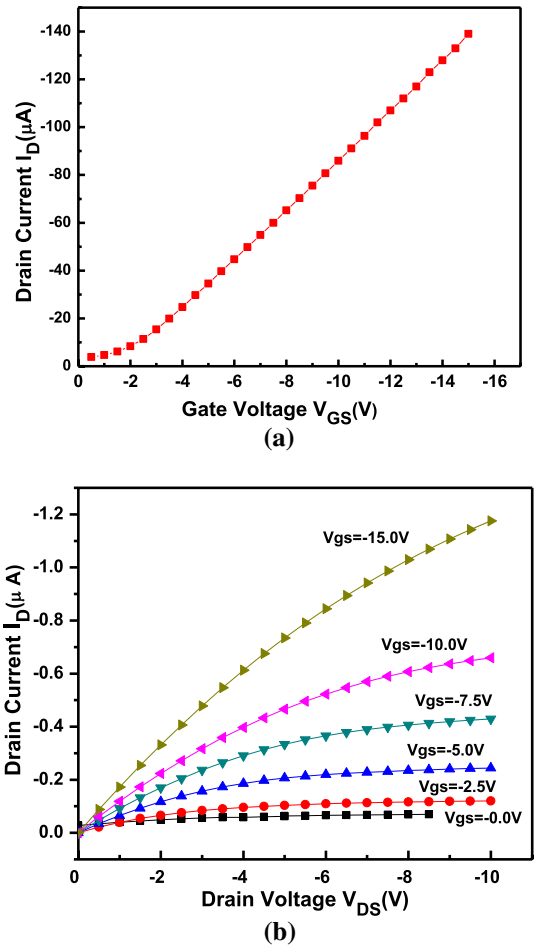


Fig. 2 Characteristics plots of DG-OTFT **a** transfer curve (I_D–V_{GS}) and **b** output curve (I_D–V_{DS})

The pseudo-inverter consists of a cascaded structure in which the output of the first inverter is fed to the second inverter. There is more than one Atlas device in the structure, so each device is simulated under mixed mode simulation in which structures are combined internally. The resulting solution file is a single file which contains both the structures. Extraction of mixed mode log files requires re-initialization of Atlas.

The electrical features of the dual gate organic transistor include the input (I_D–V_{GS}) and output (I_D–V_{DS}) characteristics shown in Fig. 2. The electrical parameters include the threshold voltage, sub-threshold slope, drain current, trans-conductance, mobility, I_{ON}/I_{OFF} ratio and trap density.

Threshold voltage (V_T) is an important factor that governs the morphology of the device. If the size of the grain particles is not uniform, then trap density increases, consuming a considerable amount of power and resulting in higher threshold voltage. Sub-threshold slope is used to reveal defects, interface states and trap density. SS is a measure of switching behavior of the transistor, which depends on the thickness of the insulating layer [3–6]. With an insulating layer thickness of 100 nm for a bottom contact device, the sub-threshold slope is measured as 5.85 V/dec. Mobility of an organic transistor gives insight into the movement of charge carriers in the device. In this work, the *p*-type material (pentacene) we used has a hole mobility of 0.62 cm^2/V s.

Drain current flows in the OTFT from source to drain due to accumulation of charges in the channel. Drain current decreases with increase in channel length '*L*' due to increased charge density from source to drain. When the depth of the layer is increased, the concentration of the holes is increased in the active layer. As a result, the charge density in the channel is increased, leading to an increase in the drain current. Trans-conductance is used to estimate the variation in the drain current due to change in gate voltage; it should be as high as possible and is determined graphically from the slope of the transfer curve. In this work the trans-conductance was obtained as 2.12×10^{-7} μS.

I_{ON}/I_{OFF} is the ratio of the current in the accumulation mode to the current in the depletion mode. A higher I_{ON}/I_{OFF} ratio indicates higher utility in switching applications. In this paper I_{ON}/I_{OFF} ratio was observed to be on the order of 10^7.

2.2 Dual gate organic thin film transistor

The dual gate based OTFT exhibits superior performance compared to its single gate counterpart. Pentacene is the preferred material due to its higher mobility as compared to other organic materials, such as CuPc, P_3HT and C_{60} [4]. Organic transistors can be further as *p*-type or *n*-type, based on the polarity (positive or negative) of the channel formed between source and drain electrodes. The main difference between *p*-type and *n*-type materials is of the mobility of respective charge carriers, which is an important parameter when integrating these transistors into a specific application (Fig. 3). Mobility of the *p*-type material is reported as higher than that of the *n*-type material. In this work, a dual gate pseudo-all-*p* inverter was realized using devices with *p*-type materials only. A pseudo-all-*p* inverter is used to improve the noise margin and propagation delay, which in turn increases the circuit robustness. The organic conventional inverter circuit and the proposed pseudo-all-*p* inverter are discussed in Sects. 3 and 4, respectively.

3 Organic conventional inverter circuit

Organic inverters may also be designed using a pair of *p* and *n*-type transistors, similar to the CMOS inverter. In this section, the dual gate-based inverter circuit is analysed. Depending on the connections between the source, gate and drain terminals, the inverter design may consist of zero-V_{gs}

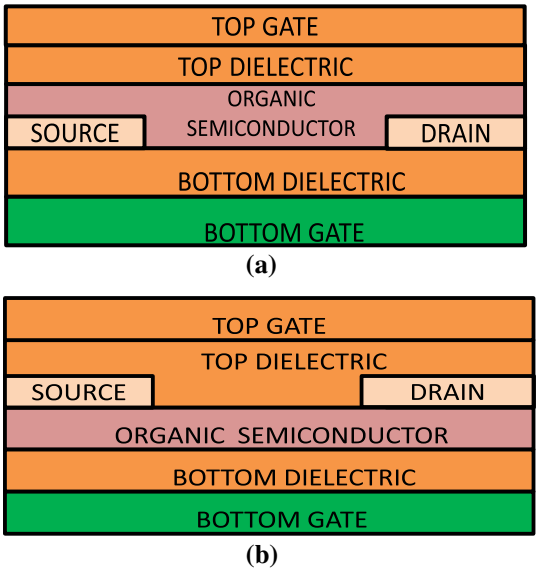


Fig. 3 Structure of DG-OTFT **a** bottom contact and **b** top contact

Table 1 Structural dimensions and materials of DG-OTFT

Thin film	Material	Dimension (nm)
Bottom gate	Ag	150
Bottom dielectric	SiO_2	100
Organic semiconductor	Pentacene	200
Top dielectric	Al_2O_3	300
Top gate	Ag	150

load logic (ZVLL) or diode load logic (DLL) [3, 9, 10]. These configurations are different in terms of their outcomes and are suitable for specific purposes. DLL is preferred for higher switching applications, whereas, ZVLL is a better choice for robustness against external noise. The DG-OTFT based ZVLL and DLL inverters are given in Fig. 4.

In ZVLL configuration, the gate and source terminals are shorted with each other to form a constant current source from V_{DD} to ground. This topology gives an advantage of higher noise margin and voltage swing, which helps to increase the robustness of the circuit. The driver transistor operates in enhancement mode and the load transistor operates in depletion mode.

At lower V_{IN}, the channel resistance of the driver transistor is much smaller than its load, causing to the driver to be in the 'ON' state, and thus V_{OUT} is obtained as near to V_{DD}. On the other side, at higher V_{IN}, comparatively a larger aspect ratio (*W*/*L*) is required for pull down action. At higher V_{IN} both the driver and load transistor form a zero-gate to source and are in the ON state. Therefore, the *W*/*L* ratio of the load transistor is around five times higher than that of the driver transistor.

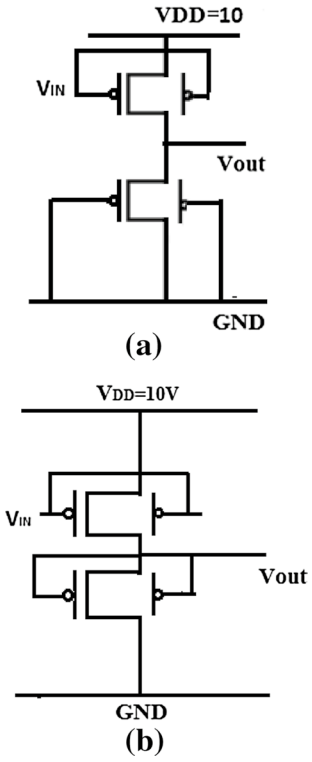


Fig. 4 DG-OTFT based inverter configuration **a** DLL and **b** ZVLL

In a DLL inverter, the drain and source terminals are connected to ground. In this configuration, both driver and load transistors operate in enhancement mode. They are compatible with both enhancement and depletion type load transistors. They are proven as the faster circuits, as compared to the ZVLL based inverter.

The *W*/*L* ratio is an important factor when selecting and designing the inverter configuration. It maximizes the noise margin (NM) and allows the capacitive load to charge and discharge in equal times by providing identical current to source and sink capabilities. If the *W*/*L* ratio is varied, then the steepness of the transfer characteristics is increased. If the aspect ratio is greater than 1, then the inverter is known as a HI-skewed inverter, and if the ratio is less than 1 it is known as a LO-skewed inverter. A HI-skewed inverter gives a better noise margin as compared to a LO-skewed inverter. To achieve more robustness in the inverter performance, a pseudo-all-*p* inverter is proposed and described in the subsequent section.

3.1 Proposed dual gate organic pseudo-all-*p* inverter

The organic pseudo-all-*p* inverter circuit topology is vital to achieving better inverter performance. It is designed to enhance the noise margin of the inverter. As observed from the literature, the circuit structure involves two configurations: pseudo-D and pseudo-E. It is a two stage inverter

structure in which the output of the first inverter is connected as the input to the second inverter. The left configuration can be either pseudo-D or pseudo-E, and the connection on the right is a level shifter. V_{DD} is the supply voltage and output is expected to attain maximum swing between V_{DD} and GND. There is a significant difference between the ZVLL, DLL and organic pseudo-all-*p* based circuit. The two stage structure keeps the higher voltage V_{OH} constant and provides a reduction in V_{OL} that is absent in the zero-V_{gs} logic circuit. This is an important influencing factor for the design of the pseudo-all-*p* inverter. Zhao et al. [10] reported the single gate based organic pseudo-CMOS inverter shown in Fig. 5.

3.1.1 Operation of pseudo-all-*p* inverter

The structure shown in Fig. 6 is the circuit for the pseudo-all-*p* inverter. It consists of four transistors: TFT1, TFT2, TFT3 and TFT4 respectively. When the input V_{IN} applied at the TFT1 is logic LOW, the TFT1 operates in linear region, while TFT2 is in saturation region. Subsequently, TFT3 and TFT4 operate in the cut-off and linear regions, respectively. With TFT1 in the linear region, the source V_{DD} appears at the output terminal; V_{01} is used as input for second stage, comprising TFT3 and TFT4. With TFT3 in cut-off and TFT4 in linear region, the output is forced to logic HIGH, which is an expected voltage transfer characteristic (VTC). When the input is HIGH, both TFT1 and TFT2 are in the linear region, while TFT3 and TFT4 are in the linear and saturation regions, respectively. With TFT

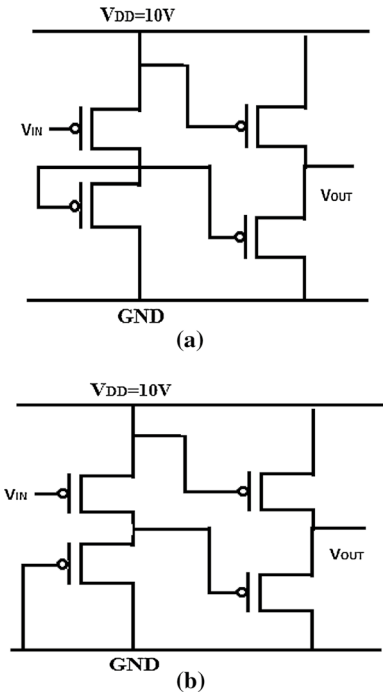


Fig. 5 SG-OTFT based pseudo-all-*p* inverter **a** ZVLL and **b** DLL configurations

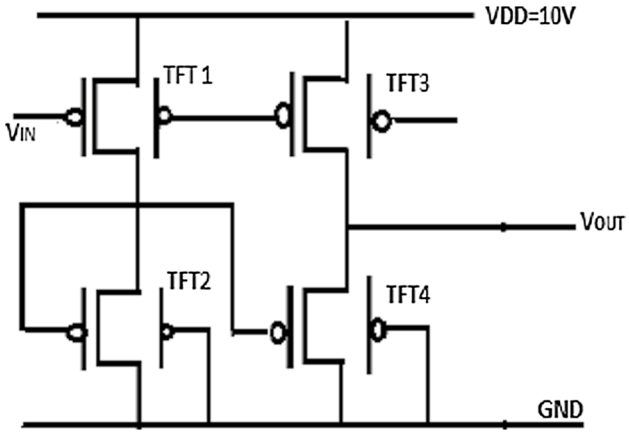


Fig. 6 Circuit schematic of dual gate based pseudo-all-*p* inverter

4 in saturation, it drains the current from V_{DD} , leading to an output voltage equal to '0'. The pseudo-all-*p* inverter is illustrated in Fig. 6.

Voltage transfer characteristics and voltage gain of dual gate based organic pseudo-all-*p* inverter are shown in Figs. 7 and 8, respectively. From the VTC of the inverter, various performance parameters, including noise margin (NM), voltage swing, gain and power consumption, were extracted and further analyzed in Sect. 4 and 5, respectively.

The noise margin of the circuit is explained as the maximum noise that can be tolerated by the input without any disturbance in the output.

$$NM_H = V_{OH} - V_{IH} \text{ and } NM_L = V_{IL} - V_{OL}$$

The noise margin of the circuit should be high for better performance of the circuit. In this work the value of NM_L is higher than the value of NM_H. The lower value for

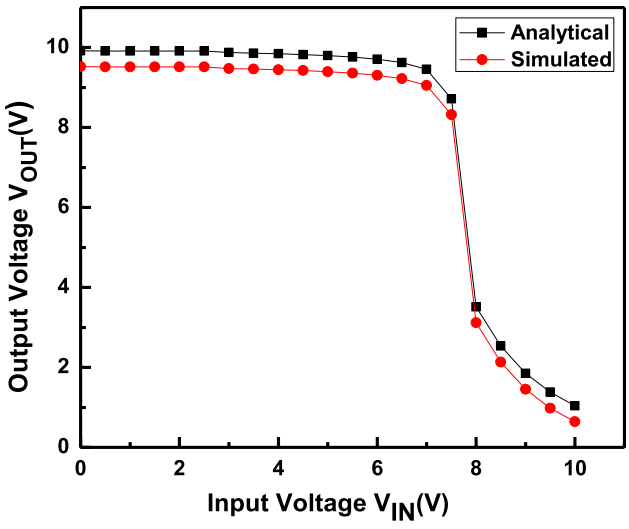


Fig. 7 Voltage transfer characteristics (VTC) of dual gate organic all-*p* pseudo-inverter circuit

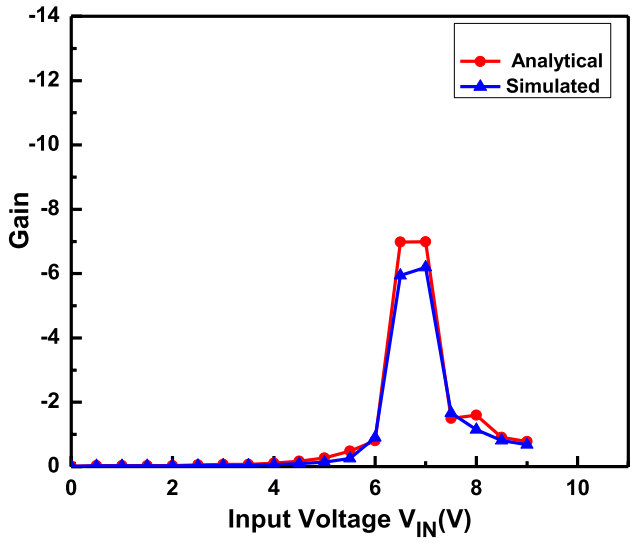


Fig. 8 Voltage gain plot of dual gate organic all-*p* pseudo-inverter in ZVLL configuration

NM_H is due to the presence of a weak pull down transistor (*p*-type). The voltage gain is depicted in Fig. 8.

The magnitude of V_{OH} must be high and close to V_{DD} to achieve high noise margin. The maximum value attained by V_{OH} is limited to (V_{DD}–V_T) owing to the threshold voltage drop of the device. The minimum pull down that can be achieved by *p*-MOS is equal to V_T , always. In order to achieve better performance threshold, voltage drop should be minimized. The dynamic behaviour of the dual gate based pseudo-all-*p* inverter is demonstrated in Fig. 9. The V_{OUT} exhibits some spikes with magnitude between 10 and 11 V at the beginning of the pull-down action. When the input is changed instantaneously from 0 to V_{DD}, the voltage over gate-drain capacitor has to change from V_{DD} to 0 V. This could happen only if an infinite amount of current flows through the capacitor [11, 12]. Due to the low *K* of the organic *p*-type pull-down transistor, it may not able perform a fast draining action, and for some instances the V_{OUT} may retain a value between V_{DD} to 2* V_{DD} unless the pull-down transistor comes completely under good conducting mode.

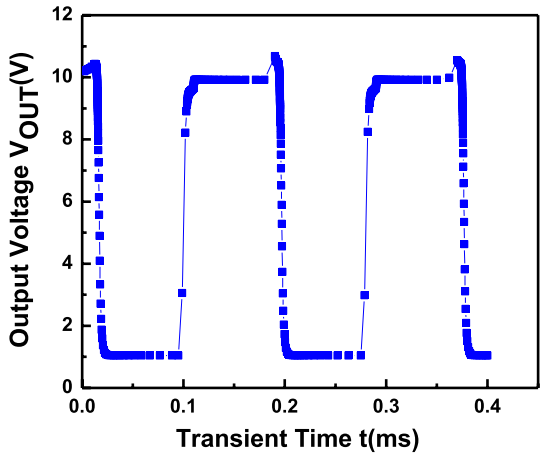


Fig. 9 Transient characteristics of dual gate organic all-*p* pseudo-inverter in ZVLL configuration

To analyse the transient characteristics, a square pulse was applied with a magnitude of 10 V and frequency of 1 kHz. From the transient characteristics, the propagation time delays are calculated. The high-to-low (τ_{PHL}) and low-to-high (τ_{PLH}) transition can be defined as the time taken by the capacitor to charge and discharge, respectively. Output rise time and fall time are calculated from the transient characteristic.

4 Parameter extraction of all-*p* organic pseudo-inverter using analytical models

In this section, the mathematic models behind extraction of pseudo-inverter parameters are analyzed to understand the circuit behaviour in depth. The extraction of critical voltage levels, noise margins, propagation delay, power consumption are described in the following subsections.

4.1 Analysis and extraction of static parameters

Critical performance parameters such as V_{OH}, V_{OL}, V_{IH}, V_{IL} and V_{TH} are extracted from the VTC plot, as demonstrated in Fig. 7.

(a) V_{OH} extraction

The value of V_{OH} is extracted on applying a zero input, wherein TFT1 operates is in linear region and TFT2 is in cut-off. The TFT1 is working as a driver transistor, whereas TFT2 is working as a load transistor. The output of the load transistor is connected as the input for TFT4. Since the transistors are placed in series, the drain current in steady state is equal. As the load transistor is in cut-off region, the drain current is zero and $V_{OUT}=V_{OH}=V_{DD}=10$ V.

(b) V_{OL} extraction

For extraction of V_{OL} the input voltage is equal to the supply voltage V_{DD}. The operating status of different transistors is summarized in the Table 2. The transistors TFT3 and TFT4 are connected in cascade. As both the transistors (TFT3 and TFT4) are acting as the buffer, thus current is same at the steady state and equal to zero. Now, the modified equation may be expressed as:

$$\mu C_{OX}\frac{W}{L}\left(\frac{\left(V_{GS}-V_{TH}\right)^2}{2}\right)=0 \tag{2}$$

On simplifying, the magnitude for V_{OL} is obtained as 2 V.

(c) V_{IL} extraction

As per the status of the transistor mentioned in Table 2 on V_{IL} input, the KCL is applied at the output node, wherein the current at the steady state are set equal and expressed as:

$$K_n\left(\frac{W}{L}\right)\left(\left(V_{GS}-V_{TH}\right)V_{DS}\frac{V_{DS}^2}{2}\right)=K_n\left(\frac{W}{L}\right)\times\left(\frac{\left(V_{GS}-V_{TH}\right)^2}{2}\right) \tag{3}$$

Using the above conditions as input to TFT4, we have:

$$\left(V_{IN}-8\right)\times\left(V_{OUT}-10\right)-\frac{1}{2}\left(V_{OUT}-10\right)^2=\frac{1}{2}\left(V_{IN}+8-V_{OUT}\right)^2 \tag{4}$$

Differentiating with regard to 'V_{IN}' on both sides and using the condition from the VTC, we get:

$$\left(V_{IN}-8\right)\times\frac{dV_{OUT}}{dV_{IN}}+\left(V_{OUT}-10\right)-\frac{1}{2}\times2\times\left(V_{OUT}-10\right)\times\frac{dV_{OUT}}{dV_{IN}}=\frac{1}{2}\times2\times\left(\left(V_{IN}+8-V_{OUT}\right)\times\left(\frac{1}{2}-\frac{dV_{OUT}}{dV_{IN}}\right)\right) \tag{5}$$

Setting $V_{IN}=V_{IL}$ we get $V_{IL}=6.28$ V.

(d) V_{IH} extraction

Similar to V_{IL}, the V_{IH} is extracted as:

$$K_n\left(\frac{W}{L}\right)\left(\frac{\left(V_{GS}-V_{TH}\right)^2}{2}\right)=K_n\left(\frac{W}{L}\right)\left(\frac{\left(V_{GS}-V_{TH}\right)^2}{2}\right) \tag{6}$$

Putting the values we get:

$$K_n\left(\frac{800}{90}\right)\left(V_{IN}-8\right)^2=2 \tag{7}$$

Differentiating with regard to V_{IN} on both sides we get 8 V. The remaining two transistors, TFT 3 and TFT4, are acting as a buffer, so the current remains same in both the transistors at the steady state.

$$K_n\left(\frac{W}{L}\right)\left(\frac{\left(V_{GS}-V_{TH}\right)^2}{2}\right)=K_n\left(\frac{W}{L}\right)\left(\frac{\left(V_{GS}-V_{TH}\right)^2}{2}\right) \tag{8}$$

$$K_n\left(\frac{800}{90}\right)\times\frac{1}{2}\left(V_{IN}-8\right)^2=K_n\left(\frac{800}{90}\right)\times\frac{1}{2}\times\left(V_{O1}+2-V_{OUT}\right)^2 \tag{9}$$

Differentiating with regard to 'V_{IN}' on both sides we get:

$$V_{IN}-8=2-V_{OUT} \tag{10}$$

Setting $V_{IN}=V_{IH}$ and $V_{OUT}=V_{OL}=2.0$,V we get $V_{IH}=8.0$ V.

4.2 Analysis and extraction of dynamic parameters

From the VTC characteristics of the inverter, it can be seen that the transition occurs for quite a narrow duration. In this region, the inverter works as an amplifier, as the gain is higher than unity only for this region. The analytical value of the gain is calculated from the earlier measured values of the critical voltages and discussed as:

Width of the transition region $=(V_{IH}-V_{IL})$ V
Height of the transition region $=-(V_{OH}-V_{OL})$ V
Gain 'A_v' $=-(V_{IH}-V_{IL})/(V_{OH}-V_{OL})=-4.65$
$NM_H=(V_{OH}-V_{IH})=(9.91–8)=1.91$ V
$NM_L=(V_{IL}-V_{OL})\ V=(6.28-2.0)=4.28$ V

It can be observed from Table 3 that the performance of the dual gate pseudo-all-*p* inverter is higher than that of conventional inverter configurations.

The gain is augmented by factors of 2 and 1.5 compared to DLL and ZVLL configurations, respectively. Additionally, the output voltage swing is increased by 25% and 19% for the pseudo-inverter compared to DLL and ZVLL, respectively. This is due to the presence of the cascaded logic design of the pseudo-all-*p* inverter, which enhances the values of the V_{OH} and V_{OL}. The static performance parameters calculated through analytics are compared with the values extracted from simulation and are listed in Table 4. The parameters show a close match with reasonable validation.

Table 2 Static parameters and mode of transistor organic all-*p* pseudo-inverter circuit

Static parameter	Mode of operation	Transistors in 'ON' state
V_{OH}	Linear and cut-off	TFT1
V_{OL}	Cut-off and saturation	TFT2
V_{IL}	Linear and saturation	TFT1 and TFT2
V_{IH}	Both in saturation	TFT1 and TFT2

Table 3 Comparison of performance parameters of proposed dual gate pseudo-all-*p* inverter with conventional inverter

Parameters	Static performance parameter extraction			Dynamic performance parameter extraction			
	V_{OH} (V)	V_{OL} (V)	A_v	NM_L (V)	τ_{PHL} (µs)	τ_{PLH} (µs)	τ_p (µs)
Conventional organic inverter							
DLL (DG)	8.1	2.1	2.2	4.1	30	19	25
ZVLL (DG)	9.1	1.6	3.1	5.1	85	54	70
Proposed dual gate pseudo-all-*p* inverter	9.95	1.05	4.65	5.45	77	12	44

Table 4 Comparison between simulated and analytical values of static performance parameters

Parameter	Simulated value (V)	Analytical value (V)
V_{OH}	9.91	10
V_{IH}	8.5	8.2
V_{OL}	1.05	2
V_{IL}	6.50	6.28
NM_H	1.93	1.91
NM_L	5.45	4.28
A_v	−4.43	−4.65

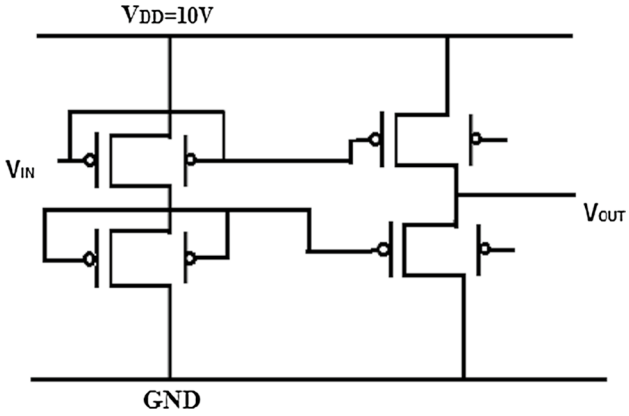


Fig. 10 Back gate biasing in DG-OTFT organic pseudo-all-*p* inverter

5 Performance improvement techniques of organic dual gate pseudo-all-*p* inverters

One of the performance limiting factors is the threshold voltage drop occurring in the circuits. To improve performance, this threshold voltage drop should be minimized. There are different techniques for improving the performance. In this work, two techniques are employed:back gate biasing and bootstrapping.

5.1 Back gate biasing in organic pseudo-all-*p* inverters without bootstrap technique

Reliable as well as robust operation of organic circuits is directly linked with the noise margin, which should be considerably high to ensure stable characteristics [13]. Threshold voltage can be limited by applying a fixed bias at the back gate, while a higher positive value of fixed gate bias is applied at the back gate of the driver [13]. Furthermore, front and back gate are coupled together for the load transistor to improve the noise margin and gain of the inverter.

By applying a fixed bias, the transistor operates in the linear region. Back gate V_{TH} control is used to limitat the performance degradation of the circuit and shift it towards $V_{DD}/2$, which is expected for the ideal behaviour of the inverter. The organic pseudo-inverter with back gate biasing is shown in Fig. 10; it has proven to be an efficient technique for enhancing the noise margin and robustness of the circuit.

5.2 Organic pseudo-all-*p* inverter with bootstrap technique

Bootstrapping is used to reduce the threshold voltage drop in the circuit [3, 12]. The circuit structure consists of a body capacitor (CB) that is used in the feedback loop between the output node and the input node. The charge stored on the capacitor is added to the voltage available at its input node, resulting in voltage boosting at its output terminal.

As a result, the threshold voltage drop is compensated and the noise margin is enhanced. The dual gate based pseudo-inverter with bootstrapping is shown in Fig. 11. All transistors are operating in sub-threshold operation using the bootstrap technique. The voltage transfer characteristic with bootstrapping is shown in Fig. 12. Bootstrapping is more efficient in reducing the leakage power and enhancing the speed of the logic circuits. It provides voltage levels above supply voltage and below the ground potential. This phenomenon leads the devices into the super cut-off region-which minimizes the leakage power dissipation.

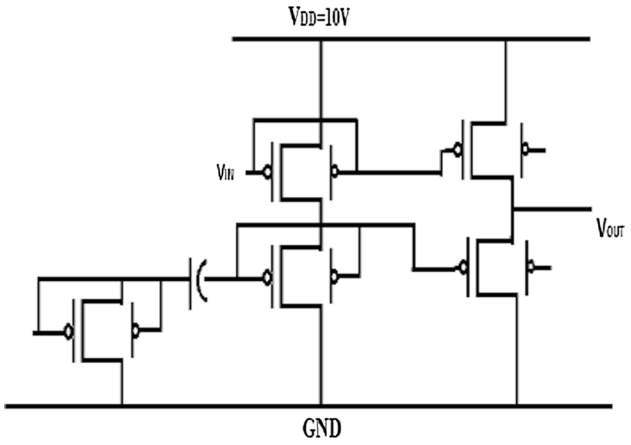


Fig. 11 Dual gate organic pseudo-all-*p* inverter with bootstrap technique

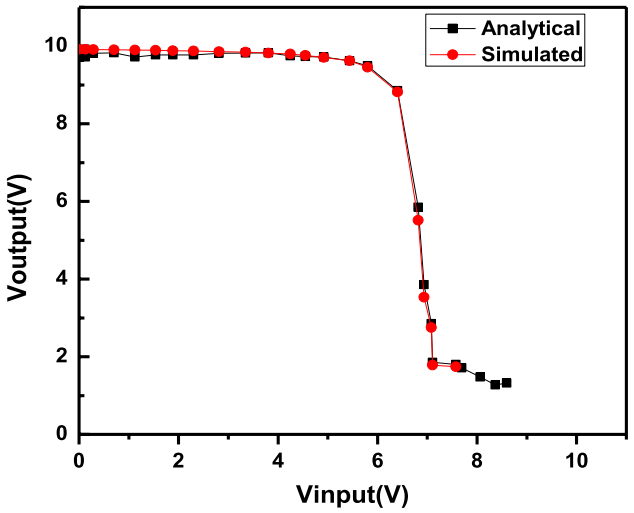


Fig. 12 VTC plot of analytical and simulated results for organic pseudo-inverter with bootstrap technique

Table 5 Comparison of extracted values of the dynamic performance parameters with and without bootstrap technique

Parameter	Extracted parameter without bootstrap technique	Extracted parameter with bootstrap technique
τ_{PHL} (μs)	77	79
τ_{PLH} (μs)	12	84
τ_{p} (μs)	44	82
P_S (μW)	22.2	31.1

Comparison of extracted values of the dynamic performance parameters with and without bootstrapping is summarized in Table 5. The output voltage level is held to a minimum value by its body transistor, which is boosted with a body capacitor. The main purpose of the body capacitor is to reduce threshold voltage (V_{TH}) sensitivity. The gain of the inverter using bootstrapping is improved by 48.47%. Voltage gain of a organic pseudo-all-*p* inverter with and without bootstrap techniques in ZVLL configuration is depicted in Fig. 13. A comparison between noise margin of organic pseudo-all-*p* inverter with and without bootstrap technique is given in Table 6, which shows that the higher value of noise margin (NM_H) improved by 25% with the bootstrap inverter.

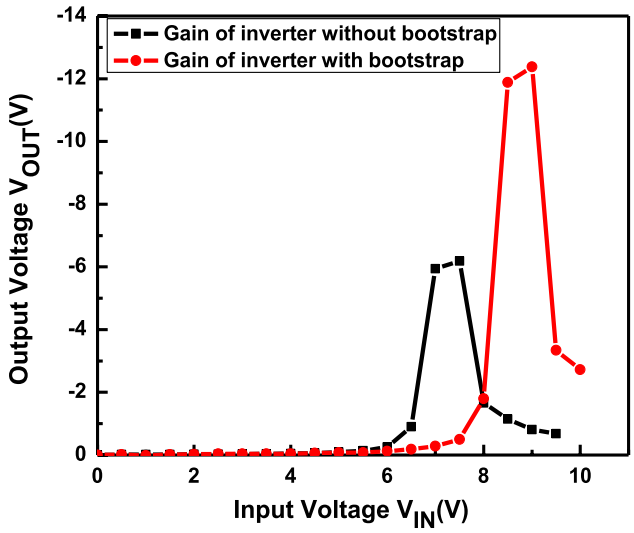


Fig. 13 Voltage gain of organic pseudo-all-*p* inverter with and without bootstrap techniques in ZVLL configuration

Table 6 Comparison of extracted parameters of DG-OTFT based pseudo-all-*p* inverter with and without bootstrap technique

Parameter	Without bootstrap (V)	With bootstrap (V)
V_{OH}	9.91	9.96
V_{TH}	7.35	7.21
V_{OL}	1.05	0.92
NM_H	1.93	1.96
NM_L	5.45	2.64

From the dynamic characteristics, it is observed that the propagation delay is increased in the organic pseudo-all-*p* inverter with bootstrap technique, as shown in Fig. 14, due to the weak pull down (*p*-type) transistor in the structure that may be increased using a high aspect ratio for the pull down transistor.

The power consumed should be estimated for the device, as it is an important parameter used to predict the reliability of the device for long term use. If the device consumes less power, it is more reliable for intended use. The power consumed by the inverter is discussed in Sect. 6.

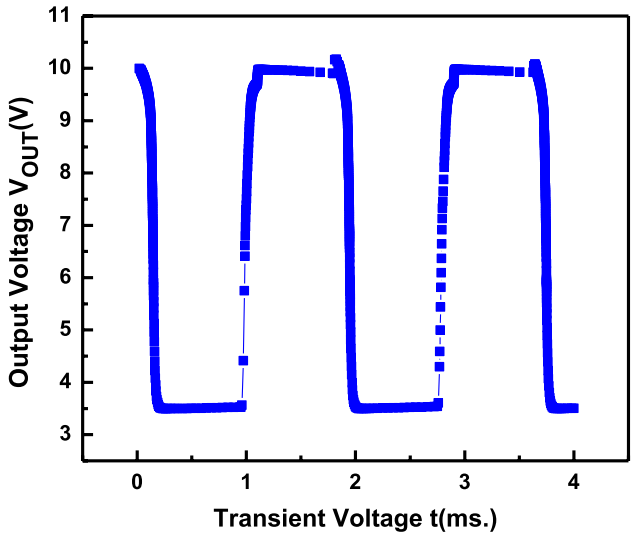


Fig. 14 Transient characteristics of organic pseudo-all-*p* inverter with bootstrap technique in ZVLL configuration

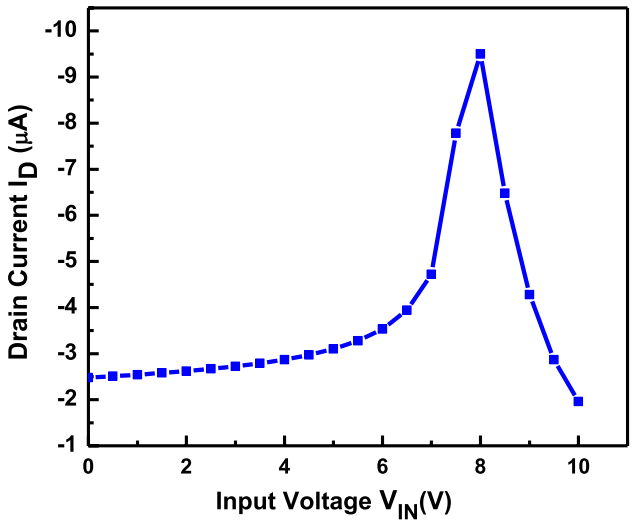


Fig. 15 Power supply current of organic pseudo-all-*p* inverter

6 Estimation of propagation delay, power consumption and area in dual gate organic pseudo-all-*p* inverter

The power consumed by a pseudo-all-*p* inverter is dependent on the amount of current flowing through the inverter. The current consumed I_S is the peak value of the current flowing through the driver transistor during the ON state. In MOSFET based devices there is considerable power consumption due to leakage, the current flowing through the inverter when it is in the OFF state.

To overcome this leakage power, the bootstrap technique is employed. Applying an input voltage above the supply but below the ground potential is said to be in the super cut-off region. As the device is operating in super-cut-off region the leakage power is minimised. The power consumed by the inverter is optimized when both the TFT3 and TFT4 are in the saturation region, as shown in Fig. 15, and compared with ZVLL and DLL based inverters in Table 7. The area consumed by the design is higher, as compared to conventional inverter, due to the presence of cascaded logic in the inverter, which results in a four transistor array.

The power consumed is given by Eqs. 11, 12 and 13.

$$P_S = I_S \times V_{DD} \tag{11}$$

$$P_L = I_L \times V_{DD} \tag{12}$$

$$\text{PDP} = P \times \tau_P \tag{13}$$

From the above equations it is observed that the power consumption of a dual gate based pseudo-all-*p* inverter is lower than that of conventional inverters. It is reduced by 84% for pseudo-all-*p* inverter without bootstrapping as compared to DLL conventional configuration, as shown in Table 7. Consequently, the power delay product (PDP) with bootstrapping is higher as compared to that without bootstrapping, due to the weak pull down transistor and higher propagation delay. Likewise, the static and transient performances of the dual gate design differ from those of the single gate design due to the difference in structure, transconductance and I_{ON}/I_{OFF} ratio.

There is a significant increase of 30% in the voltage gain of the organic pseudo-all-*p* inverter in comparison to the ZVLL (DG) of conventional inverter, as observed in Table 8. Similarly, the output swing is improved by 16.85% and 32.58% in comparison to the dual gate based ZVLL and DLL inverters, respectively.

Likewise, the propagation delay of dual gate pseudo-all-*p* inverter is reduced by 59.09% with regard to the ZVLL inverter, as depicted in Table 8. Similarly, there is improvement in the value of V_{OH} by 18.59% and 24.77% in the value of NM_H compared to conventional organic inverter, as shown in Table 9.

Table 7 Comparison of performance in terms of propagation delay power, area and PDP of pseudo-all-*p* inverter with conventional inverter

Parameter	τ_p (μs)	Power P_S (μW)	PDP = $P_S \times \tau_p$ (ps)	Area (μm^2)
DLL (DG)	25	139	3475	40,000
ZVLL (DG)	70	22	1540	40,000
DG-organic pseudo-all-*p* inverter	44	22.2	888	80,000

Table 8 Performance comparison of dual gate based circuit with single gate

Parameter	Static performance parameter extraction				Dynamic performance parameter extraction		
	V_{OH}	V_{OL}	A_v	Output swing	τ_p (µs)	τ_{PHL} (µs)	τ_{PLH} (µs)
DLL (DG)	8.1	2.1	2.2	6.0	25	54	21
DLL (SG)	7.9	2.9	1.9	5.0	93	154	32
ZVLL (DG)	9.0	1.6	3.1	7.4	70	146	38
ZVLL (SG)	9.0	2.1	2.6	6.9	159	277	40
Pseudo-all-p inverter, ZVLL (DG)	9.95	1.05	4.65	8.9	44	77	12

Table 9 Performance comparison of dual gate organic pseudo-all-*p* inverter with conventional inverter in terms of area and static parameters

Parameters	V_{OH} (V)	V_{OL} (V)	NM_L (V)	NM_H (V)	Area (µm^2)
Conventional organic inverter	8.1	2.1	5.1	4.1	40,000
Proposed dual gate pseudo-all-*p* inverter	9.95	1.05	1.93	5.45	80,000

7 Conclusion

This research paper has presented a dual gate based organic pseudo-all-*p* inverter with and without bootstrap technique. The static and transient characteristics of the inverter were observed and transition time was extracted. DG based transistors have shown better performance mainly in terms of mobility, drain current, trans-conductance, thereby improving the performance of entire circuits. These transistors have better control over the threshold voltage due to the possibility of biasing at both the gates. Significant increases of 16.85% and 32.58%, respectively, were obtained for the voltage swing of the proposed inverter over the conventional DLL and ZVLL configurations. Propagation delay and power consumption were greatly reduced as well. The extracted value of propagation delay of dual gate pseudo-all-*p* inverter was reduced by 59.09% with regard to the ZVLL inverter. Additionally, bootstrapping and back gate biasing techniques were applied. The higher value of noise margin (NM_H) was improved by 25% with the bootstrap inverter technique. Also, the gain of the inverter using bootstrapping was improved by 48.47%. Thus, the bootstrap inverter technique was proven as a significant stepping stone towards achieving high-performance flexible digital circuits.

References

1. Kumar, B., Kaushik, B.K., Negi, Y.S., Goswami, V.: Single and dual gate OTFTs based robust organic digital design. Microelectron. Reliab. **54**(1), 100–109 (2014)
2. Mittal, P., Negi, Y.S., Singh, R.K.: Mapping of performance limiting issues to analyze top and bottom contact organic thin film transistors. J. Comput. Electron. **14**(1), 360–379 (2015)
3. Kaushik, B.K., Kumar, B., Prajapati, S., Mittal, P.: Organic Thin-Film Transistor Applications: Materials to Circuits. CRC Press, Boca Raton (2016)
4. Kumar, B., Kaushik, B.K., Negi, Y.S.: Static and dynamic characteristics of dual gate organic TFT based NAND and NOR circuits. J. Comput. Electron. **13**(3), 1–12 (2014)
5. Chen, Y., Geng, D., Mativenga, M., Nam, H., Jang, J.: High-speed pseudo-CMOS circuits using bulk accumulation a-IGZO TFTs. IEEE Electron Device Lett. **36**(2), 153–155 (2014)
6. Zhao, Q., Liu, Y., Zhao, J., Guo, X., Li, H., Yang, H.: Noise margin modelling for zero-*VGS* load TFT circuits and yield estimation. IEEE Trans. Electron Devices **63**(2), 684–690 (2016)
7. Mittal, P., Negi, Y.S., Singh, R.K.: An analytical approach for parameter extraction in linear and saturation regions of top and bottom contact organic transistors. J. Comput. Electron. **14**(3), 828–843 (2015)
8. Atlas User's Manual: Device Simulation Software. Silvaco International, Santa Clara (2010)
9. Huang, T.C., et al.: Pseudo-CMOS: a design style for low-cost and robust flexible electronics. IEEE Trans. Electron Devices **58**(1), 141–150 (2011)
10. Zhao, Q., Sun, W., Zhao, J., Feng, L., Xu, X., Liu, W.: Noise margin, delay, and power model for pseudo-CMOS TFT logic circuits. IEEE Trans. Electron Devices **64**(6), 2635–2641 (2017)
11. Klauk, H., Zschieschang, U., Halik, M.: Low voltage organic thin film transistors with large transconductance. J. Appl. Phys. **102**(7), 0745141–074514-7 (2007)
12. Raval, H.N., Tiwari, S.P., Navan, R.R., Mhaisalkar, S.G., Rao, V.R.: Solution processed bootstrapped organic inverters based on P3HT with a high-k gate dielectric material. IEEE Electron Device Lett. **30**(5), 484–486 (2009)
13. Spijkman, M., Smits, E.C.P., Blom, P.W.M., Leeuw, D., Bon Salnt Come, Y., Setayesh, S., Cantatore, E.: Increasing the noise margin in organic circuits using dual gate field effect transistors. Appl. Phys. Lett. **92**(14), 143304 (1-3) (2008)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Lecture Notes on Multidisciplinary Industrial Engineering
Series Editor: J. Paulo Davim

R. Ganesh Narayanan
Shrikrishna N. Joshi
Uday Shanker Dixit *Editors*

Advances in Computational Methods in Manufacturing

Select Papers from ICCMM 2019

Springer

Lecture Notes on Multidisciplinary Industrial Engineering

Series Editor

J. Paulo Davim, Department of Mechanical Engineering, University of Aveiro, Aveiro, Portugal

"Lecture Notes on Multidisciplinary Industrial Engineering" publishes special volumes of conferences, workshops and symposia in interdisciplinary topics of interest. Disciplines such as materials science, nanosciences, sustainability science, management sciences, computational sciences, mechanical engineering, industrial engineering, manufacturing, mechatronics, electrical engineering, environmental and civil engineering, chemical engineering, systems engineering and biomedical engineering are covered. Selected and peer-reviewed papers from events in these fields can be considered for publication in this series.

More information about this series at http://www.springer.com/series/15734

R. Ganesh Narayanan · Shrikrishna N. Joshi ·
Uday Shanker Dixit
Editors

Advances in Computational Methods in Manufacturing

Select Papers from ICCMM 2019

Springer

Editors
R. Ganesh Narayanan
Department of Mechanical Engineering
Indian Institute of Technology Guwahati
Guwahati, Assam, India

Shrikrishna N. Joshi
Department of Mechanical Engineering
Indian Institute of Technology Guwahati
Guwahati, Assam, India

Uday Shanker Dixit
Department of Mechanical Engineering
Indian Institute of Technology Guwahati
Guwahati, Assam, India

ISSN 2522-5022 ISSN 2522-5030 (electronic)
Lecture Notes on Multidisciplinary Industrial Engineering
ISBN 978-981-32-9071-6 ISBN 978-981-32-9072-3 (eBook)
https://doi.org/10.1007/978-981-32-9072-3

© Springer Nature Singapore Pte Ltd. 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Conference Organizing Committee

Patron

Prof. Gautam Biswas, Director, IIT Guwahati

Chairman

Prof. S. K. Dwivedy, Department of Mechanical Engineering, IIT Guwahati

Executive Chairman

Prof. Uday Shanker Dixit, Department of Mechanical Engineering, IIT Guwahati

Organizing Secretary

Dr. R. Ganesh Narayanan, Department of Mechanical Engineering, IIT Guwahati

Joint Organizing Secretary

Dr. Shrikrishna N. Joshi, Department of Mechanical Engineering, IIT Guwahati

Treasurer

Dr. Swarup Bag, Department of Mechanical Engineering, IIT Guwahati

Members

Dr. P. S. Robi, IIT Guwahati
Dr. Pankaj Biswas, IIT Guwahati
Dr. S. Pal, IIT Guwahati
Dr. Manas Das, IIT Guwahati
Dr. Deepak Sharma, IIT Guwahati
Dr. A. K. Das, IIT Guwahati
Dr. Vinayak Kulkarni, IIT Guwahati
Dr. S. S. Gautam, IIT Guwahati
Dr. Prasenjit Khanikar, IIT Guwahati
Dr. S. Kanagaraj, IIT Guwahati
Dr. Nelson Muthu, IIT Guwahati
Dr. S. Senthilvelan, IIT Guwahati
Dr. Anil Birru, NIT Manipur
Dr. Sajan Kapil, IIT Guwahati

International Advisory Board

Amitabha Ghosh, IIT Kanpur, India
Amitava De, IIT Bombay, India
D. Banabic, Technical University of Cluj-Napoca, Romania
T. A. Dean, University of Birmingham, UK
J. S. Gunasekera, Ohio University, USA
P. Hodgson, Deakin University, Australia
P. K. Jain, IIT Roorkee, India
V. K. Jain, IIT Kanpur, India
S. S. Joshi, IIT Bombay, India
S. V. Kailas, IISc Bangalore, India
Kornel F. Ehmann, Northwestern University, USA
S. M. Mahajan, BHEL, India
S. S. Pande, IIT Bombay, India
J. Paulo Davim, University of Aveiro, Portugal
D. T. Pham, Cardiff University, UK
M. Rahman, NUS, Singapore
N. Ramesh Babu, IIT Madras, India

K. Ramesh Kumar, DRDO, India
A. Roy, Loughborough University, UK
K. P. Rao, City University of Hong Kong, Hong Kong
Santosh Kumar, IIT BHU, India
H. S. Shan, Chandigarh, India
M. S. Shanmugam, IIT Madras, India
Sumitesh Das, Tata Steel, India
Surjya Kanta Pal, IIT Kharagpur, India
S. Venugopal, NIT Nagaland, India
L. Vijayaraghavan, IIT Madras, India

Preface

Manufacturing is the backbone of the economy of a nation. It is one of the oldest activities of the human civilization. In the twenty-first century, the emphasis is on enhancing the efficacy and efficiency of manufacturing processes and systems by using the computational power. Industry 4.0 aims to make use of digital manufacturing, cloud computing and manufacturing, Internet of things, and big data analytics. It is essential to understand the manufacturing processes scientifically, develop models, and optimize them. In December 2011, the 1st International Conference on Computational Methods in Manufacturing (ICCMM 2011) was organized at the Indian Institute of Technology Guwahati. The conference was a huge success with encouraging participation from industry and academia. After a gap of 8 years, the 2nd International Conference on Computational Methods in Manufacturing (ICCMM 2019) was organized on March 8–9, 2019.

The objective of this conference was to promote the use of computational techniques in manufacturing by providing a platform for sharing ideas and expertise. We wanted to bring practicing engineers, experimentalists, and computational experts together. We received contributed papers from the researchers covering all major areas of computational manufacturing and allied fields. A total of 125 full papers were received, which were peer-reviewed. The authors have revised their papers based on the review comments. Finally, 90 papers were selected for inclusion in the book of proceedings.

This book of proceedings covers the following broad topics: material forming, joining and welding, casting, additive manufacturing, material fabrication, subtractive manufacturing, process modeling, simulation, analysis, optimization, and design. Although the focus of the conference is on computational techniques, most of the authors have invariably included experimental results as experiments are the inseparable part of modeling of manufacturing.

We hope that the book of proceedings will be a valuable document for the delegates, scientists, engineers, and research scholars. Springer has taken the responsibility of publishing and distributing it to a wider audience. We wish to

thank Ms. Swati Meherishi, Ms. Avni, and other staff members of Springer. We acknowledge the support provided by the reviewers, members of advisory and organizing committees, and colleagues at IIT Guwahati. Any feedback to the editors regarding this collection will be highly appreciated.

Guwahati, India

R. Ganesh Narayanan
Shrikrishna N. Joshi
Uday Shanker Dixit

Contents

Part I Material Forming

1 Numerical Simulation of Micro Forming of Bio-Absorbable AZ80 Magnesium Alloy 3
D. Rajenthirakumar, N. Srinivasan and R. Sridhar

2 Investigation on Tensile Testing of Tailor Welded Blanks (TWBs)—A Simulation Study 13
Shalin Marathe and Harit Raval

3 Formability Evaluation of EN-10149-2 (S700mc) Steel Under In-Plane Plane Stretching Condition 25
Dereje H/Georgis, Perumalla Janaki Ramulu and Habtamu Beri

4 Stretch Formability Prediction of AA6023-T6 Alloy Sheet Under Two Different Heating Conditions 37
Manaye Mathewos Handiso, Perumalla Janaki Ramulu and G. Somasundaram

5 FE Analysis of Single Point Incremental Forming (SPIF) for Tailor Welded Blanks (TWBs) 49
Shalin Marathe and Harit Raval

6 Joining Concentric Tubes by End Forming: A Finite Element Simulation Study 65
E. Premananda and R. Ganesh Narayanan

7 Optimization of the Forming Parameters in U-Bending for Punch Force and Springback Using Taguchi Method 75
P. Venkateshwar Reddy, D. Mohana Krishnudu, U. Pranavi and P. Janaki Ramulu

8 Experimental Analysis on Springback of JIS 3302 Grade Steel Sheet Material Under Various Heat-Treated Conditions 85
Kemal Ramato, Perumalla Janaki Ramulu and N. R. R. Anbusagar

9 **Experimental and Numerical Assessment of Residual Stresses in AA6061 After Surface Treatment by Deep Cold Rolling** 95
Rakesh Kumar, Gulshan Kumar, Vijay Gautam and Subhajit Konar

10 **Analysis of Hot Workability of Inconel Alloys Using Processing Maps** 109
Gauri Mahalle, Nitin Kotkunde, Amit Kumar Gupta and Swadesh Kumar Singh

Part II Welding and Joining Technologies

11 **Investigation on Metallographic Analysis of Electron Beam Ti6Al4V Alloy Welds** 121
Sohini Chowdhury, Yadaiah Nirsanametla and Manapuram Muralidhar

12 **Multi-objective Optimization of FSW Process Variables of Aluminium Matrix Composites Using Taguchi-Based Grey Relational Analysis** 133
Subramanya R. B. Prabhu, Arun Kumar Shettigar, Mervin A. Herbert and Shrikantha S. Rao

13 **Tool Durability and Weldability in Hybrid Friction Stir Welding of High-Strength Materials** 145
Amit Chaudhary, Chiranthan Ramesh, Viswanath Chinthapenta and Murshid Imam

14 **Swarm-Intelligence-Based Computation for Parametric Optimization of Electron Beam Fabrication** 153
Sanjib Jaypuria, Amit Kumar Das and Dilip Kumar Pratihar

15 **Effect of Pulsation on Temperature Distribution of Laser-Welded Dissimilar Joint of Thin Sheet** 165
Bikash Kumar and Swarup Bag

16 **Activated TIG Welding of AISI 321 Austenitic Stainless Steel for Predicting Parametric Influences on Weld Strength of Tensile Test—Experimental and Finite Element Method Approach** 179
S. Mohan Kumar, N. Siva Shanmugam and K. Sankaranarayanasamy

17 **Design Optimization of Welding Parameters for Multi-response Optimization in TIG Welding Using RSM-Based Grey Relational Analysis** 193
S. C. Moi, R. Rudrapati, A. Bandyopadhyay and P. K. Pal

18 **Subjective Factors Consideration in the Selection of Welding Technique for Welded Tube Manufacturing** 205
Ravindra S. Saluja and Varinder Singh

19 GMAW Investigation of AISI 201 Stainless Steel and Industry Need Optimization Using Genetic Algorithm 215
Krity Lata, M. Chandrasekaran, S. K. Tamang, R. Ramesh and N. K. Rana

20 Finite Element Modelling and Experimental Verification of Dissimilar Joining Between Inconel 718 and SS 316L by Micro-plasma Arc Welding 231
Ajit Kumar Sahu and Swarup Bag

21 Numerical Simulation and Experimental Validation of A-TIG Welding of 2.25Cr–1Mo Steel 245
A. R. Pavan, B. Arivazhagan, S. Arun Kumar, M. Vasudevan and S. Mahadevan

22 Finite Element Modeling of Hybrid Laser-TIG Welding of Type 316L(N) Stainless Steel 259
M. Ragavendran, M. Vasudevan and M. Menaka

23 Numerical Study on Electromagnetic Crimping of Aluminum Tube and Steel Profiled Rod 271
Getu Tilahun Areda and Sachin Dnyandeo Kore

24 Temperature Prediction During Self-pierce Riveting of Sheets by FEA-ANN Hybrid Model 283
Deepak Mylavarapu, R. Ganesh Narayanan and Manas Das

Part III Subtractive Manufacturing

25 Determination of Optimal Cutting and Tool Geometry Parameters for Better Surface Integrity of Hard Turned AISI 52100 Steel-Hybrid GRA-PCA 297
P. Umamaheswarrao, D. Ranga Raju, K. N. S. Suman and B. Ravi Sankar

26 Study of Cutting Edge Radius Effect on the Cutting Forces and Temperature During Machining of Ti6Al4V 309
Siddharam Mane, Shyamprasad Karagadde and Suhas S. Joshi

27 Comparative Study on Effect of Inclusions and Machinability Aspects of Alloy Steels 317
V. Krishnaraj, K. Thillairajan and R. Rajeshshyam

28 Process Parameters Optimization of Electrical Discharge Machining of Al7075/SiC/WS2 by Using MCDM 329
Rakesh Kumar Patel and M. K. Pradhan

29 CNC Turning and Simulation of Residual Stress Measurement on H13 Tool Steel 337
Ravi Butola, Qasim Murtaza and Ranganath M. Singari

30 Small-Scale Machining Simulations 349
Anuj Bisht, Anish Roy, Uday S. Dixit, Satyam Suwas and Vadim V. Silberschmidt

31 Process Parameters Optimization of Cylindrical Grinding Machining Using Taguchi Method 363
T. Nancharaiah and M. Subramanyam

32 Effect of Sequential Drilling Process on In-situ Bone Temperature 371
Varatharajan Prasannavenkadesan and Ponnusamy Pandithevan

Part IV Additive Manufacturing and Casting

33 FE-Based Heat Transfer Analysis of Laser Additive Manufacturing on Ti–6Al–4V Alloy 381
Sapam Ningthemba Singh, Sohini Chowdhury, Md. S. Mujaheed Khan, Manapuram Muralidhar and Yadaiah Nirsanametla

34 Material Perspective and Deformation Pattern of Micro-Sized Metallic Particle Using Cold Gas Dynamic Spray 393
Abdul Faheem, Faisal Hasan and Qasim Murtaza

35 Towards an Improved Understanding of Stereolithography Process—A Computational Study 403
Usharani Rath and Pulak M. Pandey

36 Experimental Investigations for Effects of Raster Orientation and Infill Design on Mechanical Properties in Additive Manufacturing by Fused Deposition Modelling 415
Pushpendra Yadav, Ankit Sahai and Rahul Swarup Sharma

37 Development of Methods for Machine Intelligence Quotient Measurement for Piston Casting by Die Casting Machines: A Study 425
Rupshree Ozah and Muralidhar Manapuram

38 Optimization of Porosity in Cast Aluminium Foams 439
R. Kumar, G. Avinash and Neha Gupta

39 To Predict Surface Roughness and Linear Shrinkage of Die Casting Process by Using of Fuzzy Algorithm Model 447
Narendra Krishania and Anil Kumar Birru

40 Prediction of Hot Spot and Hot Tear of the Al–Cu Cast Alloy by Casting Simulation Software 459
Kumara Swamy Pulisheru and Anil Kumar Birru

41 Influence of Pouring Temperature on Fluidity of Al–Cu Cast Alloy: Analyses Using Casting Simulation Software 469
Kumara Swamy Pulisheru and Anil Kumar Birru

Part V Modelling and Analysis

42 Analysis and Optimization of Tool Wear Rate in Magnetic Field-Assisted Powder-Mixed Electrical Discharge Machining of Al6061 Alloy Using TLBO 485
Arun Kumar Rouniyar and Pragya Shandilya

43 ANFIS-Based Subtractive Clustering Algorithm for Prediction of Response Parameters in WEDM of EN-31 497
Kunal Chopra, Aishwarya Payla, Guneet Kaur and Eswara Krishna Mussada

44 On Estimation of Scallop Height from CNC Part Programs 509
A. Kukreja and S. S. Pande

45 Influence of Guides on Critical Speeds of Circular Saws 519
S. Singhania, P. Kumar, S. K. Gupta and M. Law

46 Investigations on the Influence of Radial Run-Out on Cutting Forces for Serrated Cutters 531
P. Bari, M. Law and P. Wahi

47 Modeling and Optimization of WEDM Parameters Using Design of Experiments 543
P. Bharathi, G. Srinivasarao and P. G. Krishnaiah

48 Analysis on Thermal Characteristics of Micro-Drilled Glass Using Microwave Energy at 2.45 GHz 557
Gaurav Kumar and Apurbba Kumar Sharma

49 A Study on the Effect of Polarity Change on Various Parameters on Ti6Al4V in Powder-Mixed Micro-EDM Using Multi-objective Grey Fuzzy Optimization 569
Deepak Sharma, Anusha Roohi Siddique, Vishnu Kumar, Shalini Mohanty and Alok Kumar Das

50 Multi-output Response Optimization for Overall Enhancement of Mechanical Characteristics Using Utility Approach for AISI 316L Austenitic Stainless Steel Using Microplasma Arc Welding 587
Srikant Prasad, Dibakor Boruah and Pintu Thakur

51 Characterization of Geometrical Features of Ultra-Short Pulse Laser-drilled Microholes Using Computed Tomography 603
K. Kiran Kumar, G. L. Samuel and M. S. Shunmugam

52 Investigation on Rotary Electrical Discharge Machining (EDM) of EN-25 Material Using Copper Electrode for Improving Geometrical Errors 615
L. Selvarajan, R. Rajavel, F. Leo Princely, R. Aravind and T. P. Habeeb Masood

53 Computation of End-Cutting-Edge Wear of Single-Point Cutting Tool Using Image Processing 629
Vishal K. Singh and Shrikrishna N. Joshi

54 Modelling and Analysis of Multi-agent Approach for an IoT-Enabled Autonomous Manufacturing System 643
K. B. Badri Narayanan and M. Sreekumar

Part VI Material Fabrication

55 Estimation of Distortion for CFRP Antenna Reflector During Autoclave Processing 657
Shrey Patel, Dhaval Shah, Shashikant Joshi and Kaushik Patel

56 Comparative Study of Hydroxyapatite Nanocomposites Reinforced with Zirconia and Titania Produced by Using HEBM 669
Vemulapalli Ajay Kumar, Penmetsa Rama Murty Raju, Nallu Ramanaiah and Siriyala Rajesh

57 Parametric Optimization of Corrosion Resistance of Electroless Ni–Co–P Coating 685
Subhasish Sarkar, Rishav Kumar Baranwal, Ishita Koley, Rupam Mandal, Tapendu Mandal and Gautam Majumdar

58 Investigation on Dimensional Accuracy for CFRP Antenna Reflectors Using Autoclave and VARTM Processes 693
Tushar Gajjar, Dhaval Shah, Shashikant Joshi and Kaushik Patel

59 On Finite Element Analysis of Melting of Metallic Powders Using Microwave Energy at 2.45 GHz 703
Mohit Kumar, Gaurav Kumar and Vikas Kukshal

60 Tribological Behaviour of Nylon 6 Composites and Optimization of Process Parameters Using Hybrid Taguchi-GRA and CSA 719
Sudhir Kumar, Barnik Saha Roy and K. Panneerselvam

61 Fabrication of Hexagonal Bar from Aluminum Alloy AA6063 Scrap by Frictional Stir Back Extrusion on Milling Machine 733
Gizaw Yohannes, Habtamu Beri and Perumalla Janaki Ramulu

62 Fabrication, Characterization and Evaluation of Mechanical Properties of Aluminium Hybrid Matrix Composite (Al6063/SiC-Bagasse Fly-Ash) 743
Getachew Gashaw, Perumalla Janaki Ramulu and Ch. Venkatesh

63 Fuzzy-EDAS (Evaluation Based on Distance from Average Solution) for Material Selection Problems 755
Divya Zindani, Saikat Ranjan Maity and Sumit Bhowmik

64 Tensile and Flexural Property Analyses of Glass/Carbon Embedded Hybrid Natural Fibre Composites 773
Jobin G. Jose, E. Sivakumar, Jibin T. Philip, Kore Mahesh and Basil Kuriachen

65 Mechanical Characterization of Hybrid Carbon–Glass-Reinforced Polypropylene Composites 783
Ansuman Padhi and Debiprasada Sahoo

66 Effect of Temperature on the Fracture Strength of Perfect and Defective MonoLayered Graphene 793
Kritesh Kumar Gupta and Sudip Dey

67 Effect of Process Parameters of Plasma Arc Preheating on Low Carbon Steel 805
Abhishek Kumar and Pankaj Biswas

Part VII Process Simulation and Optimization

68 Reduced-Order Finite Element Analysis of Rough Surface Contact Using Reduced Integration Elements 819
Ahad Noor, Saipraneeth Gouravaraju and Sachin Singh Gautam

69 FEM-Based Modeling of Drilling-Induced Delamination in Laminated FRP Composites 829
Vishwas Divse, Deepak Marla and Suhas S. Joshi

70 Fuzzy Inference System-Based Neuro-Fuzzy Modeling of Electron-Beam Welding 839
Sanjib Jaypuria and Dilip Kumar Pratihar

71 Numerical Analysis of Erosion of Rough Surface 851
Ahad Noor, Saipraneeth Gouravaraju and Sachin Singh Gautam

72 Numerical Simulation and Experimental Validation of Nanoindentation of Silicon Using Finite Element Method 861
Borad M. Barkachary and Shrikrishna N. Joshi

73 Effect of Microchannel Pattern on Fluid Flow and Heat Transfer Characteristics 877
Dungali Sreehari and Apurbba Kumar Sharma

74 Comparative Analysis of GPU-Based Solver Libraries for a Sparse Linear System of Equations 889
Utpal Kiran, Subhajit Sanfui, Shashi Kant Ratnakar, Sachin Singh Gautam and Deepak Sharma

75 Hybridizing Evolutionary Multi-objective Algorithm Using Random Mutations and Local Searches 899
Riddhiman Saikia and Deepak Sharma

76 Local Search Heuristics-Based Genetic Algorithm for Capacitated Vehicle Routing Problem 909
Deepak Sharma, Sukhomay Pal, Aayush Sahay, Puneet Kumar, Gaurav Agarwal and Katravath Vignesh

Part VIII Design

77 Performance Evaluation of Adhesive Spur Gear with the Influence of Properties of Adhesive 923
R. Shailesh, M. Ramu, M. Govindaraju, K. Karthikeyan and V. Satheeshkumar

78 Design Analysis and Modification of Sugarcane Fibrizer Hammer at Wonji Shoa Sugar Factory (WSSF) Ethiopia 933
Tesfaye Mathewos, Perumalla Janaki Ramulu and Esmael Adem Esleman

79 FEA Validation of Experimental Results of First Ply Failure of Composite Structure 947
Ishan Sonawane and A. Muthuraja

80 Triply Periodic Minimal Surface-Based Porous Scaffold Design and Analysis Subjected to Hard Tissue Reconstruction 955
Verma Rati, Nishant Singh, Sanjay Rai and Shekhar Kumta

81 Design of Medical Device Product Using Multiphysics Simulations 967
Rajesh Kapuganti

82 Finite Element Analysis of a Disc Brake Mounted on the Axle of a Train 973
E. Madhusudhan Raju, Somayajula Kavya Geetika and Devireddy Jitesh Krishna

83 Modelling and Analysis of Composite Conical Space Adapter 985
Rajesh Addanki, Pranay Narinder Kumar, Ayush Mahajan and Manoj Kumar Buragohain

84 Parametric Study of Axisymmetric Circular Disk Subjected to Rotational Autofrettage 997
S. M. Kamal and R. Kulsum

85 Influence of Varying Fully Constrained Circular Boundary Condition Area on Damage Mechanism of GFRP Laminate Under Low Velocity Impact Loading 1011
Mahesh and K. K. Singh

86 Process Capability Index for Geometrically Distributed Quality Characteristics 1023
N. K. Singh

87 Thermomechanical Analyses of Single Sided Single Pass Submerged Arc Welding of AISI 304 Austenitic Stainless Steel 1031
P. V. S. S. Sridhar, Vishnu Nair, Pankaj Biswas and Pinakeswar Mahanta

88 Stochastic Longitudinal Vibration of Single Walled Carbon Nanorods—A Non Local Elasticity Approach 1045
Tanmoy Loha and Sudip Dey

89 Finite Element Analysis and Three-Point Bend Testing of Indian Rice Grains with Fissure Effects 1059
P. V. Gharat, D. K. Shinde and D. N. Raut

90 Finite Element Analysis of Potential Liner Failures During Operation in Spherical Pressure Vessel 1073
R. Pramod, N. Siva Shanmugam, C. K. Krishnadasan and K. Sankaranarayanasamy

Index 1089

About the Editors

R. Ganesh Narayanan is an Associate Professor at the Department of Mechanical Engineering, Indian Institute of Technology (IIT) Guwahati, India. He received his Ph.D. from the IIT Bombay, India. His research areas include metal forming and joining. He has contributed many research articles in reputed journals and international conferences. He has edited few books including 'Sustainable Material Forming and Joining' published by CRC Press USA, 'Strengthening and Joining by Plastic Deformation' published by Springer Singapore, 'Advances in Material Forming and Joining' published by Springer India, and 'Metal Forming Technology and Process Modeling' published by McGraw Hill Education, India. He has also edited special issues of journals including 'Advances in Computational Methods in Manufacturing' in the International Journal of Mechatronics and Manufacturing Systems, and 'Numerical Simulations in Manufacturing' in the Journal of Machining and Forming Technologies. He has organized three international conferences at IIT Guwahati namely International Conference on Computational Methods in Manufacturing (ICCMM) in 2011, the 5th International and 26th All India Manufacturing Technology, Design and Research (AIMTDR) Conference in 2014, and International Conference on Computational Methods in Manufacturing (ICCMM) in 2019. He has also organized a GIAN course on 'Green Material Forming and Joining' at IIT Guwahati in 2016.

Shrikrishna N. Joshi is an Associate Professor in the Department of Mechanical Engineering, Indian Institute of Technology Guwahati, India, where he joined after completing his doctoral studies from IIT Bombay in 2010. His research interests include laser based manufacturing, ultraprecision machining, computer aided design and manufacturing (CAD/CAM) and, manufacturing process modeling & optimization. He has authored/co-authored about 50 research articles in international journals/conferences and six book chapters. He has also co-edited 2 books on "Laser based manufacturing" with Springer.

Uday Shanker Dixit is a Professor in the Department of Mechanical Engineering, Indian Institute of Technology Guwahati. He completed his BE in Mechanical Engineering from the University of Roorkee (now the Indian Institute of Technology Roorkee) in 1987, and his MTech and PhD in Mechanical Engineering from the Indian Institute of Technology Kanpur (IIT Kanpur) in 1993 and 1998, respectively. Professor Dixit has previously worked on designing machines for industry, and as the Officiating Director of the Central Institute of Technology, Kokrajhar. Professor Dixit has been actively engaged in research in various areas of design and manufacturing for the last 25 years. He has authored/co-authored more than 200 research papers, 27 book chapters and 7 books on mechanical engineering. He has also co-edited 7 books on manufacturing. Presently, he is an associate editor of the Journal of the Institution of Engineers (India), Series C and Vice-President of the AIMTDR conference.

Part I
Material Forming

Chapter 1
Numerical Simulation of Micro Forming of Bio-Absorbable AZ80 Magnesium Alloy

D. Rajenthirakumar, N. Srinivasan and R. Sridhar

Abstract The characteristics of grains during the micro forming process are unpredictable. It reflects in the material behavior, interfacial friction and surface finish due to size effects. This paper describes the numerical analysis of size effects and frictional effects during the plastic deformation of magnesium AZ80 alloy in the micro extrusion process. On the basis of true stress–strain curve, the material model is used for numerical investigation. The results reveal that the flow stress increases with decreasing grain size due to increasing interfacial friction. The results are then validated with the micro extrusion experiments conducted. Further, the friction generation and load prediction can be achieved through numerical simulation is the primary criteria.

Keywords Micro forming · Extrusion · Finite element method · Size effect

1.1 Introduction

Miniaturization technology has transformed the very risky assignment into a smart task not only in electronic field but also in medical devices such as pacemakers, implantable defibrillators, catheters, stents, and neuromodulator device. The metals and glasses are replacing the polymers due to its higher strength in minimum grain size. During the transformation from macro scale to micro scale the phenomena called size effect governs the micro plastic deformation of material and its behavior [1]. When scaling down the characteristics of grains are unpredictable pertaining to material behavior. During the micro forming process, surface to volume ratio increases considerably and grain size effect ruins the flow stress of the material [2]. Hence, different methods of experiments and numerical analysis are conducted by

D. Rajenthirakumar (✉) · N. Srinivasan · R. Sridhar
Department of Mechanical Engineering, PSG College of Technology,
Coimbatore 641004, India
e-mail: rajentkumar@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_1

researchers worldwide. The primary focus of the research is to control the size effect and minimize the interfacial friction.

Piwnik and Mogielnicki simulated the influence of container roughness on material flow process in micro extrusion. They concluded the significant impact of a wave roughness on material flow in the process [3]. The authors also simulated the interfacial friction towards the improvement of surface roughness [4]. The strategies to estimate grain properties and model grain size effect are developed and the grain size effect on deformation behavior and the scatter of flow stress are modeled. The results are then validated with experiments [5]. So to eliminate the time taken and expenses for experimenting, numerical simulation is employed to predict the behavior in several aspects during the experiments.

In this work, a numerical simulation of AZ80 Mg alloy is performed and the results obtained are then validated with physical experiments. The AZ80 Mg has a huge application in biomedical field [6]. The paper mainly focuses on the numerical analysis of micro extrusion by which the prediction of size effects and friction at interfaces has been achieved. Initially, the material model is obtained by micro compressions test, which is given as input for numerical simulation. The second stage is the FEM simulation. The last stage is the validation of numerical analysis with experiments.

1.2 Micro Compression Test and Size Effects

The material for the micro compression experiment is AZ80 Mg alloy of different grain sizes such as 20, 36, and 84 µm in quasi-static condition. The strain rate is 0.1 mm/s which is kept constant. Figure 1.1 shows the true stress–strain curves for

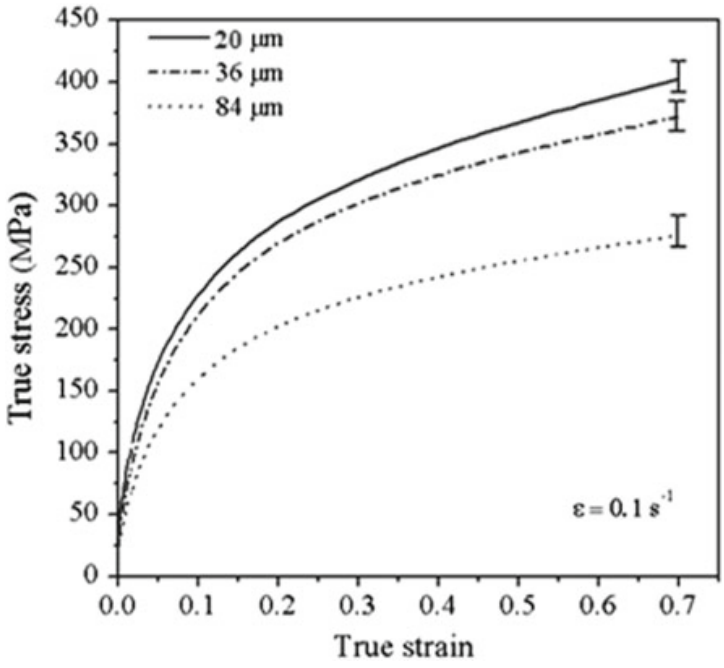


Fig. 1.1 True stress–strain curve

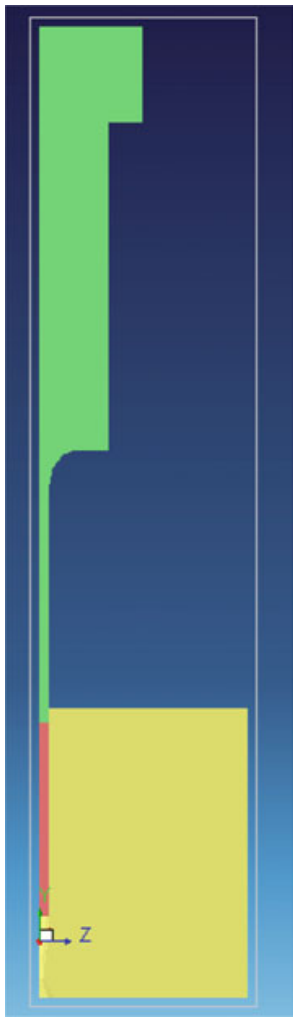


Fig. 1.2 FEM—2D axisymmetric Model

specimens in 1.2 mm diameter and three different grain sizes. The curves illustrate that the flow stress increases with decrease in grain size.

The effects of scaling down with decreasing flow stress are discussed through so-called surface model [7]. The resistance to deformation gets varied from surface layer to inner grains of the material. The difference is due to the hardness variation in surface layer to inner grains of the material. The occurrence of interfacial friction effect can be seen in flow stress curves. The existence of larger grains count possesses maximum resistance to plastic deformation. As a result, the smaller the grains, the larger the friction generated in tool-billet interfaces can be noticeable.

1.3 FEM Model

AFDEX is a metal forming simulator, meets the requirements for bulk metal-forming and simulation. The AFDEX_V17R00 is utilized for predicting the nonlinear behavior of material in 2D and 3D. Based on the previous results in literature [8], 0.3 µm is selected as coefficient of friction as it is a dry condition, as

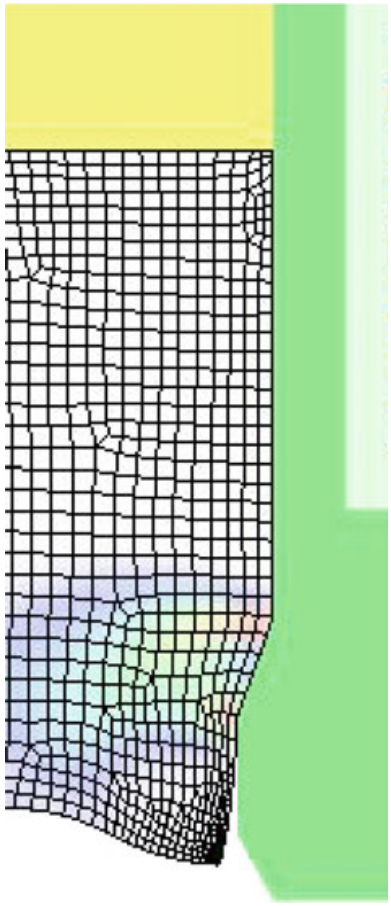


Fig. 1.3 Meshed model

an assumption to compare the frictional effect on forming load. The micro extrusion tooling system is modeled as rigid bodies while the forming material (micro pin) is an elastoplastic body. Simulations are performed with the most widely used von Mises yield criterion as it gives equal weight to all principal stresses. Considering the symmetry in both geometry and loading conditions, a single half experimental setup is modeled (see Fig. 1.2). The meshing in Fig. 1.3 is done with 8200 four nodes 2D axisymmetric quadrilateral element. These elements have maintained the accuracy in stress analysis. The adaptive meshing technique is selected to retain the mesh quality. To obtain a suitable mesh, a local element size has to be applied. This requires some trial and error method.

1.4 Physical Experiment

In order to show the effects of grain size and investigate the interfacial friction, an experimental setup is designed and micro extrusion tests with AZ80 magnesium billets of different grain sizes have been performed. Figure 1.4 shows the exploded view of micro extrusion die assembly and Fig. 1.5 shows the schematic view of micro extrusion die assembly.

Fig. 1.4 Micro extrusion—die assembly

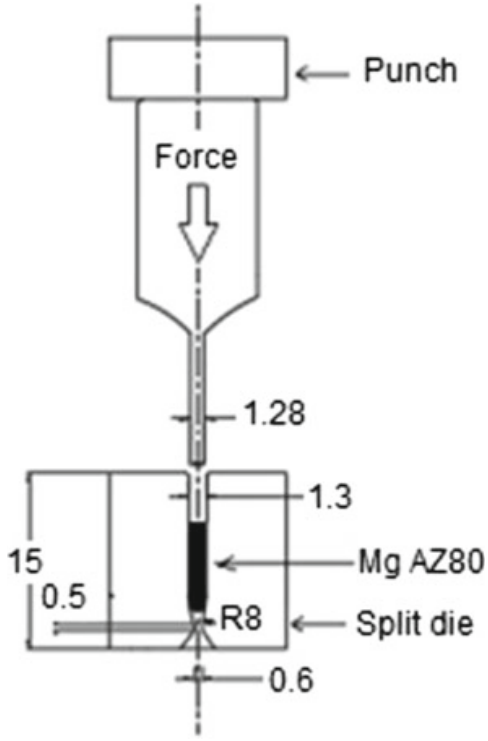


Fig. 1.5 Schematic view of microextrusion assembly

1.5 Results and Discussion

The numerical simulation of micro forward extrusion is performed to evaluate the interfacial friction. The 20 μm grain size is selected for simulations it possesses high grains count per unit volume. The Von-Mises stress distribution and strain distribution of the micro forward extrusion are shown in Figs. 1.6 and 1.7, respectively. The material undergoes the severe shear deformation at the die-billet contact surfaces. The differences in friction coefficient are observed in stress and strain distribution. The interfacial friction observed during the simulation (see Fig. 1.8) is reduced with coefficient of friction.

The experimental results are then compared with numerical simulation for validation purpose. The micro extruded parts are shown in Fig. 1.9. The billet-die interfaces are exposed to friction which increases with the decrease of specimen size. This is due to the increase of the fraction of contact area and so, the peak is observed initially in forming load. It started decreasing due to increase in proper

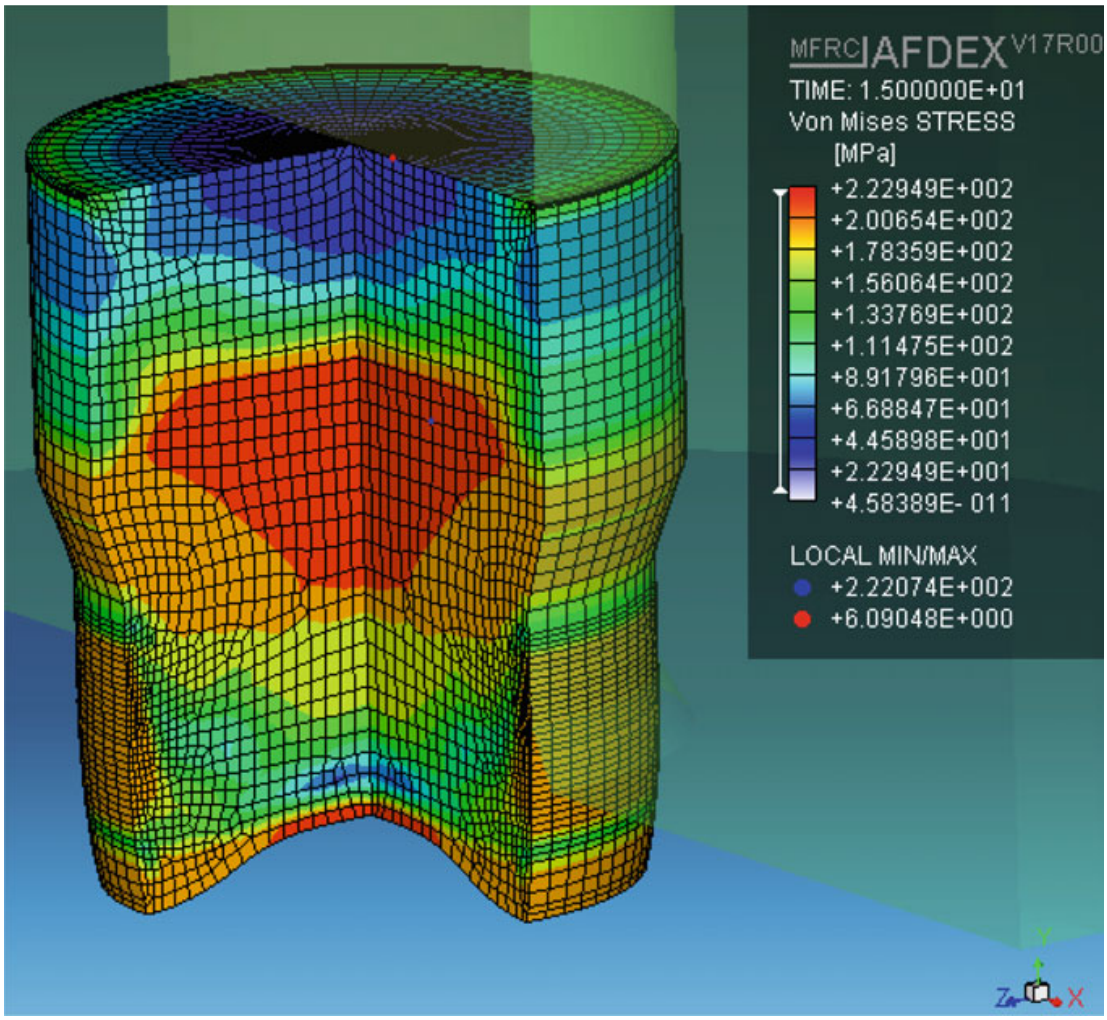


Fig. 1.6 Stress distribution (20 μm)

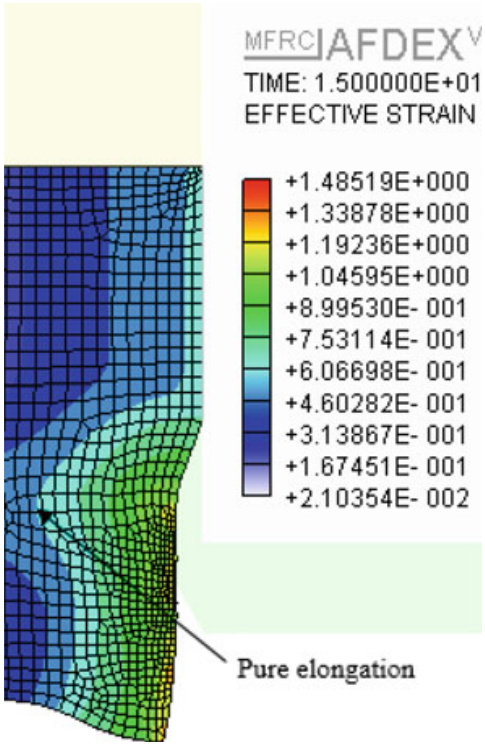


Fig. 1.7 Strain distribution (20 μm)

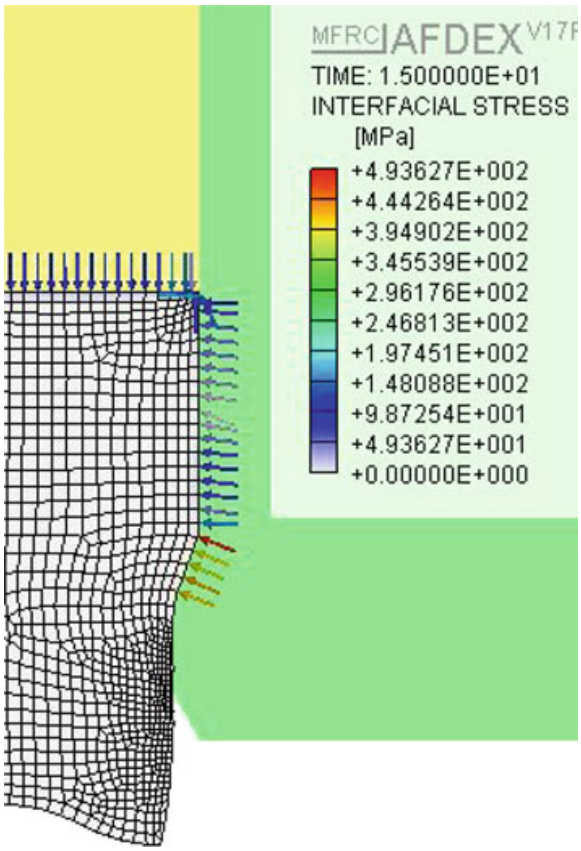


Fig. 1.8 Interfacial stress (20 μm)

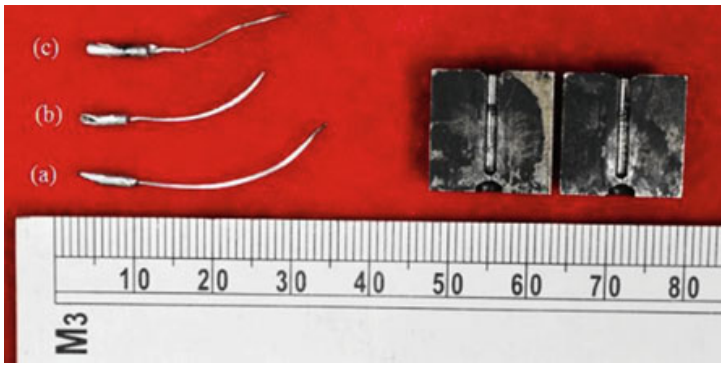


Fig. 1.9 Micro extruded parts, **a** 20 µm, **b** 36 µm and **c** 84 µm

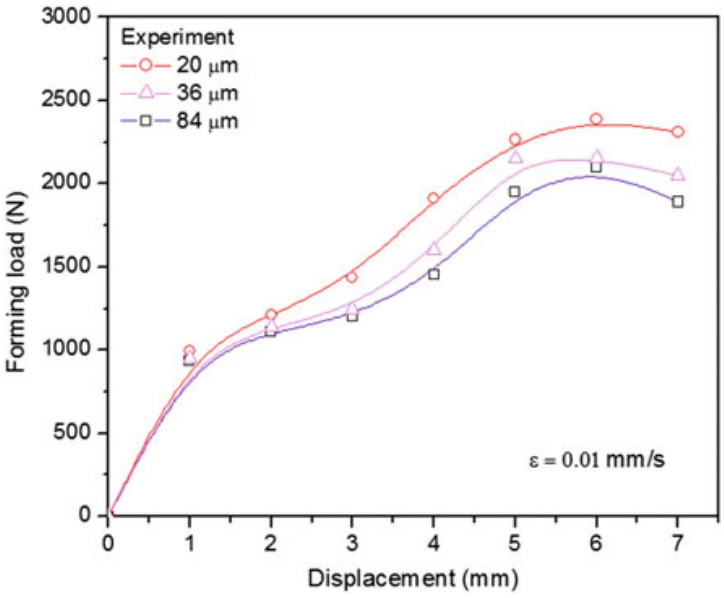


Fig. 1.10 Load versus displacement curve of AZ80 Mg

material plastic flow. The numerical simulation of the micro forward extrusion performed with 0.3 µm. The load–displacement response obtained from experiments is shown in Fig. 1.10. The maximum load consumed by 20 µm is due to high grains count per unit volume which possesses high strain resistance to deformation.

1.6 Conclusions

The numerical simulation and experimental validation has been conducted for biological grade Mg alloy AZ80 and the following conclusions are made:

- The interfacial friction increases with decrease of grain size. It is proved by observing high interfacial stress of 20 µm than the deformation of other grain size Mg AZ80 alloy.
- The flow stress of AZ80 Mg alloy increases with reduction in grain size.
- The deformation load increases with reduction in grain size.

- The numerical simulation results with AFDEX shows the in-depth friction condition effects in the micro forward extrusion process. When the friction coefficient increases the deformation load increases correspondingly.

The numerical simulation has concealed significant benefits such as load prediction for several friction conditions, existence of interfacial friction and minimizing the time and number of experiments. Even though the numerical analysis shows some deviation with experiments, it is likely to be considered for prediction.

Acknowledgements The authors gratefully acknowledge the financial support of Science and Engineering Research Board, Government of India and PSG College of Technology, Coimbatore, Tamil Nadu, India.

References

1. Engel, U., Eckstein, R.: Microforming—from basic research to its realization. J. Mater. Process. Technol. **126**, 35–44 (2002)
2. Qiu, T.T., Hou, Y.K., Cao, H.L.: Review of research progress on size effect in micro-forming. Key Eng. Mater. **717**, 118–121 (2016)
3. Piwnik, J., Mogielnicki, K.: Experimental and FE analysis of aluminium alloy plastic flow in the forward microextrusion processes. Arch. Metall. Mater. **59**, 521–525 (2014)
4. Piwnik, J., Mogielnicki, K.: The friction influence on stress in micro extrusion. Arch. Foundry Eng. **10**, 451–454 (2010)
5. Chan, W.L., Fu, M.W., Lu, J., Liu, J.G.: Modeling of grain size effect on micro deformation behavior in micro-forming of pure copper. Mater. Sci. Eng. A **527**, 6638–6648 (2010)
6. Zheng, Y.: Magnesium alloys as degradable biomaterials. CRC Press, Boca Raton (2015)
7. Chan, W.L., Fu, M.W., Yang, B.: Study of size effect in micro-extrusion process of pure copper. Mater. Des. **32**, 3772–3782 (2011)
8. Krishnan, N., Cao, J., Dohda, K.: Study of the size effects on friction conditions in microextrusion—part I: microextrusion experiments and analysis. J. Manuf. Sci. Eng. **129**, 669 (2007)

Chapter 2
Investigation on Tensile Testing of Tailor Welded Blanks (TWBs)—A Simulation Study

Shalin Marathe and Harit Raval

Abstract The goal of better fuel economy can be fulfilled by the use of technology like TWBs (Tailor Welded Blanks). TWB is a product of joining two or more than two materials before any forming operation. The limitations associated with TWBs are the weld line movement and reduced formability. The formability reduction of TWB is due to the thin/weak material involved in it and therefore the tensile study of the TWBs becomes an important aspect. In the present study, investigation of tensile testing of Similar Material Dissimilar Gauge TWB and Dissimilar Material Similar Gauge TWB has been done numerically. For the simulation purpose, ABAQUS/Explicit has been used. The results indicate that the thin/weak material governs the strength of the TWB because the failure or necking is observed in that particular side only. Also, it has been observed that the failure locations are found to be at the boundary of thin/weak base material and weld zone.

Keywords Tailor welded blanks (TWBs) · Friction stir welding (FSW) · ABAQUS/explicit · Simulation · Tensile testing

2.1 Introduction

Nowadays, automobile, aerospace, transportation, and shipbuilding industries are aiming at weight reduction of the components without compromising the strength and Tailor Welded Blanks (TWBs) have been found to fulfill these requirements. TWB is a product of joining two or more than two material prior to any forming operation. TWB involves joining of two or more than two materials having different properties, thicknesses and/or coatings by any of the conventional joining processes. As it involves joining of materials having different densities the overall weight of the product decreases [1]. Also, TWBs are subjected to forming after the

S. Marathe (✉) · H. Raval
Department of Mechanical Engineering, Sardar Vallabhbhai National Institute of Technology, Surat 395007, India
e-mail: shalin.marathe89@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_2

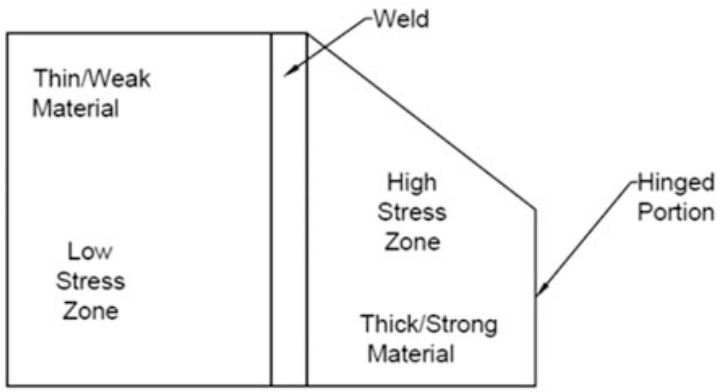


Fig. 2.1 Schematic representation of TWB used at inner door panel of an automobile body

joining has been done and because of that the number of dies required for the forming decreases [1] which also leads to decrease the product lead time and cost as well. Figure 2.1 indicates the schematic of TWB which is used at the inner door panel of an automobile body.

Despite of such tremendous advantages, the application of TWBs has been not up to the mark as limitations like formability reduction [2] and weld line movement [3] are associated with TWBs. In TWBs, as there is a presence of weld between the two parent metals, the whole product becomes inhomogeneous and when TWBs are subjected to forming operations these inhomogeneities result in uneven strain [2] and hardness distribution [2, 4, 5] across the weld. This uneven distribution of hardness and strain leads to failure of the TWBs resulting in the low formability of the component. A lot of research has been done to investigate the formability behavior of TWBs during the conventional forming [2] operations and it has been observed that the formability of TWBs is found to be decreasing in comparison to the parent metal participating in it. Now in order to improve the formability of the TWBs lots of efforts have been done by researchers.

TWBs are generally developed from the materials like steel [6–8], aluminum and its alloys [4, 9] and titanium as more application is found in automobile, ship-building, aerospace and transportation industries. The application of steel to develop TWBs has been reduced as the aluminum material gives a higher amount of weight saving [10]. TWBs involving steel as parent material are generally developed using the laser welding [6, 9, 11–13] processes and if aluminum and its alloys are used as parent material than solid-state welding processes like FSW [9, 14–16] is found more suitable. As aluminum has low laser absorption capacity and it is very reflective in nature, it is quite difficult to weld using laser welding and other conventional welding techniques.

In order to decide the quality of the TWBs, there are many testing methods available and tensile testing is one of them. So, in this present work tensile testing of TWBs has been performed using a simulation tool ABAQUS/Explicit. The simulation methods save important resource like time, material and cost. So in present work, testing of TWBs made from similar material having different thickness/gauge and dissimilar materials having similar thickness/gauge are investigated using a simulation tool. In this work, TWB is considered to be made

from FSW process which is known as Friction Stir Tailor Welded Blank (FSTWB). Results like plastic strain, Von-misses stresses, Load versus Displacement diagram, Stress–Strain behavior and True strain distribution across the weld are extracted from the simulation and compared with published literatures.

2.2 Numerical Conditions

In this present study, all the simulations are carried out using ABAQUS/Explicit [17]. This study has been divided into two sections. First section contains the simulation study on similar material dissimilar gauge TWBs made from FSW process and second section contains the same investigation for dissimilar material similar gauge TWBs. For all the simulations "Weld Zone" approach is adopted to define the weld properties in the tensile test specimen.

2.2.1 Section-I (Similar Material Dissimilar Gauge TWBs)

In this section, as a Base Material AA 5754 of 1.06 and 1.66 mm is used. The different material properties used during the simulation for the mentioned material is given in Table 2.1.

The material has been modeled as 3 dimensional deformable. The dimensions of the tensile test specimen and the cross-section of the welded joint are shown in Fig. 2.2a, b, respectively. It was assumed that the base metal and weld metal were following the power-law equation.

Figure 2.3a indicates the cross-section (isometric view) and Fig. 2.3b indicates dissimilar gauge tensile test specimen (isometric view) modeled in the ABAQUS environment. The entire simulation was completed in total two steps namely Initial and step-1. In the initial step of the simulation one end of the specimen was encastre

Table 2.1 Properties considered for simulation [18]

Property	Base material	Weld metal
Density (ρ) (kg/m^3)	2600[a]	2600[a]
Young's modulus G (N/m^2)	71×10^9	68×10^9
Yield stress (σ_y) (N/m^2)	132×10^6	146×10^6
Poisson's ratio (μ)	0.33[a]	0.33[a]
Strength coefficient (K) (N/m^2)	551.5×10^6	453.5×10^6
Strain hardening index (n)	0.33	0.23

[a]Assumed

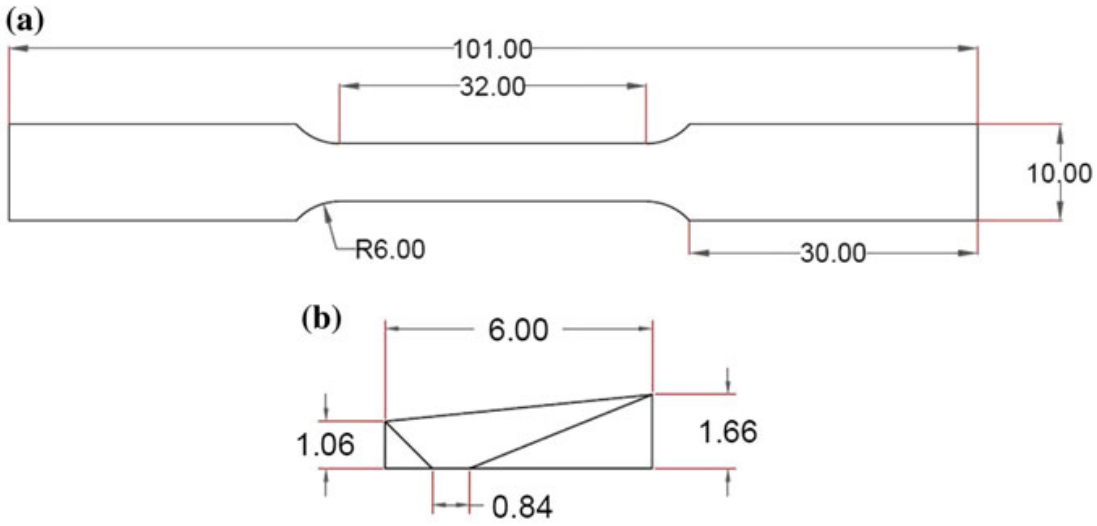


Fig. 2.2 Details of the tensile test specimen, **a** dimensions of the specimen, **b** cross-section of the welded joint (all the dimensions are in mm) [18]

(All DOF (Degree of Freedom) constrained) and in step-1 another end was given the displacement in the X direction only (transverse to the weld).

As shown in Fig. 2.4 thin base metal was given encastre boundary condition and thick base metal was given displacement in the X-direction. Meshing was done using linear tetrahedral elements. Table 2.2 indicates the meshing details used during the simulations.

2.2.2 Section-II (Similar Gauge Dissimilar Material TWBs)

In this section, tensile testing of TWB made from dissimilar material having similar gauge is performed using the ABAQUS/Explicit tool. Figure 2.5 indicates the tensile specimen modeled in the ABAQUS environment. The tensile testing was done as per the ASTM-E8 standards using a sub-sized tensile test specimen. For this section, the material properties were used as per Table 2.3.

In this investigation, one end of specimen was given encastre boundary conditions and another was given the displacement. The meshing details used for the simulation purpose are tabulated in Table 2.4. Figure 2.6 indicates the meshed model of tensile test specimen in the ABAQUS environment. Similar to case-I, this simulation was also completed using total 2 number of steps and materials were assumed to follow the Power law equation.

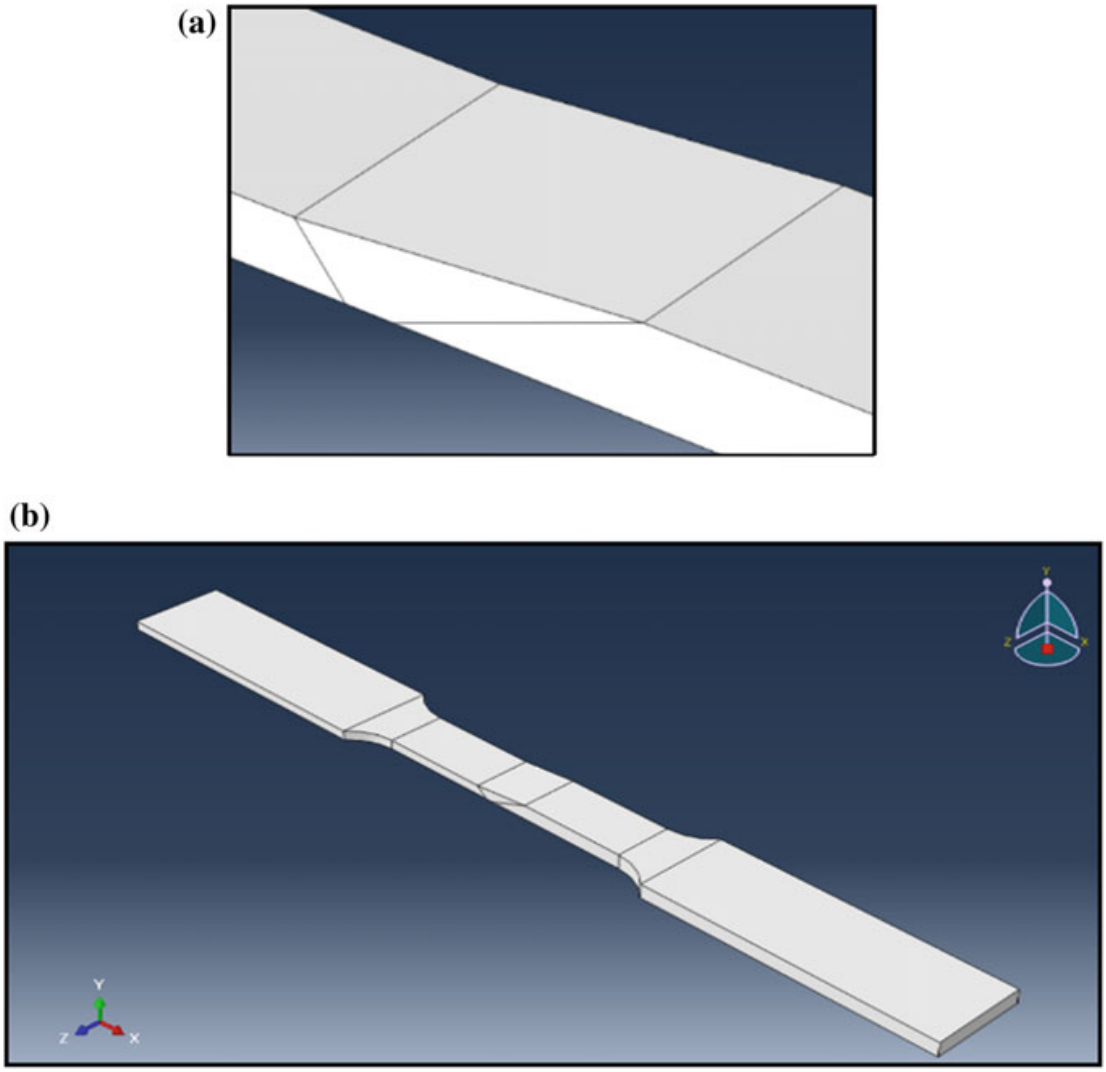


Fig. 2.3 Details of the tensile test specimen, **a** cross-section, **b** overall tensile test specimen

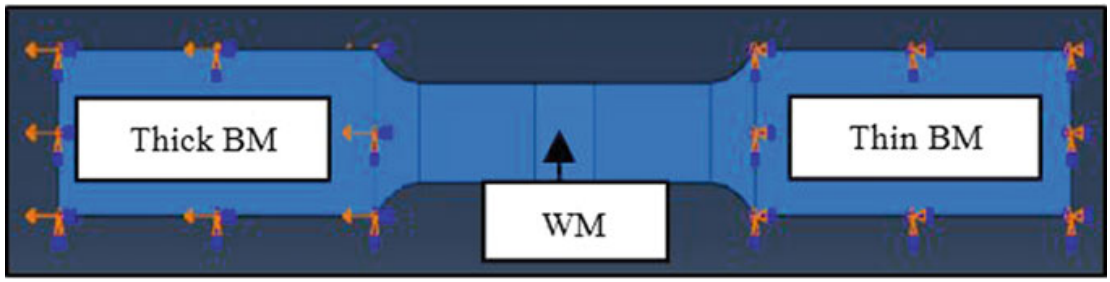


Fig. 2.4 Boundary Conditions applied to the tensile specimen (WM = Weld Metal, BM = Base Metal)

Table 2.2 Mesh statistics

Total number of nodes	3613
Total number of elements	13,355
Linear tetrahedral elements of type	C3D4

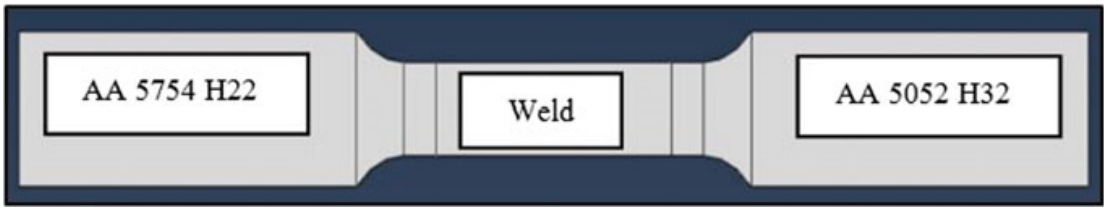


Fig. 2.5 TWB tensile specimen modeled in ABAQUS

Table 2.3 Properties used during simulation [2]

Property	AA 5754 H22	AA 5052 H32	Weld[a]
Density[a] (ρ) (kg/m^3)	2670	2680	2670
Young's modulus[a] G (N/m^2)	70.3×10^9	70.3×10^9	70.3×10^9
Yield stress (σ_y) (N/m^2)	122×10^6	140×10^6	131×10^6
Poisson's ratio[a] (μ)	0.33	0.33	0.33
Strength coefficient (K) (N/m^2)	252.04×10^6	258.7×10^6	255.37×10^6
Strain hardening index (n)	0.112	0.125	0.1185

[a]Assumed

Table 2.4 Mesh statistics

Total number of nodes	3390
Total number of elements	2016
Linear tetrahedral elements of type	C3D8R

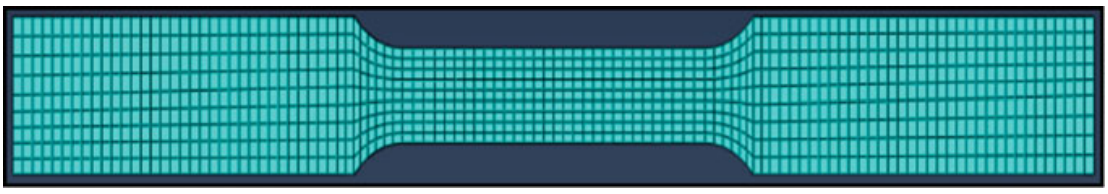


Fig. 2.6 Meshed model of the specimen

2.3 Results and Discussions

In the upcoming sub-sections, the results extracted from the simulation study have been discussed and compared with the same of available literatures. Responses like stress–strain curve, load versus displacement curve, failure location in the specimen and strain developed in the tensile test specimen are compared.

2.3.1 Section-I (Similar Material Dissimilar Gauge TWBs)

For the given boundary conditions, load versus displacement graph was extracted from the simulation and plotted as shown in Fig. 2.7 which also compares it with the published literature.

It can be observed that the natures of both the mentioned results are in good agreement with each other which indirectly authenticates the methodology adopted for the simulation purpose. From the simulations, results of plastic strain were also extracted as indicated in Fig. 2.8.

The maximum plastic strain was observed in the thin material as indicated in Fig. 2.8 and it was found at the boundary region of the thin base material and the weld zone. The value of maximum plastic strain predicted by the present study of simulation was 0.2519 and it was 0.2004 by the published study [18] which seems to be in very good agreement with each other.

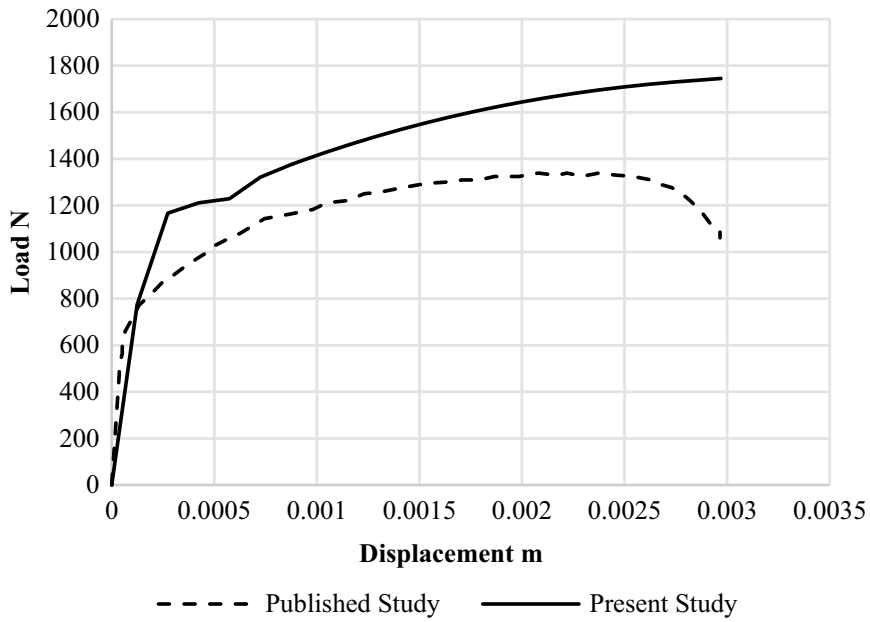


Fig. 2.7 Comparison of load versus displacement graph between published [18] and present study

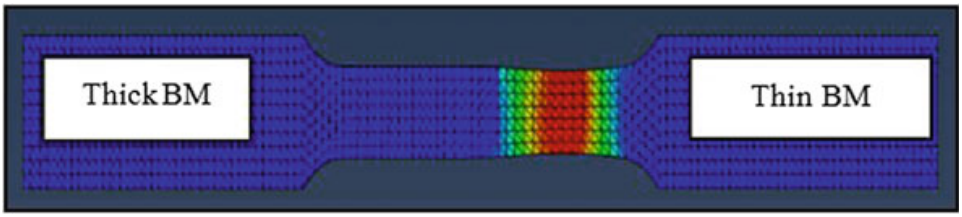


Fig. 2.8 Plastic strain (in X direction only) PE11 for present study

2.3.2 Section-II (Dissimilar Material Similar Gauge TWBs)

From the given boundary conditions in the section-II, the results like failure/necking locations, strain distributions across the weld and stress–strain behavior of TWB were extracted and compared with the published literature. Figure 2.9 indicates the present simulation result for necking location in the tensile test specimen. The failure location from weld center was observed at 11 mm in experimental study of published literature [2] and that was observed at 13 mm for the present simulation study. So, it can be stated that the results were in reasonably good agreement with each other.

The engineering stress–strain behavior of TWB was also plotted and compared with the published results as indicated in Fig. 2.10. From Fig. 2.10, it can be seen that the yield point predicted by present study and published study are in good agreement but for the plastic deformation part of the graph, a slight mismatch is found. This disagreement can be reduced by providing more appropriate boundary conditions and material properties during the simulations.

One more result of strain distribution across the weld was extracted from the simulation and compared with the published results as indicated in Fig. 2.11. The maximum strain value is observed at 13 mm from weld center for the present simulation study and that is the failure/necking point in the tensile test specimen of TWB. There was a disagreement between the present and published study regarding strain distribution which can be resolved using better boundary conditions and material properties.

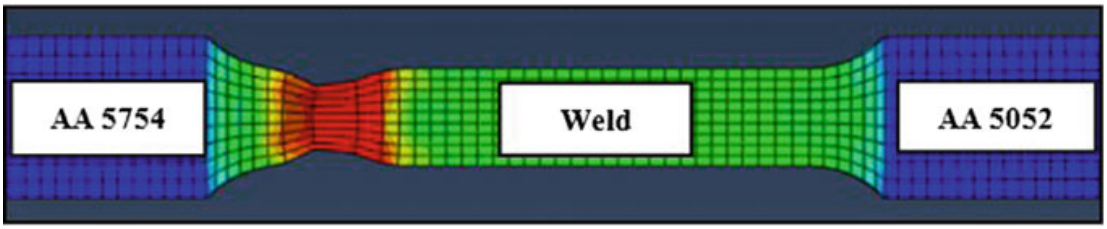


Fig. 2.9 Thinning/necking location for present simulation study (at 13 mm from weld centre)

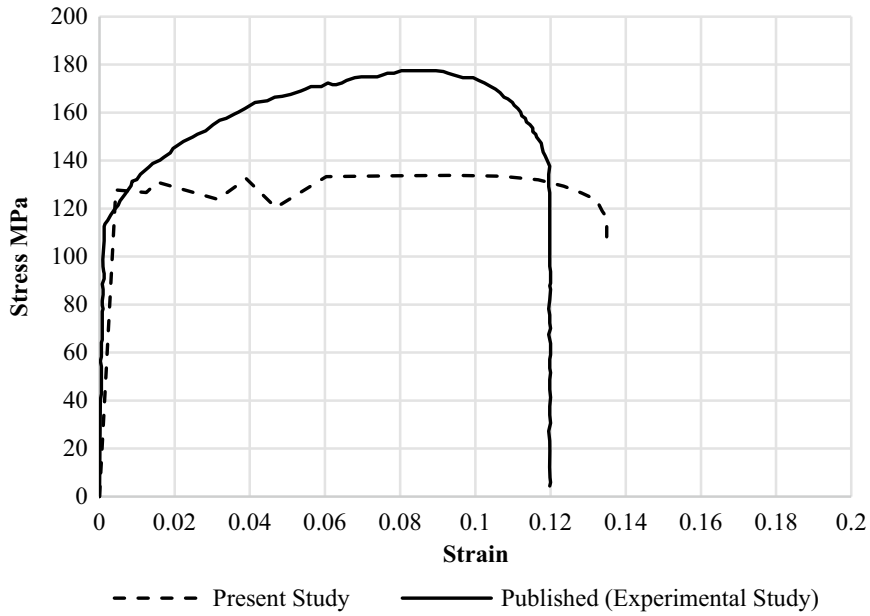


Fig. 2.10 Comparison of engineering stress–strain curve between present (simulation) and published [2] experimental results

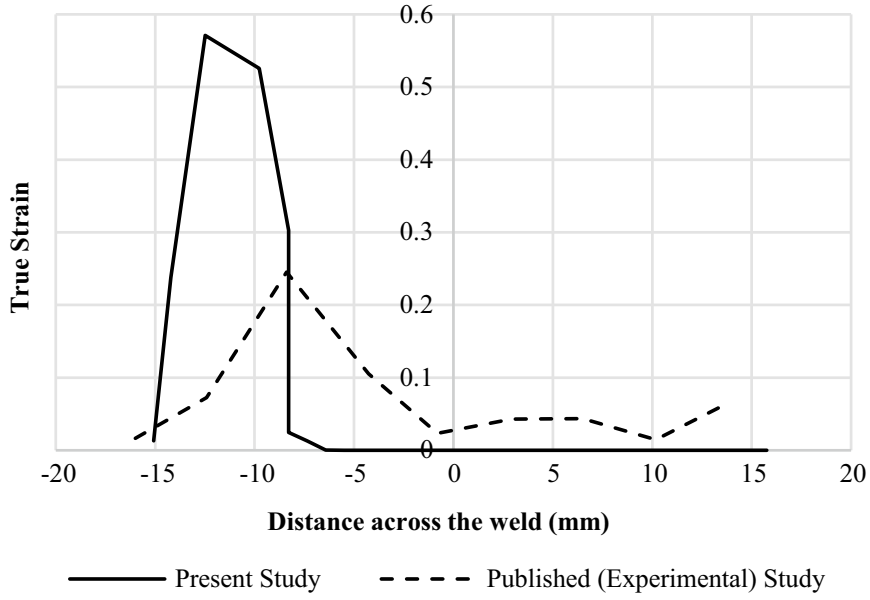


Fig. 2.11 Comparison of strain distribution across the weld for published [2] and present study

2.4 Conclusions

In the present research work, study of tensile behavior of the TWBs through FE analysis has been carried out. Following points can be concluded from the reported study:

- The "Weld Zone" approach used to define the weld properties during the simulation gives reasonably good results.
- The strength of the TWBs depends upon the weak/thin base material participating in it.
- To improve the strength and quality of the TWBs, strain distribution across the weld should be as even as possible.
- The boundary region between the weld zone and the base material is found to be the weakest.
- The simulation methodology not only saves the time, cost and resources but it also leads to reasonably good results for material behavior. The only limitation associated with the simulation methodology is that it demands for the accurate and proper input data in terms of boundary conditions and material properties.
- It is not necessarily required to perform the experimental/physical testing for accurate results. As represented in results of section-II, accurate results can be obtained to predict the material behavior during any deformation process by adopting appropriate approach for the simulation which is "Weld Zone" in this case.

References

1. Parente, M., Safdarian, R., Santos, A., Loureiro, A., Vilaca, P., Jorge, R.: A study on the formability of aluminum tailor welded blanks produced by friction stir welding. Int. J. Adv. Manuf. Technol. **83**, 2129 (2016)
2. Kesharwani, R., Panda, S., Pal, S.: Experimental investigations on formability of aluminum tailor friction stir welded blanks in deep drawing process. J. Mater. Eng. Perform. **24**(2), 1038 (2015)
3. Panda, S.K., Kumar, D.R.: Improvement in formability of tailor welded blanks by application of counter pressure in biaxial stretch forming. J. Mater. Process. Technol. **204**(1–3), 70–79 (2008)
4. Zadpoor, A.A., Sinke, J., Benedictus, R.: The effects of friction stir welding on the mechanical properties and microstructure of 7000 series aluminium tailor-welded blanks. Int. J. Mater. Form. **1**, 1311–1314 (2008)
5. Marathe, S.P., Mistry, H.J., Raval, H.K.: Parametric study of Friction Stir Welding (FSW) of AA 6061 using Taguchi method. In: Proceedings of National Workshop on Industrial Problems on Machines and Mechanical Similar Materials: "Challenges in Manufacturing". NIT, Nagpur (2016)
6. Song, Y., Hua, L., Chu, D., Lan, J.: Characterization of the inhomogeneous constitutive properties of laser welding beams by the micro-Vickers hardness test and the rule of mixture. Mater. Des. **37**, 19–27 (2012)
7. Song, Y.L., Hua, L.: Formability of tailor welded blanks based on the improved rule of mixture considering the welding residual strain and stress. Appl. Mech. Mater. Trans Tech Publ. **152**, 1406–1412 (2012)
8. Jia, Q., Guo, W., Li, W., Zhu, Y., Peng, P., Zou, G.: Microstructure and tensile behavior of fiber laser-welded blanks of DP600 and DP980 steels. J. Mater. Process. Technol. **236**, 73–83 (2016)

9. Ebrahimzadeh, P., Baseri, H., Mirnia, M.J.: Formability of aluminum 5083 friction stir welded blank in two-point incremental forming process. Proc. Inst. Mech. Eng. Part E J. Process Mech. Eng. **232**, 267 (2017)
10. Shibata, K., Iwase, T., Sakamoto, H., Kasukawa, M., Chiba, K., Saeki, H.: Welding of aluminium tailored blanks by Nd: YAG lasers. Weld. Int. **17**(4), 282–286 (2003)
11. Chatterjee, S., Saha, R., Shome, M., Ray, R.K.: Evaluation of formability and mechanical behavior of laser-welded tailored blanks made of interstitial-free and dual-phase steels. Metall. Mater. Trans. A **40**(5), 1142–1152 (2009)
12. Reis, A., Teixeira, P., Duarte, J.F., Santos, A., da Rocha, A.B., Fernandes, A.A.: Tailored welded blanks—an experimental and numerical study in sheet metal forming on the effect of welding. Comput. Struct. **82**(17), 1435–1442 (2004)
13. Rojek, J., Hyrcza-Michalska, M., Bokota, A., Piekarska, W.: Determination of mechanical properties of the weld zone in tailor-welded blanks. Arch. Civil Mech. Eng. **12**(2), 156–162 (2012)
14. Shalin, M., Hiten, M.: Experimental analysis on effect of tool transverse feed, tool rotational speed and tool pin profile type on weld tensile strength of friction stir welded joint of AA 6061. Mater. Today: Proc. **5**(1), 487–493 (2018)
15. Tronci, A., McKenzie, R., Leal, R.M., Rodrigues, D.M.: Microstructural and mechanical characterisation of 5XXX-H111 friction stir welded tailored blanks. Sci. Technol. Weld. Join. **16**(5), 433–439 (2011)
16. Zadpoor, A.A., Sinke, J., Benedictus, R.: Finite element modeling of transition zone in friction stir welded tailor-made blanks. AIP Conf. Proc. AIP **908**(1), 1457–1462 (2007)
17. Cheng, C.H., Chan, L.C., Tang, C.Y., Chow, C.L.: Determination of true stress-strain curve for the weldment of aluminum laser-welded blanks. J. Laser Appl. **17**(3), 159–170 (2005)
18. Garware, M., Kridli, G.T., Mallick, P.K.: Tensile and fatigue behavior of friction-stir welded tailor-welded blank of aluminum alloy 5754. J. Mater. Eng. Perform. **19**(8), 1161–1171 (2010)

Chapter 3
Formability Evaluation of EN-10149-2 (S700mc) Steel Under In-Plane Plane Stretching Condition

Dereje H/Georgis, Perumalla Janaki Ramulu and Habtamu Beri

Abstract The current effort is intended to evaluate the forming behavior of the EN-10149-2 (S700mc) steel of 2.42 mm thickness under in-plane plane-stretching (IPPS) condition at ambient conditions. EN-10149-2 (S700mc) steel has good mechanical properties and lightweight that increase the application in automotive, aerospace, and industrial structural design compared to other metals. For which, the tensile and plain stretching test samples were prepared in three rolling directions 0°, 45°, and 90°. The mechanical properties such as the yield strength, tensile strength, uniform elongation, total elongation, hardening exponent (*n*-value), material strength coefficient (*K*) and plastic strain ratio (*R*-value) of EN-10149-2 (S700mc) steel are evaluated from the tensile testing. The microstructure and hardness of the material were also evaluated. IPPS evaluation test has been carried out by stamping circular grids at the middle of the specimen by which measured the manor and major strain using thickness gradient necking criterion (TGNC). The obtained results from experiments found that the maximum limit strain is appeared for the rolling direction of 45° and followed by 90° and 0°. This indicates that the maximum formability is in 45° rolling direction during the IPPS test.

Keywords Sheet metal forming · Forming limits strain · Mechanical properties · IPPS · TGNC · Circular grids · Major strain · Minor strain

D. H/Georgis · H. Beri
Program of Mechanical Design and Manufacturing Engineering, School of Mechanical, Chemical and Materials Engineering, Adama Science and Technology University, Adama, Ethiopia
e-mail: haleluyaa.haile1@gmail.com

H. Beri
e-mail: habtamu_beri@yahoo.com

P. J. Ramulu (✉)
Program of Mechanical Design and Manufacturing Engineering & CoEAME, School of Mechanical, Chemical and Materials Engineering, Adama Science and Technology University, Adama, Ethiopia
e-mail: perumalla.janaki@astu.edu.et

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_3

3.1 Introduction

Basically, metalworking processes are grouped into two classes, namely, bulk-forming and sheet metal forming. The sheet metal blank is plastically deformed into a three-dimensional shape without any critical changes in the thickness of the sheet materials. In other deformation cases, elastic recapturing or springback may be significant because the amounts of plastic and elastic deformations are equivalent. Some of the important sheet metals forming processes are deep drawing, stretching, bending and rubber pad forming [1]. Moreover, numerous car businesses vehicles composed under the thought of lightweight, crash execution, vitality sparing, and ecological perspective are of unique intrigue which is much essential to drive the vehicles' material selection process. Among all the materials existed in the market, steel material has a significant impact on many automotive parts where strength is more important. In this needy, steel industries are developing new grades of advanced high strength steels (AHSS) with high strength to weight ratio as lightweight structures [2]. Likewise, EN-10149-2 (S700mc) steel sheet has many application areas in modern manufacturing industries. It has also good mechanical properties and lightweight that increases the application in automotive, aerospace, and industrial structural design compared to other metals [3, 4]. In-plane tensile tests were conducted to produce strain paths from uniaxial tension to plane strain in mild steel and high strength steel. An in-plane forming limit test that was independent of tooling geometry and friction was created. Results showed that a wide sample with a wide but short gauge section needed to produce forming limit strains close to plane strain. The sample had sufficient area for clamping to avoid fracture outside the plane strain region in the center of the sample was suggested [5]. There were other studies on the in-plane plane-stretching behavior of different materials and their limiting strains were evaluated [6–11].

Although many studies have been investigated the formability of high strength steel low alloy (HSLA) both experimentally and numerically, the formability prediction of HSLA under the in-plane plane stretching condition not been studied much. Researchers published so far described the formability prediction based on the strain path diagrams, yield criteria. The detailed description of the formability of EN-10149-2 (S700mc) steel with 2.42 mm thickness under in-plane plane-stretching (IPPS) condition has not discussed. Based on the applications of the EN-10149-2 (S700mc) steel at different level industries, the present work generates the data for forming characteristics of this particular material which can be used directly without further formability tests.

3.2 Materials and Methodologies

3.2.1 Materials

The material selected for this study was high strength low alloy or micro-alloyed steel, EN 10149-2 steel sheet which has 2.42 mm thickness. Chemical composition of material is shown in Table 3.1.

3.2.2 Methodologies

Mechanical Properties and Micro-Structure Evaluation of EN-10149-2

The microstructure of EN-10149-2 steel sheet was studied by optical microscope of model Huvitz HR-300 series. Hardness was measured using Vickers's hardness testing machine. The mechanical properties of EN-10149-2 steel sheet were measured using tensile tests as per ASTM- B557 M by cutting sheet in different rolling directions such as 0°, 45°, and 90° and evaluated the mechanical properties and plastic strain ratio (R) as per ASTM- E517 through tensile tests. Tensile tests were performed at a nominal cross-head speed of 1 mm/min, at room temperature. On an ISO 68982-1 universal testing machine and repeated thrice for each set to check the reproducibility. Failure from grip or slippage during testing was not observed. Figure 3.1 shows the base metal tensile specimen. The mechanical properties like yield strength, ultimate tensile strength, uniform elongation, strain hardening coefficient (n) and strength coefficient (K) were evaluated as per the standard procedure after testing the samples till failure. The load-stroke behavior obtained during testing and converted into engineering stress–strain and true stress–strain

Table 3.1 EN-10149-2 (S700mc) steel chemical composition (in wt%)

C	Si	Mn	P	S	Cr	Mo	Ni	Cu	W	Al	Nb	Ti	Fe
0.06	0.04	1.46	0.003	0.002	0.01	0.02	0.03	0.01	0.04	0.05	0.03	0.10	98.15

Fig. 3.1 Specimens prepared for tensile tests

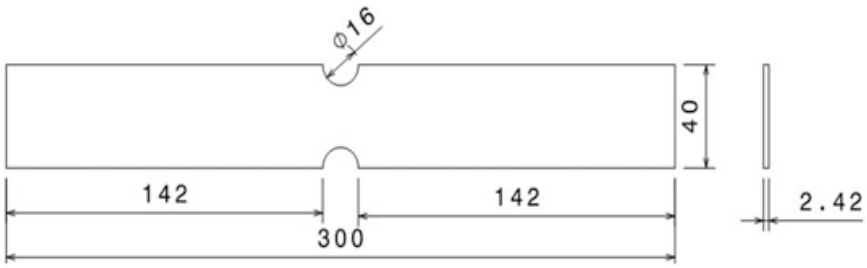


Fig. 3.2 Dimensions for plane–strain tension test

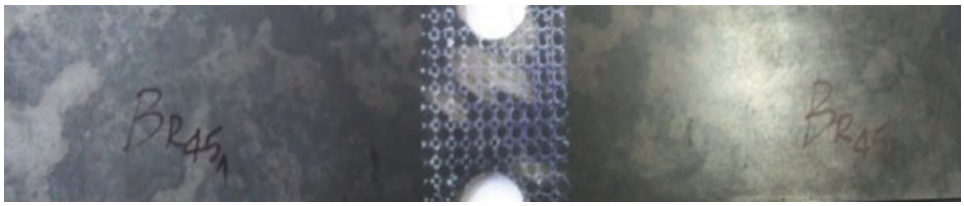


Fig. 3.3 Typical plane–strain tension specimen before test

plot for evaluating the mechanical properties. The plastic strain ratios at different rolling directions were obtained after testing the base material to 10% plastic strain.

Plane–Strain Tension Testing

In this study, plane stretching testing samples were prepared based on the width constraints method. Figures 3.2 and 3.3 show the schematic and original in-plane plane-stretching test specimens with circular grids of 3 mm in diameters on experiment specimen.

3.3 Results and Discussion

3.3.1 Mechanical Properties of Sheet Metal

Figures 3.4 and 3.5 are shown the material EN-10149-2 steel uniaxial tensile test specimen before and after the testing, respectively. Figure 3.5 representing the tested specimen at different rolling direction 0°, 45°, and 90°.

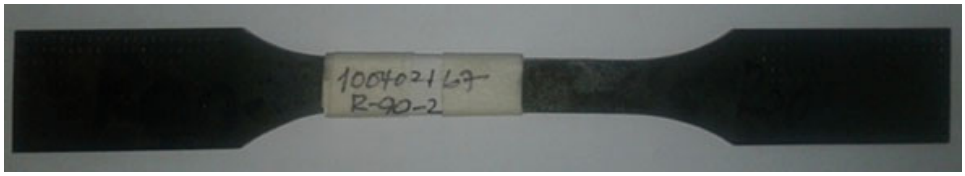


Fig. 3.4 EN-10149-2 specimen before tensile test

Fig. 3.5 EN-10149-2 specimen after tensile test

The mechanical properties of the EN-10149-2 sheet metal are evaluated from the uniaxial tensile tests tabulated in Table 3.2. The evaluated mechanical properties like yield stress, uniform elongation, tensile stress, material strength coefficient (K) and strain hardening coefficient (n). The base material engineering stress–strain behavior in different rolling directions is shown in Fig. 3.6. Table 3.2 shows the tensile properties of all three reputations and Table 3.3 summarizes the mechanical properties of EN-10149-2 base material. The yield strength and ultimate tensile strength are found to be within 972–1020 MPa and 1018–1054 MPa, respectively. The uniform elongation and strain hardening exponent are about 6.8% and 0.22 respectively.

Table 3.4 indicates summarized mechanical properties of EN-10149-2 (S700mc) sheet steel. The mechanical properties and anisotropy value are shown in Table 3.4.

Table 3.2 Tensile test results of EN-10149-2 sheet metal

Rolling direction	Trial	YS (MPa)	Uniform elongation (%)	Tensile strength (MPa)	Material strength coff. (K) (MPa)	Strain hardening exponent (n)
0°	1	951.5	5.4	999.3	1512.6	0.142
	2	997.2	7.6	1021.3	1744.5	0.211
	3	983.0	7.6	1036.1	1763.6	0.209
45°	1	927.1	5.2	973.8	1425.7	0.126
	2	968.3	8.2	1023.6	1776.1	0.222
	3	1020.7	6.5	1060.9	1603.1	0.1436
90°	1	1058.8	6.8	1090.5	1628.7	0.1378
	2	1009.2	6.8	1039.62	1796.1	0.215
	3	994.3	6.9	1034.6	1676.1	0.1804

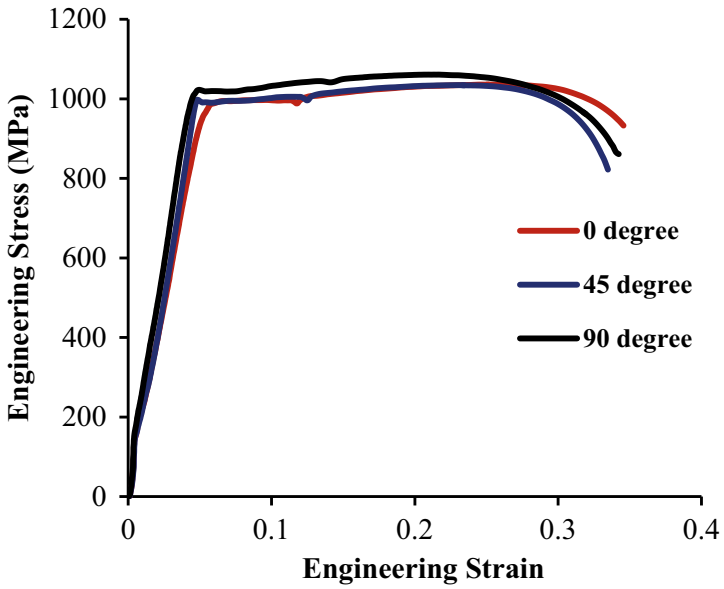


Fig. 3.6 Tensile behavior of base metal in different rolling directions

Table 3.3 Average results of mechanical properties of EN-10149-2 material

Rolling direction	YS (MPa)	Uniform elongation (%)	Tensile strength (MPa)	Material strength coff (K) (MPa)	Strain hardening exponent (n)
0°	977.3 ± 20	6.87 ± 10	1018.9 ± 20	1673.6 ± 50	0.187 ± 0.014
45°	972.03 ± 30	6.63 ± 3	1018.4 ± 31	1601.6 ± 70	0.164 ± 0.048
90°	1020.7 ± 20	6.8 ± 0.1	1054.9 ± 40	1700.3 ± 70	0.178 ± 0.037

3.3.2 Microstructure of EN-10149-2

Metallographic investigations were carried out on the sample of base metal. Etching with Nital 2% for 10 s; this etchant was used to reveal the banded structure, and the grain boundaries and carbide containing phases. A typical microstructure of EN-10149-2 is shown in Fig. 3.7.

From the microstructure of base metal in Fig. 3.7 indicates the ferrite, pearlite structures along with carbon content. Ferrite and pearlite structure is responsible for the better mechanical properties.

3.3.3 Evaluation of Hardness

Hardness was measured using Vickers hardness (VH) testing machine at five different locations with 10 kg load. The average hardness number of the sample is noted as 124.94 VHN.

Table 3.4 Mechanical property and Plastic Anisotropy (r-value) of EN-10149-2 sheet metal

Rolling direction	YS (MPa)	Uniform elongation (%)	Tensile strength (MPa)	Material strength coff. (K) (MPa)	Strain hardening exponent (n)	R_0	R_{45}	R_{90}
0°	977.3 ± 20	6.87 ± 10	1018.9 ± 20	1673.6 ± 50	0.187 ± 0.014	0.992	1.287	1.268
45°	972.03 ± 30	6.63 ± 3	1018.4 ± 31	1601.6 ± 70	0.164 ± 0.048			
90°	1020.7 ± 20	6.8 ± 0.1	1054.9 ± 40	1700.3 ± 70	0.178 ± 0.037			

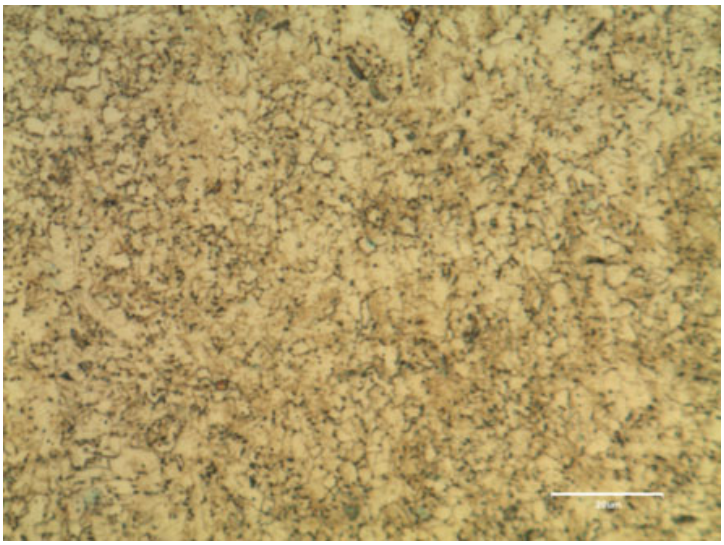


Fig. 3.7 Microstructure of EN-10149-2 at room temperature

3.3.4 In-Plane Plane Stretching (IPPS) Test

The samples were prepared in three rolling directions namely, 0°, 45° and 90° in number two per each rolling direction. The limit strains are evaluated by measuring the length of the major and minor axes of deformed circular grids at the failure region of EN-10149-2 sheet blanks. Figure 3.8 shows circular grids printed on the surface of base materials before deformation and Fig. 3.9 grids elongation after deformation, during IPPS formability tests. The digitized caliper two digits decimal was used to measure the circular grids before and after deformation of base materials for evaluating true major and minor limit strains.

Failure Criteria used for Limit Strains Evaluation

Thickness gradient-based necking criterion (TGNC)

Thickness gradient-based necking criterion is used to obtain forming limit strains of EN-10149-2 sheet blanks [12–15]. Necking can be understood as a thinning phenomenon in any sheet metal during the deformation process. It mostly distinguishes necking as a localized thinning phenomenon, i.e., a localized region within the deforming sheet where a steep gradient in thickness develops. A local critical

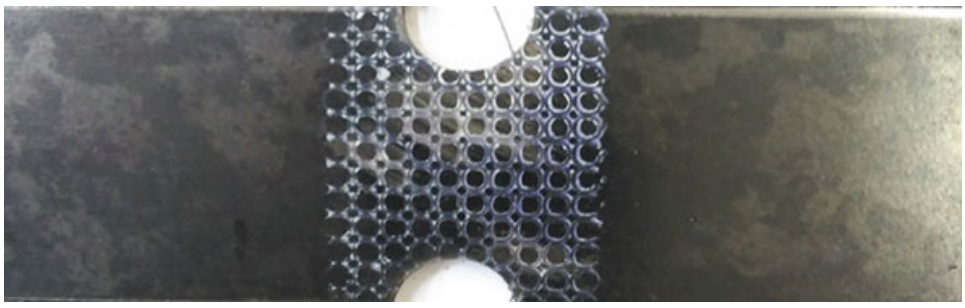


Fig. 3.8 Method of grid measurement before deformation

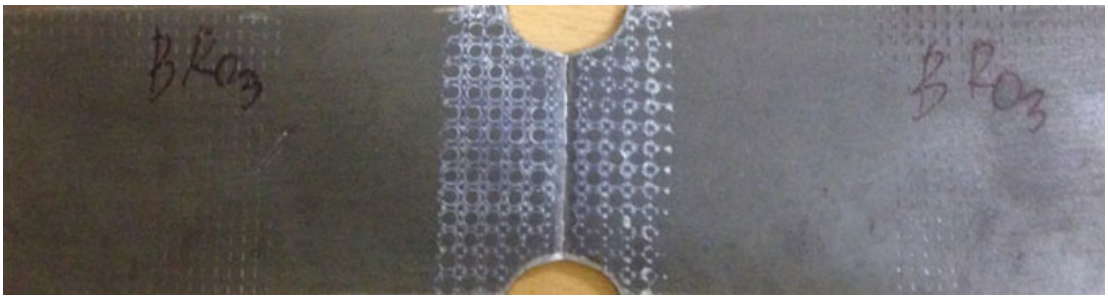


Fig. 3.9 Method of grid measurement after deformation

thickness gradient (Rc) must exist at the onset of a locally visible neck. The maximum major and minor strain in the thinner element is then reported as the forming limit strain. It is found from the earlier experiments that necking occurs in the sheet metal when the thickness gradient (ratio of the thickness of neighboring elements or circles) falls below 0.92 [12–15].

$$R_{\text{thickness gradient}} = \frac{\text{thickness}_{\text{Necking element}}}{\text{thickness}_{\text{Neighboring element}}} = \frac{t_n}{t_{n-1}} \text{or} \frac{t_n}{t_{n+1}} \leq R_{\text{Cri}} \Rightarrow \text{material failure or necking}$$

$$R_{\text{Critical}} = 0.92$$

Table 3.5 shows the limit strain results using TGNC from in-plane plane-stretching tested samples in three different rolling directions. Figure 3.10

Table 3.5 Limit strains results using TGNC from in-plane plane-stretching tested

Orientations	Trials	Measured circle diameter (mm)	Measured value- original diameter (mm)	$R_{\text{thickness gradient}}$	Major strain	Minor strain
0°	1	3.124	0.124	0.9193	0.124	0.104
		3.114	0.114			
	2	3.231	0.231	0.9166	0.252	0.114
		3.252	0.252			
45°	1	3.572	0.572	0.91783	0.57	0.541
		3.525	0.525			
	2	3.424	0.424	0.91973	0.464	0.192
		3.461	0.464			
90°	1	3.451	0.451	0.92011	0.451	0.353
		3.416	0.416			
	2	3.294	0.294	0.91588	0.321	0.264
		3.321	0.321			

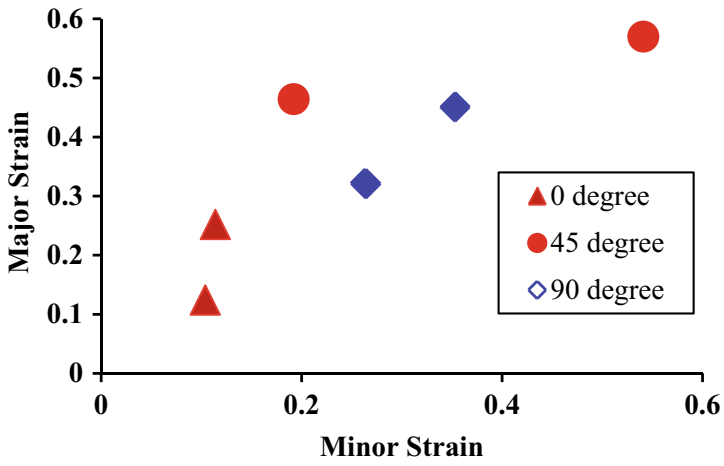


Fig. 3.10 Comparison of limit strain of in-plane plane-stretching test on three rolling directions with two reputations

shows the limit strain comparison of three (0°, 45°, 90°) different rolling direction with two reputations. Limit strain is evaluated using the TGNC after the experiment on three rolling directions. The maximum limit strain is appeared for the rolling direction of 45° and followed by 90° and 0°. This indicates that the maximum formability is in 45° rolling direction during the in-plane plane-stretching test.

3.4 Conclusions

In this work, in-plane plane stretching (IPPS) test is conducted in order to determine limit strains of the EN-10149-02 steel material. From the results, the following conclusions are drawn as follows:

- EN-10149-02 steel material had the highest yield and tensile strength in the transverse direction (90°) and a diagonal direction (45°). The strain handing exponent and total elongation is more at 0° and decreases in the diagonal direction (45°).
- The microstructure result shows that the material constitutes more pearlite matrix grain structure with small carbide boundaries which leads to material more ductile and formable. Vickers's hardness of EN-10149-02 steel material value is 122.24 VHN.
- In-plane plane stretching experimental tests the limit strain evaluated using the TGNC after the experiment on three rolling directions. The maximum limit strain is appeared for the rolling direction of 45° and followed by 90° and 0°. This indicates that the maximum formability is in 45° rolling direction during the in-plane plane-stretching test.

References

1. Handbook, A.M.: Forming and forging. ASM Int. 14 (1993)
2. Panich, S., Barlat, F., Uthaisangsuk, V., Suranuntchai, S., Jirathearanat, S.: Experimental and theoretical formability analysis using strain and stress-based forming limit diagram for advanced high strength steels. Mater. Des. **51**, 756–766 (2013)
3. Garcia, C.I.: High strength low alloyed (HSLA) steels. In: Automotive Steels, pp. 145–167 (2017)
4. Ozturk, F., Lee, D.: Experimental and numerical analysis of out-of-plane formability test. J. Mater. Process. Technol. **170**(1–2), 247–253 (2005)
5. Holmberg, S., Enquist, B., Thilderkvist, P.: Evaluation of sheet metal formability by tensile tests. J. Mater. Process. Technol. **145**(1), 72–83 (2004)
6. Keeler, S., Kimchi, M.: Advanced high-strength steels application guidelines V5. World Auto Steel (2015)
7. Agnew, S.R., Duygulu, Ö.: Plastic anisotropy and the role of non-basal slip in magnesium alloy AZ31B. Int. J. Plast. **21**(6), 1161–1193 (2005)
8. Drewett, L., Bremer, S., Liebeherr, M., De Waele, W., Martin-Meizoso, A., Brozda, J., Gubeljak, N.: HIPERC: a novel, high performance, economic steel concept for line pipe and general structural use. Publications office of the European Union (2010)
9. Janaki Ramulu, P., Satish Babu, A., Ganesh Narayanan, R., Deva Prasad, S., Srinivasa Rao, P.: The behavior of friction stir welded (FSW) sheets of AA6061-T6 during in-plane stretching test. Proced. Eng. **64**, 862–867 (2013)
10. Kumar, M., Kailas, S.V., Ganesh Narayanan, R.: In-plane plane-strain formability investigation of friction stir welded sheets made of dissimilar aluminium alloys. Appl. Mech. Mater. **446–447**, 301–305 (2014)
11. Satheesh Kumar, V., Ganesh Narayanan, R.: In-plane plane strain formability of adhesive-bonded steel sheets: influence of adhesive properties. Int. J. Adv. Manuf. Technol. **76**, 993–1009 (2015)
12. Janaki Ramulu, P., Ganesh Narayanan, R.: Weld zone representation during the formability prediction of friction stir welded blanks with similar thickness sheets. J. Strain Anal. Eng. Des. **46**, 456–477 (2011)
13. Ramulu, P.J., Narayanan, R.G.: Weld zone representation methods during the stretching of friction stir welded blanks with dissimilar sheet thickness: a study using numerical simulations. Materialwiss. Werkstofftech. **43**(3), 241–252 (2012)
14. Ramulu, P.J., Narayanan, R.G.: Experimental evaluation and prediction of forming limit of FSW blanks made of AA 6061 T6 sheets at different weld orientations and weld locations. Materialwiss. Werkstofftech. **44**(6), 527–540 (2013)
15. Ramulu, P.J., Kailas, S.V., Narayanan, R.G.: Formability of friction stir welded sheets made of AA 6061–T6 at different weld orientations and weld locations. Int. J. Mater. Prod. Technol. **50**(2), 147–160 (2015)

Chapter 4
Stretch Formability Prediction of AA6023-T6 Alloy Sheet Under Two Different Heating Conditions

Manaye Mathewos Handiso, Perumalla Janaki Ramulu and G. Somasundaram

Abstract Sheet metal forming is a process widely used in the manufacturing industry. There are several sheet metals forming processes are existing including stretch forming process. In the present work, stretch formability of AA6023-T6 sheet of 2 mm thickness at room temperature and annealed sheet at 400 °C has been performed. For which mechanical properties were evaluated through tensile test for all the conditions. For stretching operations, simulations were performed using limit dome height (LDH) test using PAM STAMP 2G software. During LDH test, the maximum value of LDH for annealed sheets of AA6023-T6 at 400 °C and minimum value for sheets at room temperature is observed. Uniform thickness distribution for annealed blanks at 400 °C is observed compared to room temperature. The room temperature stretch formability of AA6023-T6 alloy sheet has 39.5 maximum dome height and annealed sheet at 400 °C has 43 maximum dome heights. By which, annealed AA6023-T6 sheet has good stretch formability compared to room temperature stretch formability.

Keywords Mechanical properties · Formability · Forming limit diagram · Limit dome height · Thickness distribution

M. M. Handiso · G. Somasundaram
Program of Mechanical Design and Manufacturing Engineering, School of Mechanical, Chemical and Materials Engineering, Adama Science and Technology University, Adama, Ethiopia
e-mail: manayehandiso@gmail.com

G. Somasundaram
e-mail: somasundarram@gmail.com

P. J. Ramulu (✉)
Program of Mechanical Design and Manufacturing Engineering and CoEAME, School of Mechanical, Chemical and Materials Engineering, Adama Science and Technology University, Adama, Ethiopia
e-mail: perumalla.janaki@astu.edu.et

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_4

4.1 Introduction

Sheet metal forming is one of the most used processes in the manufacturing industry. There are many sheet metals forming processes including stretch forming process. In stretch forming process the sheet metal is clamped around its edge and stretched over a die or forming block, which moves upward, downward or sideways, depend up on the machine used. Stretch forming is applicable to produce aircraft-wing skin panel, automotive doors panel and window frames etc. The design of components needs the analysis of instability and determining of possible deformation of sheet metal. The determination of possible forming limits and analysis of instability of sheet metal was predicted from forming limit diagrams (FLDs) [1]. The applied load on sheet metal causes the plastic deformation when it was greater than yield strength of the metal. The stress induced during the process was greater than yield strength, but less than fracture strength of the material [2]. Another important factor which has influence on sheet metal forming is anisotropy. It is the directional variation of mechanical properties. In other words, the reaction of materials to applied stress at different direction has different reaction to the applied stress. The strength and formability of the materials are the main properties which are very important for industrial applications of sheet metals [3]. The forming behavior of a certain sheet metals always depends up on the forming temperature during forming process. For instance, aluminum series of 6xxx alloys had quite less level of formability at room temperature than aluminum series of 5xxx alloys [4, 5]. The amount of strain hardening in the material decided the formability of the materials. The tensile test was one of the mechanical properties' evaluation process in which the strain hardening exponent of the sheet metal evaluated. If the tensile tests of the sheet metal had maximum percentage of elongations, it can be predicted as the sheet metal with a good formability [6]. Forming temperature had significant influence on formability. Many researches were developed on the effect of temperature on the formability of different sheet metals. For example, Nolan et al. [7] studied the formability of a 7000 series alloy under different forming temperature. They performed the tensile tests and deep drawing tests to evaluate the forming behavior. The experiment was taken at both elevated and room temperature. In the drawing process of the metal, the formability of the given sheet metal increased with the increase of forming temperature. This warm forming behavior of the sheet metal indicated that the condition of formability of the alloy improved with that of the forming temperature. For the warm forming condition of 7000 series alloy in general the forming temperature observed in the range of 200–600 °C. The forming condition of stretch flange has a good formability at warm forming condition than cold forming process. Tokita et al. (2017) [8] studied the forming conditions of the stretch formability of two different high strength steels. The forming was preferred on warm forming by spherical stretch forming tests at various forming temperature including room temperature up to 600 °C. Their experiment showed the suitable forming temperature for stretch formability of steels. The observations from the experiments indicated the value of

maximum dome height related to the maximum formability indications. There was maximum dome height for test of steel at elevate temperature and minimum dome height at room temperature forming process. High initial forming temperature was one of the affecting factors which causes the loss of quality in the forming process and also the coefficient of friction has a significant influence on the forming behavior of the alloy. Ali and Jumah [2] investigated the forming behavior of annealed aluminum alloy at different forming temperature. The formability of alloy increased with the increasing of the forming temperature. When temperature was at maximum level, the formability was significantly good but the drawback at this forming temperature was the excessive ductility. In this case, the application of alloy limited to its strength. In general, elongation increased, with increase of temperature of AA6061-T6 alloy during tension test. Moreover, at elevated temperature greater or equal to 280 °C, the elongation increases with decreasing strain rate. With spreading of the deformation area, the deformation was more uniform than that at high strain rate; elongation increase gradually [9]. The increase or reduction of strain and plastic anisotropies related to texture intensity and texture distribution should be the main factors for good stretch formability. Refining strain (ductility) of the materials also influence the stretch formability [10, 11]. Forming limit diagram (FLD) was used to predict the materials formability behavior under different forming condition. It played a great role in industries, because it was also the most applicable method to evaluate the formability of sheet metals [12–15]. The value for these parameters calculated with respect to the specimen just before fracture. Forming limit diagrams (FLDs) are a suitable and often used mechanism for classification of the formability and investigation of sheet metal forming process. Forming limit diagram (FLD) also used to indicate the forming limit of a given sheet metals at different state of deformation. For different materials the upper limit curve (ULC) of the forming limit diagram provides breaking, and they calculated with respect to specimens just before breaking. The increase in local necking causes the increase in deformation of a certain specimen at stress concentration area. While the local necking increased further there was no change in stress to the rest of part of the specimen which was not subjected to the necking [16]. The effects of punch and die temperature on the formability of a cup shaped AA3003 aluminum alloy sheet. The formability improves as the temperature difference between the punch and the die increases. Also, the forming speed does not affect localized necking for the given punch velocities in the study. The simulations were captured for both punch force and failure location. The important parameters to control the thickness reduction in the blank were identified as friction and blank holder force. The finite element analysis model was validated with the selected experimental cases. Stretch formability of sheet metals depends on temperature of sheets. Increase in sheet metals temperature increase the stretch formability, but ductility of sheet metals was observed at elevated sheet temperature from the existing literature. To overcome the excessive ductility of sheet metal, annealing in between elevated temperature and room temperature was needed [17]. Based on the above observations, the present study objective is to predict the stretch formability of AA6023-T6 sheet of 2 mm thickness at room temperature and annealed at 400 °C.

4.2 Materials and Methodologies

4.2.1 Base Material

The Material selected for this study was AA6023-T6 alloy sheet which has 2 mm thickness. AA6023-T6 alloy sheet. Table 4.1 shows the nominal chemical composition of this material.

4.2.2 Tensile Test

The mechanical properties of base metal were measured using tensile tests as per ASTM- B557 M. The base metal was cut in different rolling directions like 0°, 45° and 90° and the mechanical properties and plastic strain ratio (R) were evaluated through tensile tests as per ASTM- E517. Tensile tests were performed at a nominal cross-head speed of 1 mm/min, at room temperature and annealed at 400 °C specimens on an ISO 68982-1 universal testing machine tensile testing machine and repeated thrice for each set to check the reproducibility. Failure from grip or slippage during testing was not observed. Figure 4.1 shows base metal tensile specimen. The mechanical properties of base material like yield strength, ultimate tensile strength, total elongation, uniform elongation, strain hardening coefficient (n) and strength coefficient (K) were evaluated as per the standard procedure after testing the samples till failure. The load-stroke behavior obtained during testing was converted into engineering stress–strain and true stress–strain plot for evaluating the mechanical properties. The plastic strain ratios at different rolling directions were obtained after testing the base material to 10% plastic strain.

Table 4.1 Chemical composition of AA6023-T6 alloy sheet (in wt%)

Chemical composition								
Si	Fe	Cu	Mn	Mg	Bi	Sn	Residuals	Al
0.57	0.5	0.19	0.24	0.41	0.3	1.2	0.59	96

Fig. 4.1 Typical specimen for tensile tests

4.2.3 Simulation Details

Finite element method was used for numerical analysis of sheet metal forming operations. The numerical analysis of stretch forming process was done by PAM STAMP 2G software. The input model consists of die, blank, blank holder, draw and punch were constructed in pre-processor. The punch, die and blank holder had considered as rigid body. The fine meshing was done on the surface of die, blank holder and punch, since, the blank has been considered as a deformable body. In limit dome height (LDH) test, the blank with 2 mm thickness is deformed by a hemispherical punch (diameter: 101.6 mm) inside opening of a die with 105.9 mm diameter. For simulations process, five tests with different size of blanks (100 mm × 200 mm, 125 mm × 200 mm, 150 mm × 200 mm, 175 mm × 200 mm, and 200 mm × 200 mm) were used for each (0°, 45° and 90°) orientation of samples. The tools required for LDH test, punch, die, blank holder, draw bead, and blank generated in solid works, a CAD packages, and meshed using delta mesh facility in PAM STAMP 2G. Both Hollomon's strain hardening law and Hill's 1948 isotropic hardening yield criterion was used as the plasticity model. The base material comprised quadrilateral shell elements of Belytschko–Tsay formulation with five through-thickness integration points. A uniform mesh size of 1 mm was used for all the simulations. The friction coefficient m was assumed to be 0.12 and is kept constant throughout the study. The blank was as the active body and the punch force used was 80 kN to avoid the drawing and tearing near draw beads during forming prediction. For prediction, the forming limit strain thickness gradient necking criterion (TGNC) was used [18–21].

4.3 Results and Discussion

4.3.1 Mechanical Properties of Base Material

The base material true stress–strain behavior at different rolling directions of two different heating conditions are shown in Figs. 4.2 and 4.3. Tables 4.2 and 4.3 summarize the mechanical properties of AA 6061 base material obtained in the present work. The yield strength and ultimate tensile strength are found to be within 180–200 MPa and 193–206 Mpa, respectively.

Heat treatment of the aluminum alloys causes change in mechanical and anisotropy values. Table 4.3 indicates the value of mechanical properties and plastic anisotropy of AA6023-T6 alloy sheet after annealing at 400 °C with one and half hours soaking time.

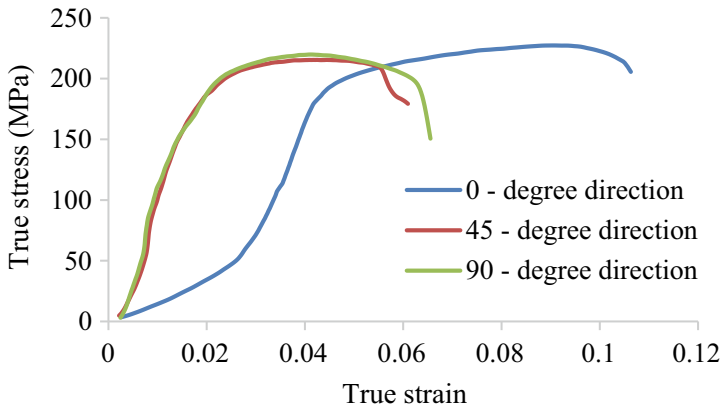


Fig. 4.2 True stress—true strain curve for samples at room temperature

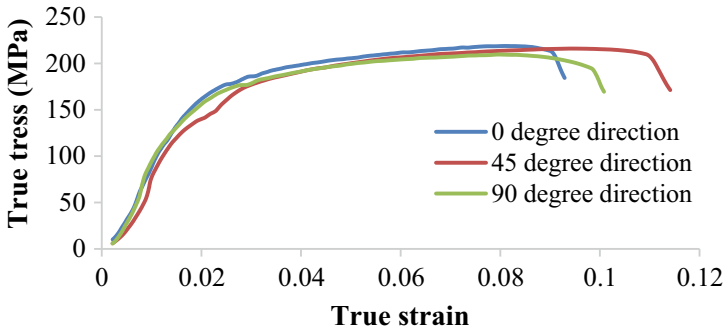


Fig. 4.3 True stress—true strain curve for annealed samples at 400 °C

Table 4.2 Mechanical property of AA6023-T6 alloy sheet at room temperature

Orientation of samples	UE (mm)	UTS (MPa)	σ_{ys} (MPa)	K (MPa)	n	R_0	R_{45}	R_{90}
0°	0.38	200 ± 9	180 ± 3	313 ± 40	0.132	0.49	0.68	0.56
45°	0.54	208 ± 4	190 ± 18	338 ± 20	0.134			
90°	0.52	212 ± 12	200 ± 12	358 ± 20	0.133			

Table 4.3 Mechanical property of AA6023-T6 alloy sheet after heat treatment

Orientation of samples	UE (mm)	UTS (MPa)	σ_{ys} (MPa)	K (MPa)	n	R_0	R_{45}	R_{90}
0°	0.19	206 ± 13	170 ± 17	326 ± 33	0.149	0.43	0.58	0.47
45°	0.21	197 ± 2	168 ± 5	323 ± 16	0.162			
90°	0.25	193 ± 6	166 ± 7	315 ± 28	0.155			

4.3.2 Evaluation of Forming Limit Diagram

The forming limit diagram indicates the stretch formability characteristics of the sheet metal. The forming limit diagram is also the forming limit curve, which is one of necessary value during practical assessment of sheet stretch forming process. Figures 4.4 and 4.5 show the comparison between the form limit curves of AA6023-T6 alloy sheets at room temperature and annealed alloy sheets and Table 4.4 indicate major and minor strain obtained by thickness gradient-based necking criterion at LDH.

Figure 4.4 shows the form limit curves of a sheet that predicts the stretch formability of AA6023-T6 alloy sheet at room temperature. The distribution of the curves of blanks with different dimension to the upper limit curve (ULC) is scattered. There are some blanks lie above the upper limit curve ULC which indicates fracture before the necking. The stretch formability of these blanks which have fracture earlier than the forming limit is not acceptable. The earlier formation of fracture caused due to the cold forming condition. The maximum limit dome height LDH observed in the stretch formability of blanks under room temperature forming

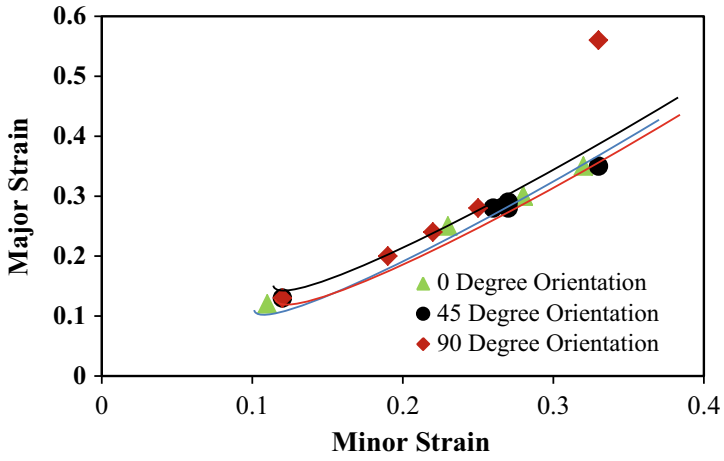


Fig. 4.4 Form limit curve of AA6023-T6 alloy sheet at room temperature forming condition

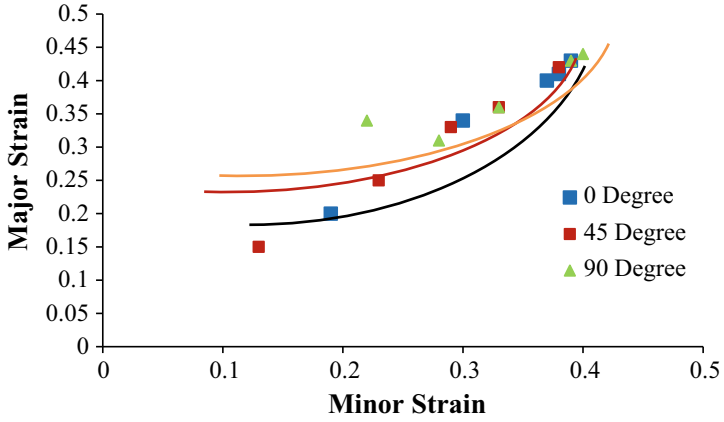


Fig. 4.5 Form limit curve of AA6023-T6 alloy sheet after annealing condition

Table 4.4 Major and minor strain at LDH for room temperature blanks

	0° oriented					45° oriented					90° oriented				
Blanks	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Major strain	0.12	0.25	0.30	0.30	0.35	0.13	0.29	0.35	0.28	0.28	0.13	0.24	0.56	0.20	0.28
Minor strain	0.11	0.23	0.28	0.28	0.32	0.12	0.27	0.33	0.27	0.26	0.12	0.22	0.33	0.19	0.25
LDH	26	37.2	39	39	39	29.5	38.5	39	39.5	39.5	32	39	39.5	39.5	39

condition for blanks with dimension of 150 mm × 200 mm, 175 mm × 200 mm and 200 mm × 200 mm is 39.5 mm. From the value of maximum limit dome height LDH, these blanks have good formability compared to the others blanks with minimum value of limit dome height. Table 4.5 indicates major and minor strain obtained by thickness gradient-based necking criterion at LDH for annealed blanks.

Figure 4.5 shows the form limit curves of sheet that predicts the stretch formability of annealed AA6023-T6 alloy sheet. The distribution of the curves of blanks with different dimension to the upper limit curve is scattered. The distribution of the curves of the blanks is under limit of the upper forming limit diagram. This indicates there is no earlier fracture occurred before necking. The stretch formability of these blanks which are scattered blow the upper forming limit curve and near to the upper limit curve is recommended blanks for attaining a good formability. The maximum limit dome height LDH observed in the stretch formability of blank which are formed after annealing condition with dimension of 150 mm × 200 mm, 175 mm × 200 mm and 200 mm × 200 mm is 43 mm. From the value of maximum limit dome height LDH, these blanks have good formability compared to the others blanks with minimum value of limit dome height. The comparison between the maximum limit dome heights of the blanks at room temperature annealed alloy sheet shows the stretch formability of AA6023-T6 alloy sheets. The average maximum limit dome height LDH for room temperature forming condition of the blanks is 39.5 mm and the average maximum limit dome height LDH is 43 mm. Therefore, annealed AA6023-T6 alloy sheet has a good formability at annealing temperature of 400 °C and 90 min soaking time than sheet at room temperature.

Table 4.5 Major and minor strain at LDH for annealed blanks

	0° oriented					45° oriented					90° oriented				
Blanks	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Major strain	0.20	0.34	0.40	0.43	0.41	0.15	0.25	0.33	0.36	0.42	0.34	0.31	0.36	0.43	0.44
Minor strain	0.19	0.30	0.37	0.39	0.38	0.13	0.23	0.29	0.33	0.38	0.22	0.28	0.33	0.39	0.40
LDH	31	40.5	41.5	42	41.5	32	42.5	43	43	43	30.5	41.5	43	43	43

4.4 Conclusions

From the present work, the following conclusions are drawn.

- LDH from simulation justified the stretch formability behavior of AA6023-T6 alloy sheet. A blank at room temperature with dimensions of 150 mm × 200 mm, 175 mm × 200 mm and 200 mm × 200 mm have 39.5 mm maximum dome height and blank with dimension of 100 mm × 200 mm and 125 mm 200 mm at room temperature have 26 mm minimum dome height.
- Annealed blanks at 400 °C with 150 mm × 200 mm, 175 mm × 200 mm and 200 × 200 dimensions have maximum dome height compared to 100 mm 200 mm and 125 mm × 200 mm blanks at room temperature. The maximum dome height observed in simulation is 43 mm and minimum 31 mm. By which, the blanks with dimension of 150 mm × 200 mm, 175 mm × 200 mm and 200 mm × 200 mm have good stretch formability compared to blanks with dimension of 100 mm × 200 mm and 125 mm × 200 mm at room temperature.
- The comparison between stretch formability of AA6023-T6 alloy sheet at room temperature and annealed sheet at 400 °C shows, AA6023-T6 alloy sheet at room temperature has 39.5 mm maximum dome height and annealed sheet at 400 °C has 43 mm maximum LDH, therefore, annealed sheet has good stretch formability.

References

1. Stanislav, N.A.S., Emil, E.C.S.: Forming limit diagrams as an important indicator of progressive high-strength steel sheets forming. Transfer inovácií 17–20 (2013)
2. Ali, W.J., Th Jumah, O.: Warm forming of aluminum alloy 2024 at different temperatures. Al-Rafadain Eng. J. **20**(2) (2012)
3. Cardoso, R.P.R., Adetoro, O.B.: A generalisation of the Hill's quadratic yield function for planar plastic anisotropy to consider loading direction. Int. J. Mech. Sci. **128**, 253–268 (2017)
4. Mohamed, M.S., Ismail, A.: Review on sheet metal forming process of aluminium alloys. In: 17th International Conference on Applied Mechanics and Mechanical Engineering (2016)
5. Klos, A., Kahrimanidis, A., Wortberg, D., Merklein, M.: Experimental and numerical studies on the forming behavior of high strain Al–Mg–Si (–Cu) sheet alloys. Proced. Eng. **183**, 95–100 (2017)
6. Abovyan, T., Kridli, G.T., Friedman, P.A., Ayoub, G.: Formability prediction of aluminum sheet alloys under isothermal forming conditions. J. Manuf. Process. **20**, 406–413 (2015)
7. Nolan, R.A., Kulas, M.A., Prangnell, P.B., Quinta da Fonseca, J.: The effect of temperature on the formability of a high strength aluminium automotive alloy. University of Manchester, Manchester (2014)
8. Tokita, Y., Nakagaito, T., Tamai, Y., Urabe, T.: Stretch formability of high strength steel sheets in warm forming. J. Mater. Process. Technol. **246**, 77–84 (2017)
9. Chen, Z., Fang, G., Zhao, J.Q.: Formability evaluation of aluminum alloy 6061-T6 sheet at room and elevated temperatures. J. Mater. Eng. Perform. **26**(9), 4626–4637 (2017)

10. Suzuki, K., Chino, Y., Huang, X., Yuasa, M., Mabuchi, M.: Enhanced room-temperature stretch formability of Mg–0.2 mass% Ce alloy sheets processed by combination of high-temperature pre-annealing and warm rolling. Mater. Trans. **56**(7), 1096–1101 (2015)
11. An, Y.G.: Strain path change effects on stretch formability of 6082 aluminium alloy. Mater. Sci. Technol. **17**(3), 258–263 (2001)
12. Chiba, R., Takeuchi, H., Nakamura, R.: Forming-limit prediction of perforated aluminium sheets with square holes. J. Strain Anal. Eng. Des. **50**(6), 391–404 (2015)
13. Lopez, A.M., van den Boogaard, A.H.: Formability limit curves under stretch-bending. Res. Progr. Mater. Innov. Inst. M2i 1–2 (2011)
14. Shao, Z., Bai, Q., Li, N., Lin, J., Shi, Z., Stanton, M., Dean, T.: Experimental investigation of forming limit curves and deformation features in warm forming of an aluminium alloy. Proc. Inst. Mech. Eng. Part B. J. Eng. Manuf. **232**(3), 465–474 (2018)
15. Butuc, M.C., Teodosiu, C., Barlat, F., Gracio, J.J.: Analysis of sheet metal formability through isotropic and kinematic hardening models. Eur. J. Mech. A/Solids **30**(4), 532–546 (2011)
16. Holmberg, S., Enquist, B., Thilderkvist, P.: Evaluation of sheet metal formability by tensile tests. J. Mater. Process. Technol. **145**(1), 72–83 (2004)
17. Bagheriasl, R.; Formability of aluminum alloy sheet at elevated temperature (2012)
18. Janaki Ramulu, P., Ganesh Narayanan, R.: Weld zone representation during the formability prediction of friction stir welded blanks with similar thickness sheets. J. Strain Anal. Eng. Des. **46**(6), 456–477 (2011)
19. Ramulu, P.J., Narayanan, R.G.: Weld zone representation methods during the stretching of friction stir welded blanks with dissimilar sheet thickness: A study using numerical simulations. Materialwiss. Werkstofftech. **43**(3), 241–252 (2012)
20. Ramulu, P.J., Narayanan, R.G.: Experimental evaluation and prediction of forming limit of FSW blanks made of AA 6061 T6 sheets at different weld orientations and weld locations. Materialwiss. Werkstofftech. **44**(6), 527–540 (2013)
21. Ramulu, P.J., Kailas, S.V., Narayanan, R.G.: Formability of friction stir welded sheets made of AA 6061–T6 at different weld orientations and weld locations. Int. J. Mater. Prod. Technol. **50**(2), 147–160 (2015)

Chapter 5
FE Analysis of Single Point Incremental Forming (SPIF) for Tailor Welded Blanks (TWBs)

Shalin Marathe and Harit Raval

Abstract Fuel economy along with cost reduction is the ultimate target of the automobile industries and in order to achieve that, reduction in weight of the automobile body parts should be done. The homogeneous blanks are replaced by Tailor Welded Blanks (TWBs) while manufacturing the components of an automobile vehicle in order to reduce the weight and cost of the components. TWB is the outcome of joining two or more than two materials having different properties. Despite of having extremely important advantages, technology of TWB is associated with the limitation of weld line shift and formability reduction during the forming process. During the conventional forming process, formability of the TWBs is found to be decreasing and weld line shift is observed to be on higher side. In order to improve the formability of homogeneous blanks, Single point incremental forming process (SPIF) is found to be the excellent solution. However, the combination of SPIF and TWBs technology is not much explored by the researchers. In this present work, simulation of SPIF process for TWBs has been attempted. Effect of different tool initial position on the responses like weld line shift, Plastic Equivalent strain (PEEQ) is investigated. It has been found that the tool dragging effect and deformation of weak material is responsible for the weld line shift during the SPIF process. Particularly for SPIF process, nature and magnitude of weld line shift is affected by the tool initial position.

Keywords Tailor welded blanks (TWBs) · Weld line shift · Formability · Finite element analysis

S. Marathe (✉) · H. Raval
Department of Mechanical Engineering, Sardar Vallabhbhai National Institute of Technology, Surat 395007, India
e-mail: shalin.marathe89@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_5

5.1 Introduction

The automobile industries are aiming at vehicles with more fuel economy along with reduction in the cost of manufacturing the automobile body parts. Reduction in weight without compromising the strength of the component is possible by adopting technology like TWBs. TWB is a product of joining of two or more than two material having different properties which are subjected to forming operations. The joining of material with different densities will lead to the reduction in the overall weight of the component. A schematic representation of TWB used in inner door panel of a car body is shown in Fig. 5.1.

Automobile, aerospace, transportation, and shipbuilding industries are aiming at weight reduction of the components without compromising the strength and TWBs have been found to fulfill these requirements. TWB is a product of joining two or more than two materials prior to any forming operation. TWB involves joining of two or more than two materials having different properties, thicknesses and/or coatings by any of the conventional joining processes. As it involves joining of materials having different densities the overall weight of the product decreases [1]. Also, TWBs are subjected to forming after the joining has been done and because of that the number of dies required for the forming decreases [1] which also leads to decrease the product lead time as well. Figure 5.1 indicates the typical example of TWB.

In conventional technologies, forming of two different materials was done separately and then joining was performed. In the case of TWBs, joining is followed by forming of welded blank using a single die setup and that leads to the manufacturing cost reduction [1, 2]. The joining method adopted for development of TWB depends upon the parent material involved in it. As the application of aluminum instead of steel leads to the 20% reduction of weight [3], use of aluminum is increasing in the automobile applications. For welding of aluminum, friction stir welding process is found to be producing sound and defect-free welds [4].

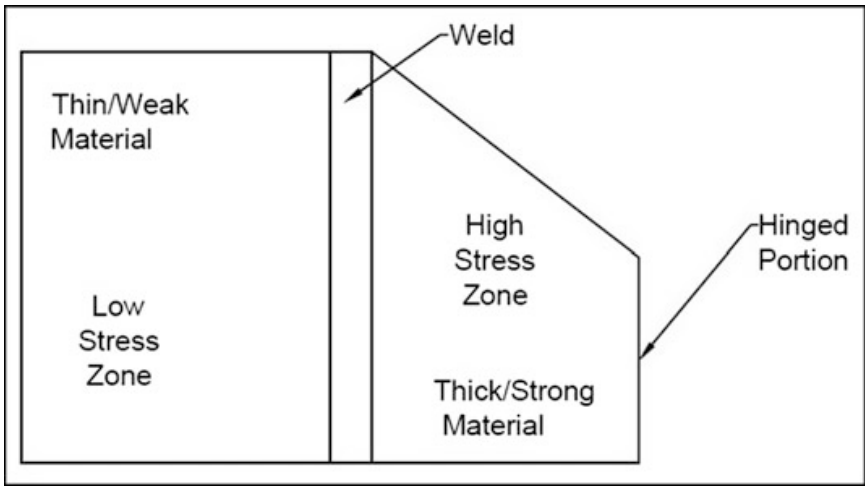


Fig. 5.1 Schematic representation of TWB used in inner door panel of a car body

Despite of having such tremendous advantages, the application of TWBs is limited as they are associated with the limitation of formability reduction [5, 6] and weld line shift [7, 8] during the forming process. In order to improve the forming behavior of TWBs during forming process, many efforts have been made. Such effort was made by Suresh et al. [8] in which the selective heating method was used to heat the strong/thick material to reduce the weld line shift during the stamping process. During the conventional forming process like deep drawing, the weld line shift is observed towards thick/strong material side [5] and formability of TWBs is found to be decreasing [5] than the same of parent material participating in it. However, for homogeneous blank, SPIF results in the better formability compared to the conventional forming processes [9]. So, the combination of SPIF and TWBs technology should be explored which is done by very few researchers.

So, in this present simulation study, FE analysis of SPIF process is carried out for the TWBs developed from the friction stir welding process. The effect of different initial position of tool on the responses like weld line shift, Plastic Equivalent Strain (PEEQ) and formed geometry is investigated.

5.2 Simulation Strategy

5.2.1 Challenges in Simulation of TWBs

For the simulation of TWBs, the major challenge is the consideration of the weld metal properties. During simulation, if the properties of the weld metal are not considered then results of simulation and experimental work are found to be deviating from each other [6, 7, 10]. In order to consider the weld metal properties, testing of welded joint is required to be performed and from that, different properties are extracted and used during the simulation. Rule of Mixture (ROM) can also be adopted to determine the weld metal properties [11]. During simulation of TWBs, if the appropriate properties of the weld metal are considered than actual behavior of the TWBs during that particular process can be predicted [12–14]. In order to consider the weld metal properties during simulation, an approach called "Weld Zone" is used [15]. In this approach, a different zone for weld metal is modeled and separate properties are assigned to it (see Fig. 5.2). In present simulation study, weld zone approach is used during the simulation in order to define the weld zone properties.

FSW results into the different zones in the welded blanks like WN (Weld Nugget), HAZ (Heat Affected Zone), TMAZ (Thermo Mechanically Affected Zone) and Base Metal. In the present simulation WN, HAZ and TMAZ are combined and considered as Weld Metal/Weld Nugget/Weld Zone. The properties of all

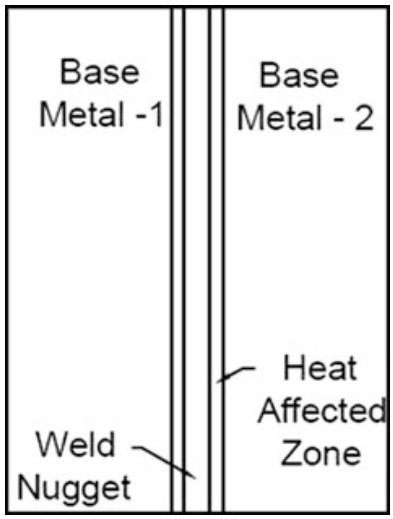


Fig. 5.2 Schematic representation of weld and heat affected zones

the mentioned welding zones can be considered during simulations but for that mechanical testing is required. However, consideration of such properties will lead to more accurate results.

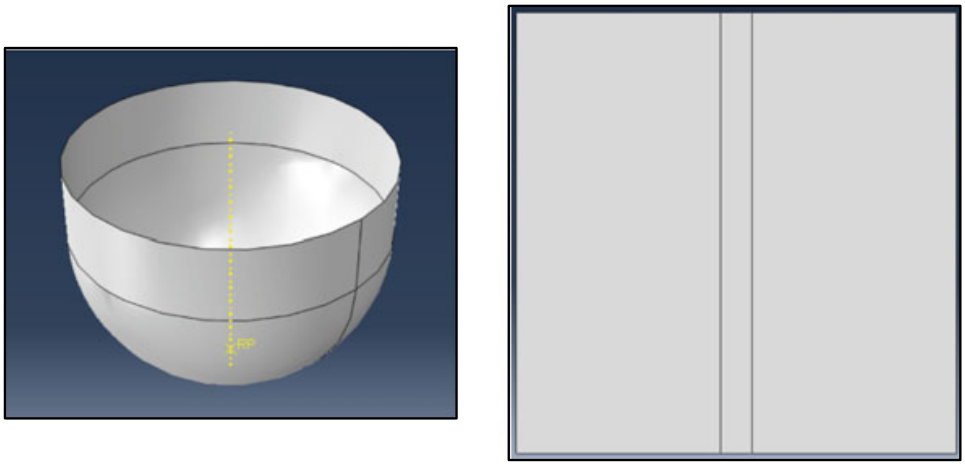
5.2.2 Simulation Methodology

In this work, ABAQUS/Explicit is used for the simulation purpose. In the modeling module, two instances are modeled namely blank and tool. In this present work, it is assumed that two blanks are welded using FSW process in order to develop a TWB. TWB is having dimensions of 70 × 70 mm^2 and 2 mm of thickness. The width of weld zone is assumed to be of 5 mm. Particularly, for FSW, the range of welding zones depends upon the diameter of the FSW tool shoulder. However, this present study is focused on the forming parameters and its effect of the weld line movement. In order to weld the plate of 2 mm thickness, FSW tool having 2 mm pin diameter is required. Considering that, shoulder diameter will be in the range of 4–6 mm [16]. So, the HAZ produced after welding will be in the same range. So, in present simulation study, 5 mm width of weld zone/material is considered.

The tool is considered as analytical rigid body and the blank is considered as shell type deformable body (see Fig. 5.3). The blank material is considered as isotropic in nature for the ease of simulation. However anisotropic properties can be considered and accuracy of results can be further improved.

The material properties used during simulations are represented in Table 5.1. The properties of the weld metal were assumed. The TWB is considered to be fabricated from AA 5754 H22 and AA 5052 H32 using the FSW process. Now onwards AA 5754 H22 and AA 5052 H32 is considered as weak base metal/weak material and strong base metal/strong material respectively.

The material having high value of yield strength is considered as strong material and material having low value of yield strength is considered as weak material.



(a) Hemispherical tool (rigid body, 6 mm radius)

(b) TWB – shell deformable body (70*70 mm^2,2 mm thick)

Fig. 5.3 Instances modeled in ABAQUS/explicit

Table 5.1 Properties used during simulation [5]

Property	AA 5754 H22	AA 5052 H32	Weld
Density[a] (ρ) (kg/m^3)	2670	2680	2670
Young's modulus[a] G (N/m^2)	70.3×10^9	70.3×10^9	70.3×10^9
Yield stress (σ_y) (N/m^2)	122×10^6	140×10^6	131×10^6
Poisson's ratio[a] (μ)	0.33	0.33	0.33
Strength coefficient (K) (N/m^2)	252.04×10^6	258.7×10^6	255.37×10^6
Strain Hardening index (n)	0.112	0.125	0.1185

[a]Indicates assumed properties

During simulation, material of TWB is assumed to follow the Power-Law equation as indicated below.

$$\sigma = K\varepsilon^n \tag{1}$$

In Eq. 1, σ and ε indicates true stress and true strain respectively. After the modeling, assembly of the SPIF tool and the blank is done. After assembly, interaction between SPIF tool and blank is provided in terms of friction. A surface to surface contact is defined between tool and blank and friction coefficient of 0.1 [17] is defined in order to consider the effect of friction (see Fig. 5.4).

The flat TWB is deformed using SPIF process and converted on to the target geometry frustum of pyramid. The dimensions of the pyramid are presented in Table 5.2.

During the simulation, the edges of the blank were fixed by defining encastre boundary (see Fig. 5.5) condition. The encastre boundary condition ensures pure stretching of the blank during forming operation and it eliminates the chance of drawing of material.

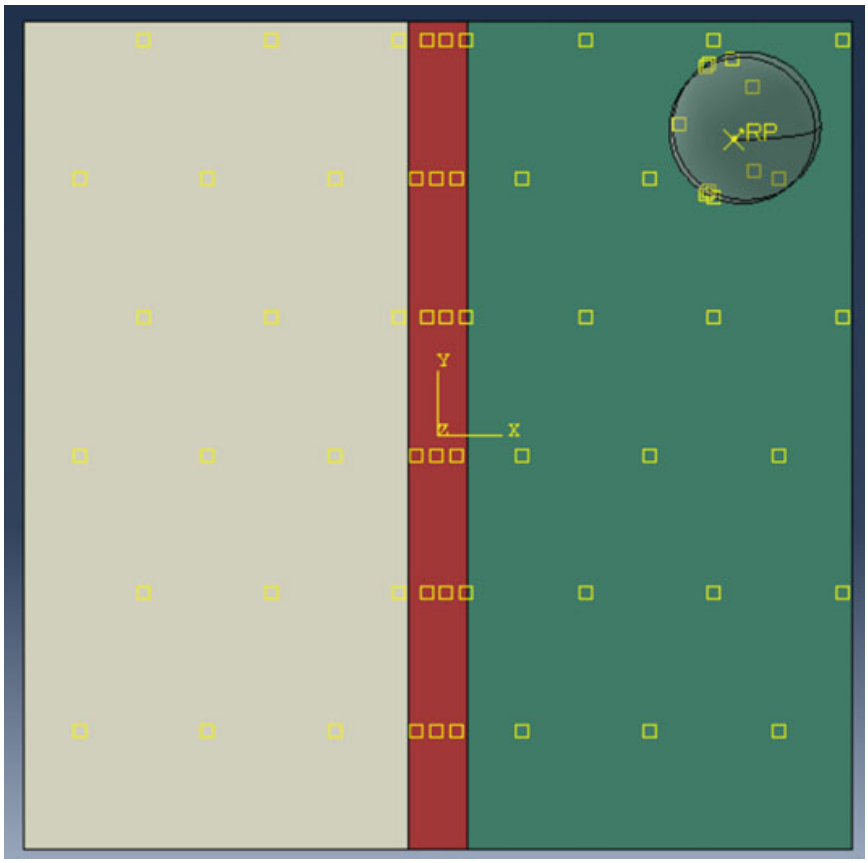


Fig. 5.4 Assembly of tool and blank and interaction between tool and blank

Table 5.2 Dimensions of the target geometry/frustum of pyramid

S. No.	Particular	Unit	Magnitude
1	Wall angle	Degree	45
2	Height of the pyramid	mm	20
3	Top dimensions	mm^2	50 × 50
4	Bottom/base dimensions	mm^2	10 × 10

The tool was given boundary condition in terms of movement in order to produce the target geometry (see Fig. 5.5). Tool starts the movement and goes through the path 1–2–3–4 and then it deforms the blank in the direction of thickness (perpendicular to plane of this paper) by 0.5 mm in step 5. The magnitude of movement of tool in step 5 depends upon the wall angle of target geometry. This cycle is repeated till the desired geometry is produced. Total 205 different boundary conditions were defined along with total 43 different amplitudes for the tool in order to produce the target geometry. The values of SPIF parameters adopted during the simulation are represented in Table 5.3.

After successful implementation of boundary conditions, meshing of blank is done with the help of S4R shell elements by keeping mesh size of 0.5 mm. For the present simulation study, mesh sensitivity analysis was carried out and based on

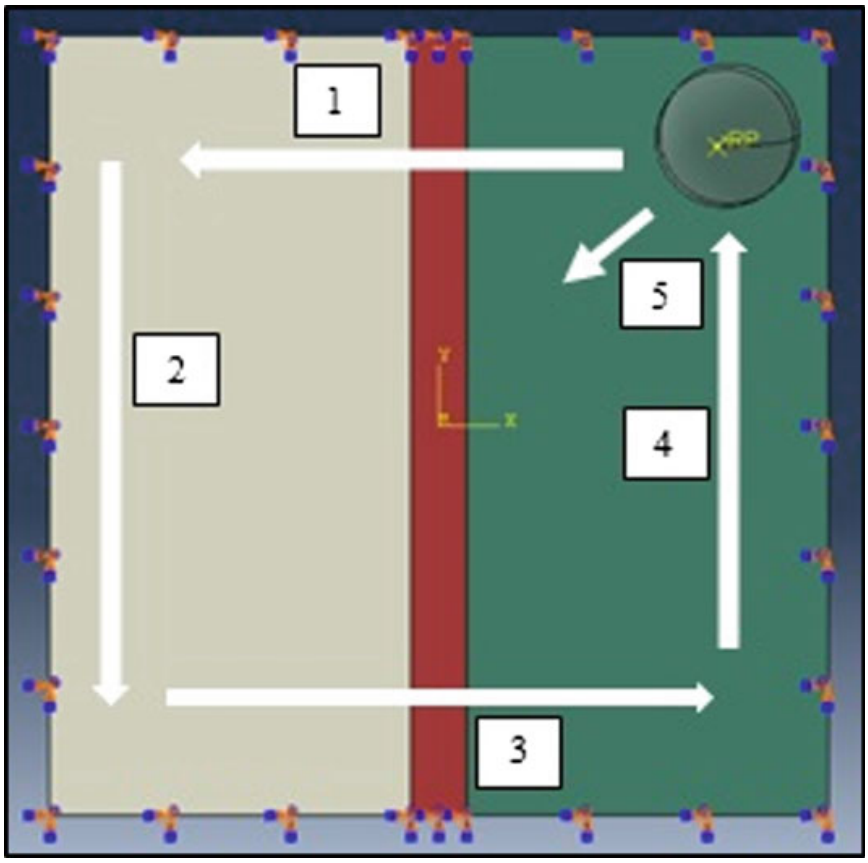


Fig. 5.5 Encastre boundary condition to the edges of the blank and direction of tool movement

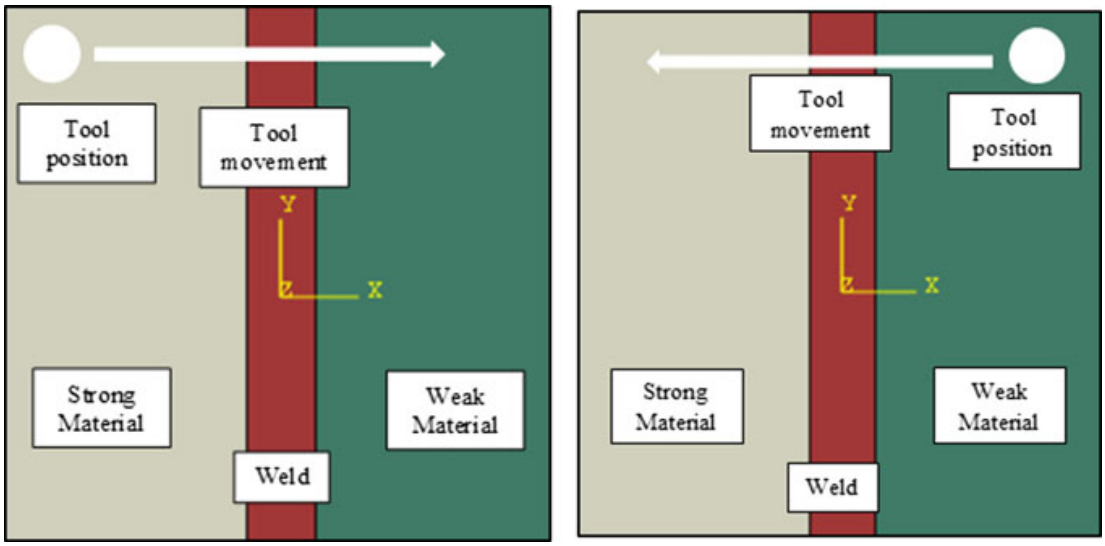
Table 5.3 Constant SPIF process parameters for present work

S. No.	Particular	Unit	Magnitude
1	Tool feed	mm/sec	10
2	Tool rotation	RPM	0
3	Tool path	–	Inward tool path
4	Tool radius compensation	mm	Not considered
5	Incremental depth	mm	0.5

computational time for one complete simulation and considering the available computational resources 0.5 of mesh size was selected. The meshing process is followed by the submission of job for the analysis.

The main aim of the present simulation study is to check the effect of different tool starting position on weld line shift during forming. In order to study this, two starting position of tools are considered as indicated in Fig. 5.6.

If the tool is starting its movement from strong material it will be considered as SFS condition (see Fig. 5.6a) and if it is starting from weak material it will be



(a) Starting From Strong (SFS) material

(b) Starting From Weak material (SFW)

Fig. 5.6 Two different starting position of SPIF tool

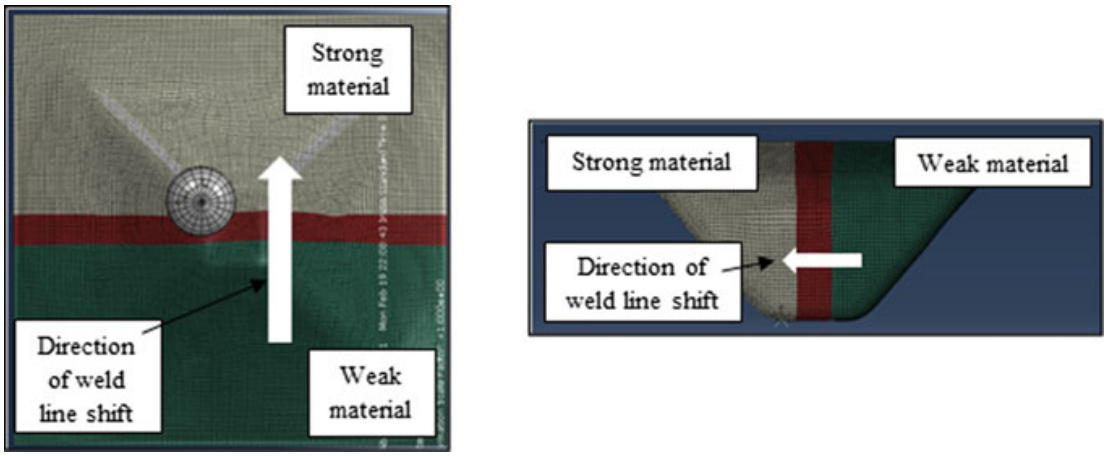
considered as SFW condition (see Fig. 5.6b). In the next sections, the effect of tool starting position on different responses is discussed in detail.

5.3 Results and Discussion

After the successful completion of simulations, results like weld line shift, PEEQ and formed geometry are studied and discussed below.

5.3.1 Formed Component

The formed component is analyzed for the weld line shift. Figure 5.7a, b indicates the top view and side view of the component formed using the SPIF process respectively. It can be observed from Fig. 5.7 that the weld line is shifted towards strong material by marginal amount. In order to investigate the weld line shift in detail, a node path was generated on the weld line and displacement of each node was recorded.



(a) Top view (b) Side view

Fig. 5.7 Formed component—frustum of pyramid

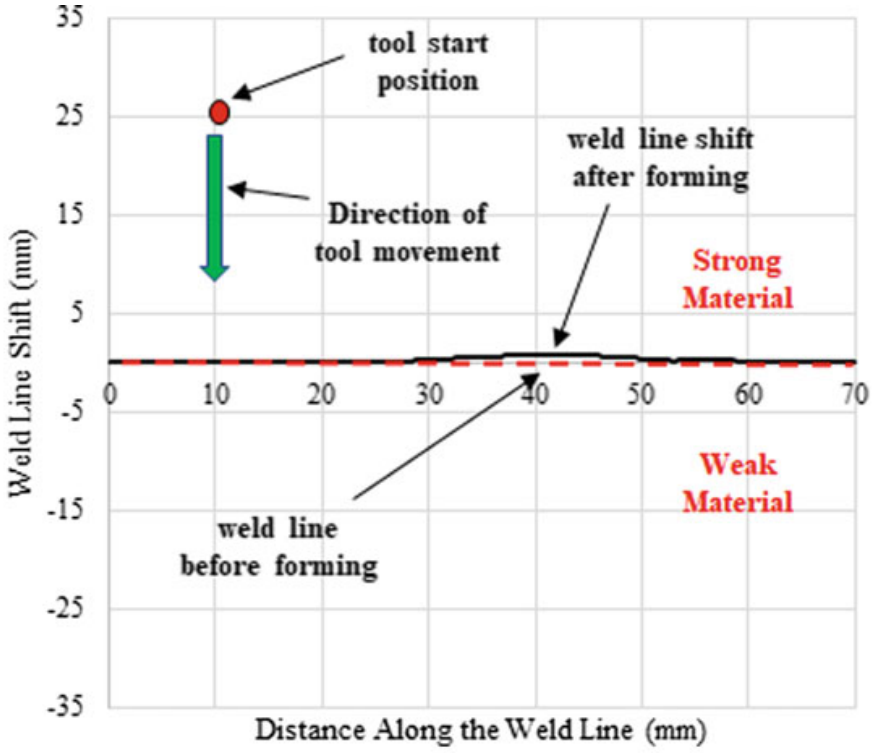
5.3.2 Weld Line Shift

During any forming process, same amount of deformation is applied by tool on strong, weak and weld material. As the deformation resistance capacity of the weak material is less compared to strong material, more deformation is experienced by the weak material compared to strong material. Due to this, weld line shifts towards strong material during any forming operation.

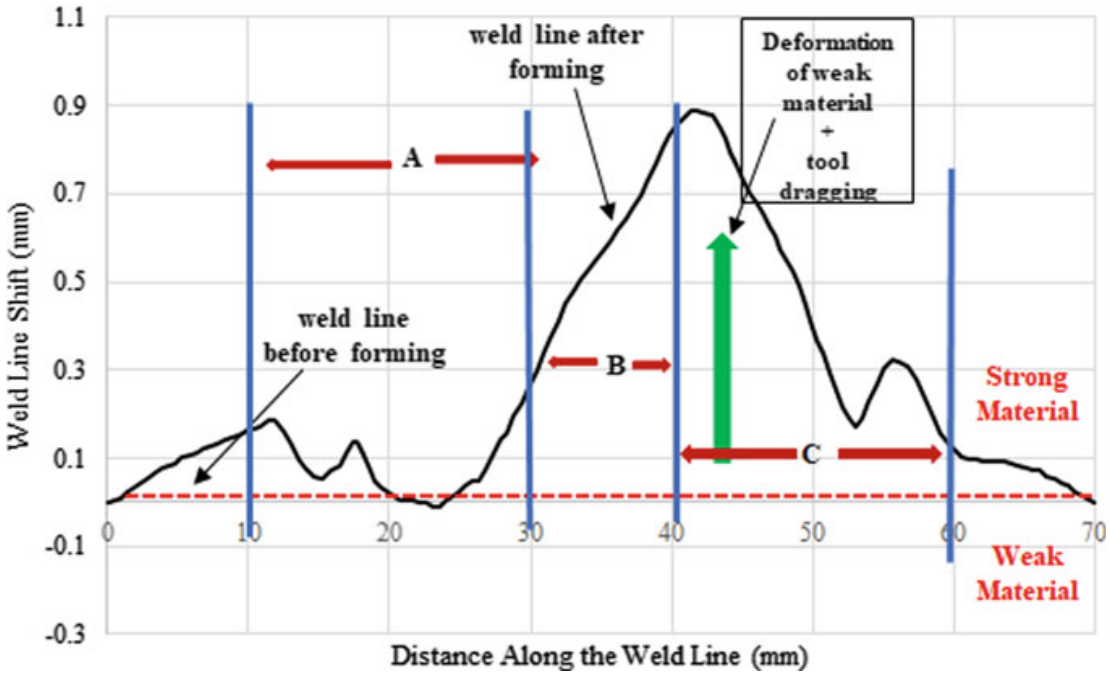
As the SPIF process involves local deformation of material due to tool movement, the weld line shift depends upon the initial position of tool. In order to understand the mechanism of weld line shift particularly for SPIF process, consider Fig. 5.8.

Figure 5.8a represents the overall view of the blank and position of weld line shift before and after forming process for SFS case. Figure 5.8b represent the magnified view with different sections. Section B can be called as untouched section, as there is no interaction between tool and blank in it. Section A and C represent the weld line shift in the wall of the pyramid. It should be observed that the weld line shift in section A is less compared to the same in section C. In SFS case, the strong material is above the mean line and weak material below the mean line. When the tool comes from section B to section C, deformation of weak material is supported by the tool dragging effect. Due to which more weld line shift is observed in section C. Now, in section A, deformation of weak material is opposed by the tool dragging and because of that the amount of weld line shift observed in section A is less. The weld line shift observed in section B is only due to section A and C.

Figure 5.9 represents the magnified view of the weld line shift observed for the case SFW. The sections represented in the figure are same as Fig. 5.8. The nature and magnitude of weld line shift of SFW case are different from the same of SFS case. It clearly indicates that the magnitude and nature of weld line shift depends



(a) Weld line shift overall view



(b) Magnified view of weld line shift

Fig. 5.8 Weld line shift (max. 0.88 towards strong side) for case SFS (Starting From Strong)

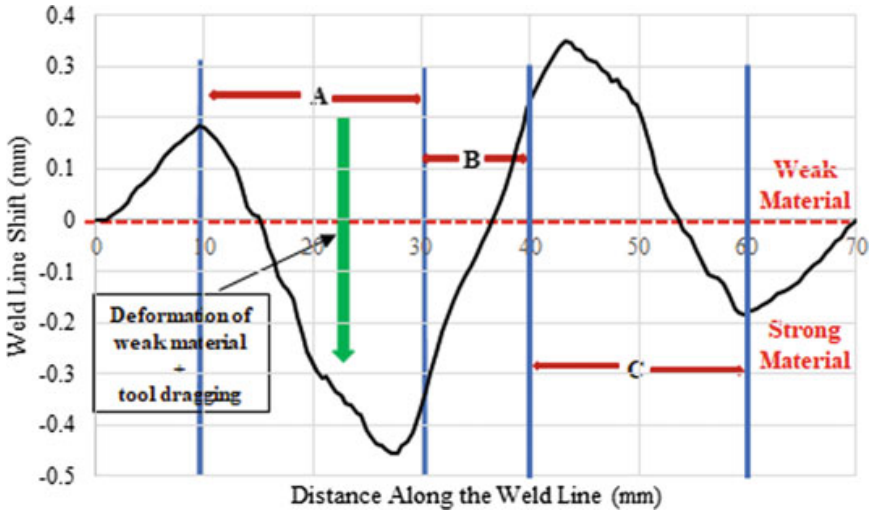


Fig. 5.9 Weld line shift (max. 0.45 towards strong side) for case SFW (starting from weak)—magnified view

upon on the initial position of the tool particularly for the SPIF process. In Fig. 5.9, on comparison of section A and C, it can be observed that the magnitude of weld line shift in section A is higher than the same of section C. For SFW case, the weak material is above the mean line and strong material is below the mean line. In section A, weld line shift is supported by the tool dragging effect and in section C same is opposed by the tool dragging, due to that, section A is experiencing more weld line shift.

In SPIF process, the deformation of the blank is localized unlike other conventional forming processes. Due to this localized deformation mechanism, the weld line shift of TWB during the SPIF process depends upon the initial position of the forming tool.

5.3.3 Plastic Equivalent Strain (PEEQ)

PEEQ result represents the permanent strain present in the body after the forming has been completed. Figure 5.10a, b represent the results of PEEQ for the case SFS and SFW, respectively. These results are based on the von-mises stress criteria.

For the SFS case, more plastic strain is observed in weak material compared to strong material. Due to this, for SFS case, weld line shift is observed only on one side (see Fig. 5.8). More strain leads to excessive thinning and that may lead to the failure of the blank during the forming. Due to in-homogeneity near the weld material, more value of PEEQ is observed there as well.

For the SFW case, value of PEEQ on strong and weak side is observed to have nearly same value. In Fig. 5.10b, two sections are highlighted which are responsible for the weld line shift in section A and C. Due to that, the nature of weld line shift is found to be sinusoidal in nature for SFW case (see Fig. 5.9).

(a) SFS – Max. 4.96 (b) SFW – Max. 6.58

Fig. 5.10 PEEQ in the formed component

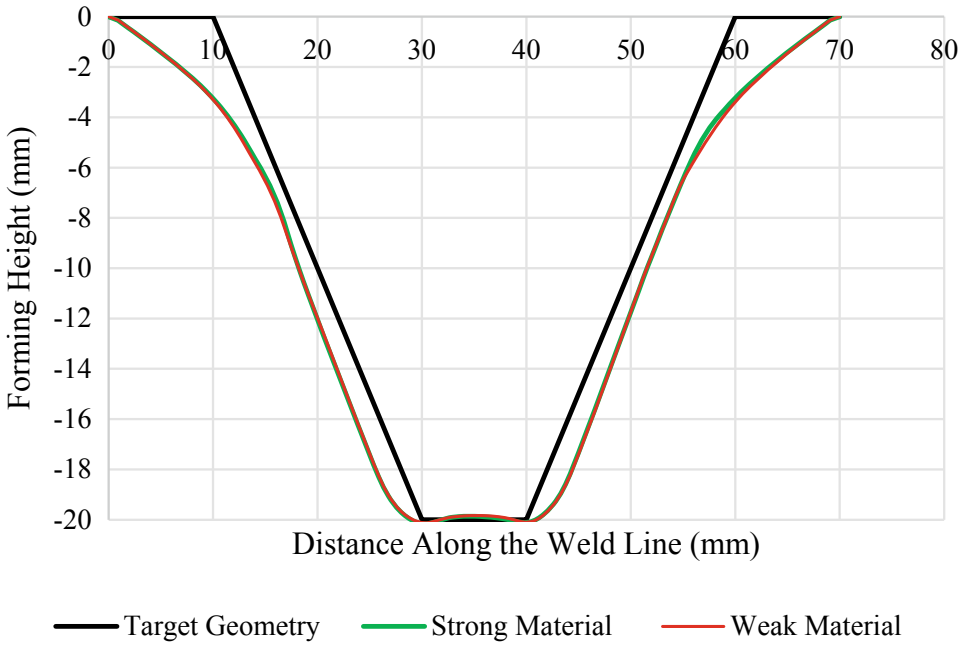
5.3.4 Comparison with the Target Geometry

The formed geometry is compared with the target geometry for SFS and SFW case in Fig. 5.11a, b, respectively. It can be observed that the formed geometry is deviating from the target geometry. In order to resolve this effect, proper tool radius compensation should be considered. For both SFS and SFW case, it should be noted that strong material is experiencing more spring back compared to the weak material because strong material is having high yield strength compared to weak material.

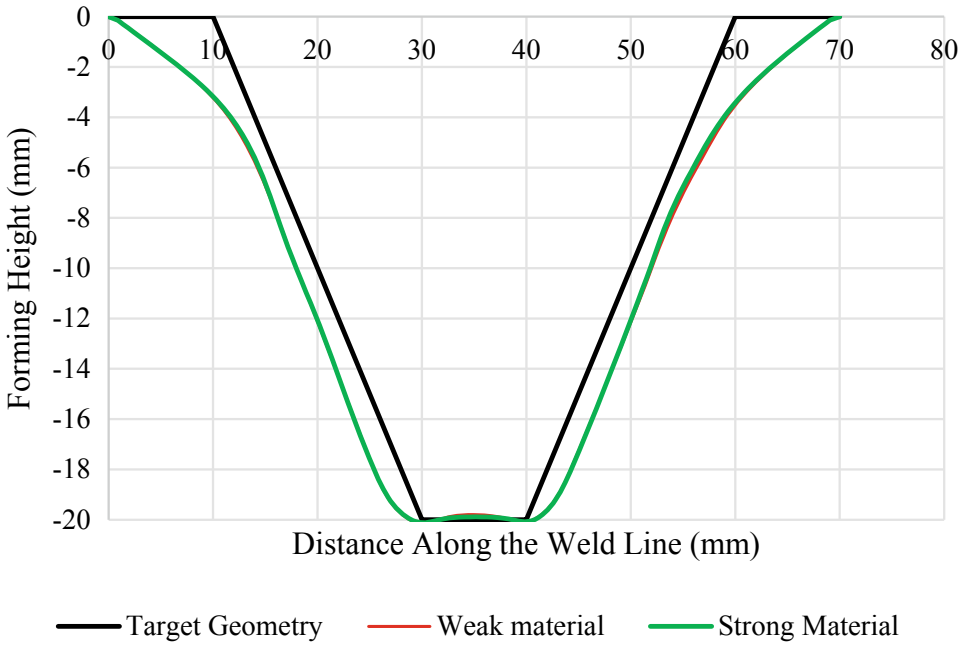
5.4 Conclusions

In the present research work, FE analysis of SPIF process is attempted for the TWBs. The effect of two different initial positions of tool on weld line shift, PEEQ and target geometry has been investigated. From the reported simulation study, the following points can be concluded:

- FE analysis can be adopted to study the forming behavior of the TWBs during SPIF process.
- Other than process parameters, weld line shift is affected by the initial position of the tool for the SPIF process. Initial position of tool affects the nature and magnitude of the weld line shift during the SPIF process.
- For the presented simulation conditions SFW results into the less weld line shift compared to the SFS case.



(a) For SFS case



(b) For SFW case

Fig. 5.11 Comparison with target geometry

- More weld line shift is observed in the wall section of the pyramid and that is due to the high amount of strain produced in it.
- If the deformation of weak material is supported by the tool dragging during the forming process, it leads to the maximum weld line shift.

- Formed geometry is nearly matching with the target geometry. However, the gap between them can be reduced by adopting proper tool radius compensation during forming.

References

1. Parente, M., Safdarian, R., Santos, A., Loureiro, A., Vilaca, P., Jorge, R.: A study on the formability of aluminum tailor welded blanks produced by friction stir welding. Int. J. Adv. Manuf. Technol. **83**. https://doi.org/10.1007/s00170-015-7950-0
2. Riahi, M., Amini, A.: Effect of different combinations of tailor-welded blank coupled with change in weld location on mechanical properties by laser welding. Int. J. Adv. Manuf. Technol. **67**, 1–9 (2013). https://doi.org/10.1007/s00170-012-4620-3
3. Shibata, K., Iwase, T., Sakamoto, H., Kasukawa, M., Chiba, K., Saeki, H.: Welding of aluminium tailored blanks by Nd: YAG lasers. Weld. Int. **17**(4), 282–286 (2003)
4. Shalin, M., Hiten, M.: Experimental analysis on effect of tool transverse feed, tool rotational speed and tool pin profile type on weld tensile strength of friction stir welded joint of AA 6061. Mater. Today: Proc. **5**(1), 487–493 (2018)
5. Kesharwani, R., Panda, S., Pal, S.: Experimental investigations on formability of aluminum tailor friction stir welded blanks in deep drawing process. J. Mater. Eng. Perf. **24**(2), 1038 (2015). https://doi.org/10.1007/s11665-014-1361-5
6. Panda, S.K., Kumar, D.R.: Improvement in formability of tailor welded blanks by application of counter pressure in biaxial stretch forming. J. Mater. Process. Technol. **204**(1), 70–79 (2008). https://doi.org/10.1016/j.jmatprotec.2007.10.076
7. Panda, S.K., Baltazar Hernandez, V.H., Kuntz, M.L., Zhou, Y.: Formability analysis of diode-laser-welded tailored blanks of advanced high-strength steel sheets. Metall. Mater. Trans. A **40**(8), 1955–1967 (2009). https://doi.org/10.1007/s11661-009-9875-4
8. Suresh, V.S., Regalla, S.P., Gupta, A.K.: Combined effect of thickness ratio and selective heating on weld line movement in stamped tailor-welded blanks. Mater. Manuf. Process. **32** (12), 1363–1367 (2017). https://doi.org/10.1080/10426914.2016.1257128
9. Micari, F., Ambrogio, G., Filice, L.: Shape and dimensional accuracy in single point incremental forming: state of the art and future trends. J. Mater. Process. Technol. **191**(1–3), 390–395 (2007). https://doi.org/10.1016/j.jmatprotec.2007.03.066
10. Reis, A., Teixeira, P., Duarte, J.F., Santos, A., da Rocha, A.B., Fernandes, A.A.: Tailored welded blanks—an experimental and numerical study in sheet metal forming on the effect of welding. Comput. Struct. **82**(17), 1435–1442 (2004). https://doi.org/10.1016/j.compstruc.2004.03.039
11. Abdullah, K., Wild, P.M., Jeswiet, J.J., Ghasempoor, A.: Tensile testing for weld deformation properties in similar gage tailor welded blanks using the rule of mixtures. J. Mater. Process. Technol. **112**(1), 91–97 (2001). https://doi.org/10.1016/S0924-0136(01)00555-6
12. Song, Y., Hua, L., Chu, D., Lan, J.: Characterization of the inhomogeneous constitutive properties of laser welding beams by the micro-Vickers hardness test and the rule of mixture. Mater. Des. **37**, 19–27 (2012). https://doi.org/10.1016/j.matdes.2011.12.029
13. Jia, Q., Guo, W., Li, W., Zhu, Y., Peng, P., Zou, G.: Microstructure and tensile behavior of fiber laser-welded blanks of DP600 and DP980 steels. J. Mater. Process. Technol. **236**, 73–83 (2016). https://doi.org/10.1016/j.jmatprotec.2016.05.011
14. Chu, G.N., Liu, G., Liu, W.J., Yuan, S.J.: An approach to improve thickness uniformity within tailor-welded tube hydroforming. Int. J. Adv. Manuf. Technol. **60**(9–12), 1247–1253 (2012). https://doi.org/10.1007/s00170-011-3792-6

15. Zadpoor, A.A., Sinke, J., Benedictus, R.: Finite element modeling of transition zone in friction stir welded tailor-made blanks. AIP Conf. Proc. **908**(1), 1457–1462 (2007)
16. Elangovan, K., Balasubramanian, V.: Influences of tool pin profile and tool shoulder diameter on the formation of friction stir processing zone in AA6061 aluminium alloy. Mater. Des. **29** (2), 362–373 (2008). https://doi.org/10.1016/j.matdes.2007.01.030
17. Saidi, B., Boulila, A., Ayadi, M., Nasri, R.: Prediction of the friction coefficient of the incremental sheet forming SPIF. In: The 6th International Congress Design and Modelling of Mechanical Systems CMSM (2015)

Chapter 6
Joining Concentric Tubes by End Forming: A Finite Element Simulation Study

E. Premananda and R. Ganesh Narayanan

Abstract Concentric tubes are used in heat exchangers, chemical and bioreactors, air conditioners, etc. for transmitting fluid from one location to another. Butt joining of tubes is generally performed and this is accomplished by welding and mechanical fastening. The joining of concentric tubes is mostly performed by fusion welding processes. In the present work, an alternative approach to join two concentric tubes using end forming is proposed. A die groove is designed using finite element (FE) analyses to optimize the process parameters for successful joining. FE code ABAQUS/CAE is used for the FE simulations. The tubes are joined by end forming within the die groove. Such FE simulations helped in practical demonstration of the joining process as well.

Keywords Tube joining · End forming · Finite element simulation

6.1 Introduction

Tubes form an integral part of automotive industries, chemical factories, air conditioning system, railway structures, food processing sectors, construction and shipbuilding industries. Concentric tubes are used in heat exchangers, chemical and bioreactors, air conditioners, etc. for transmitting fluid from one location to another. They act as a structural member and hence they are joined and supported suitably to form a mechanical system. Butt joining of tubes is generally performed to take care of length requirements and this is accomplished by welding and mechanical fastening. The joining of concentric tubes is seen in certain applications, in which the ends are joined for sealing and supporting purposes. This is mostly done by

E. Premananda
Department of Production Engineering, VSSUT, Burla, Sambalpur 768018, India

R. Ganesh Narayanan (✉)
Department of Mechanical Engineering, IIT Guwahati, Guwahati 781039, India
e-mail: ganu@iitg.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_6

fusion welding processes viz., arc welding and flame welding. Green and sustainable manufacturing and fabrication of parts, as a theme and at a process level, plays a vital role in the development of new materials and processes. One can quote many examples for this. Such a development is meant for cleaner manufacturing and healthy environment. Some examples in which 'welding' is replaced by 'forming or plastic deformation' to accomplish the joining operation are discussed below.

Agrawal and Narayanan [1] proposed a method to join a steel to a thin sheet by end forming of tube. FE simulations were conducted to optimize the parameters. Some criteria were suggested to select the successful joints for fabrication and testing. Alves et al. [2] proposed a method of joining a tube to a sheet using end forming of tube. The forming component involves two localized bends in the tube between which the sheet is placed. FE simulations were conducted to check the possibility of the joining method. A practical example of joining a lever to a fulcrum in an automotive braking system has been demonstrated using the proposed method. A similar method of localized bending of tubes to join two tubes was presented by Alves et al. [3]. In this, joining of carbon steel welded tube (n = 0.06; k = 616.4 MPa) with Aluminium 6060T6 tube was demonstrated. The slenderness ratio and angle of inclination of chamfered edges are optimized to yield a successful product. Silva et al. [4] presented an environmental friendly joining of tubes involving localized bending of tubes. This is similar to tubular lap joints that have a significant overlay with the counterfacing surfaces of the coupling tubes to be joined. The proposal has been demonstrated experimentally and through numerical simulations. The joints performed well in tensile, bending, torsion, and water tightness destructive tests. A new method of fabricating gas and liquid storage reservoirs using nosing of a tubular preform has been demonstrated by Alves et al. [5]. The cases with and without mandrel have been compared. The case in which a mandrel is used for fabrication has been successful. Zhang et al. [6] had used rotary swaging technology for joining two tubes, in which a concave arc region has been formed to take care of the joint efficiency. The ratio of tensile load of the joined tubes to the maximum tensile load of the single tube can reach up to 68% for the inner tube and 47% for the outer tube, depicting the acceptable strength of the joint for practical applications. In separate work, Zhang et al. [7] demonstrated that copper and aluminium tubes can be joined by rotary swaging. The length of the overlapped region plays a vital role in deciding the joint strength. Many such examples can be quoted in the context.

From the discussion, it is imperative that the end forming of tubes is either used to fabricate tubular components or used to join tubes or to join tubes to sheets. By adopting these processes, usage of filler materials, shielding gases, flux in fusion welding processes are ruled out. Thus, joining by end forming can be categorized under green and sustainable joining processes.

In this context, the main objective of the present work is to propose and analyze a simple manufacturing method for joining of concentric tubes having dissimilar properties using end forming. An experimental setup having a die groove has been designed using FE simulations in Abaqus/CAE, FE code. The tubes undergo plastic deformation at their ends in the die groove and get mechanically interlocked to form a joint at the tube end. The influence of three important parameters is studied through FE simulations and suggestions on the sensitivity of those parameters are delivered in the article. Such parameters are optimized within chosen range for practical demonstration.

6.2 Proposed Joining Method

Figure 6.1 shows the proposed method of joining concentric tubes. The tubes of dissimilar qualities are located just above the die groove. The punch is given downward displacement such that the tubes move along the die groove and plastically deform through end curling. During tube end curling, the ends get mechanically interlocked to form a joint at the ends. The joining of concentric tubes by end forming is studied with the inner tube of length 70 mm, internal diameter 37.8 mm, and thickness 1.64 mm, and an outer tube of internal diameter 42.8 mm, and thickness 1.92 mm. The study is carried out to verify process feasibility as a function of gap between the tubes 'd', ratio of length of tubes 'h_1/h_2', and Coulomb's friction coefficient, μ. The input parameters are given in Table 6.1. The range for these parameters was selected after several FE simulation trials.

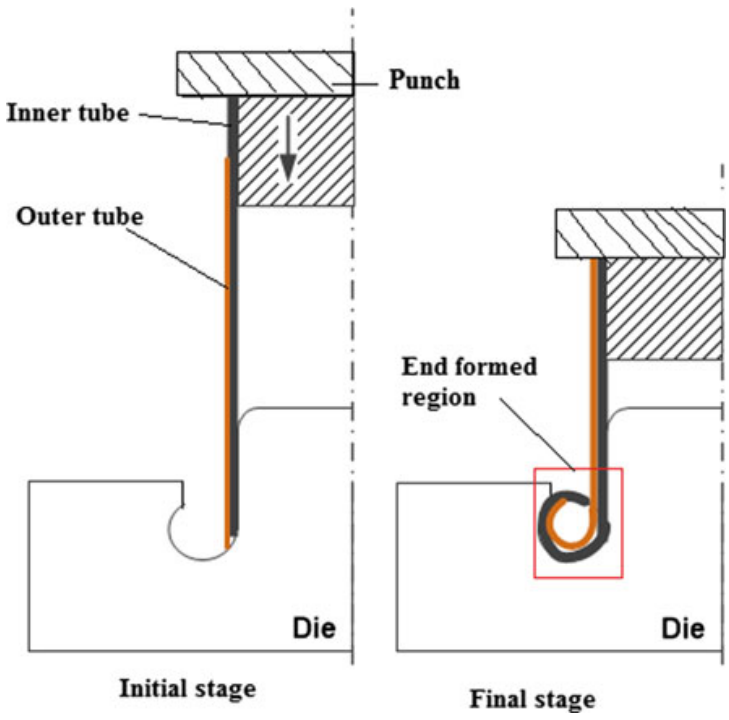


Fig. 6.1 Schematic of joining of concentric tubes by end forming

Table 6.1 List of varying input parameters for concentric tubes joining

Coulomb's friction coefficient (μ)	Ratio of length of tubes (h_1/h_2)	Gap between the tubes (d) in mm
0.05	70/44	0.10
0.08	70/46	0.50
0.10	70/48	0.86
0.15	70/50	1.00
0.20	70/52	1.50

h_1 = length of internal tube referred to as tube1
h_2 = length of external tube referred to as tube2

6.3 FE Simulation and Experiment Trial

The numerical simulation of joining of concentric tubes by end forming is carried out by using the FE code Abaqus/CAE, version 6.7. The selected mesh size for punch and tube is 4 mm and for die is 3.5 mm. The element type of die and punch (rigid bodies) are C3D4 and tube (deformable) is C3D8R. The element size has been chosen after sensitivity analyses and for minimum computational time. The objective of this study is to join the concentric tubes by end forming process. By using the simple die which is used in case of single tube, the locking of tubes did not take place. So a die having groove is designed to fabricate the joint. The entire process can be divided into three stages, i.e., stage 1, stage 2 and stage 3.

Stage 1: As both the tubes are of different length, i.e., internal tube is larger than the external tube, when axial compression takes place the internal tube starts to deform at the die surface. Since there is no contact between the punch and external tube, it rests at the bottom portion without having any deformation which is shown in Fig. 6.2.

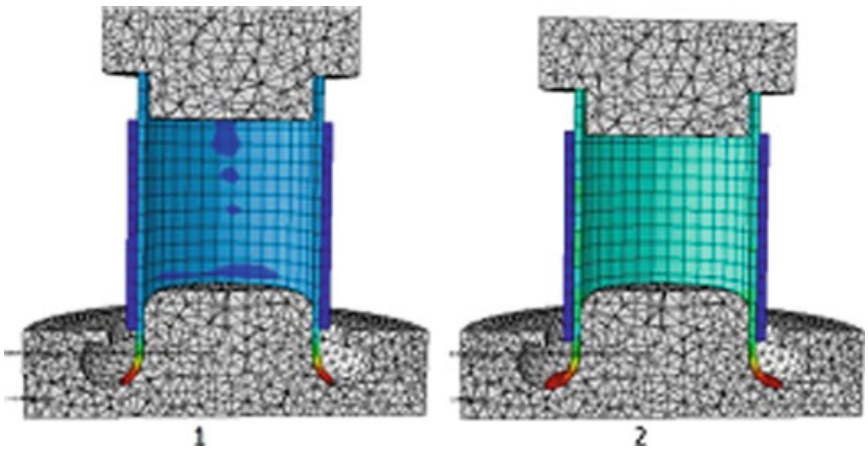


Fig. 6.2 Stage 1 of the proposed joining method

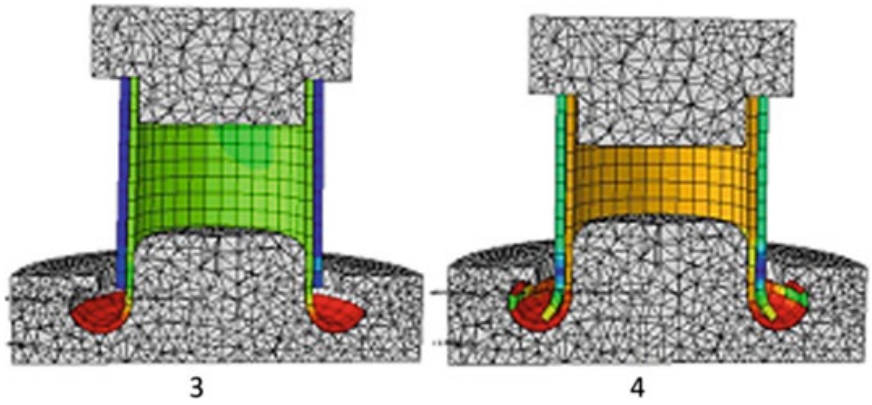


Fig. 6.3 Stage 2 of the proposed joining method

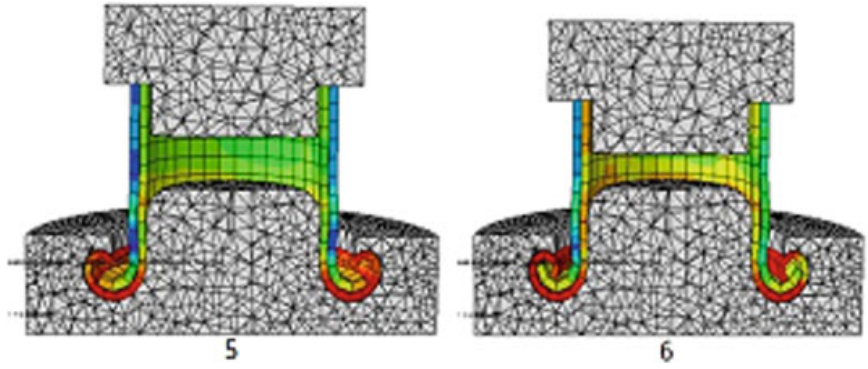


Fig. 6.4 Stage 3 of the proposed joining method

Stage 2: In this stage, the external tube comes in contact with the punch and it also starts deforming above the internal tube. By the time, the internal tube keeps on deforming and acquires the shape of groove of the die. The process of stage 2 is shown in Fig. 6.3.
Stage 3: As the upper portion of the die groove is heading inward, the leading edge of internal tube starts deforming in the inward direction. After a particular stage, the tubes get mechanically locked as shown in Fig. 6.4.

Some experimental trials were conducted to demonstrate the joining method at lab scale. Figure 6.5 shows the experimental setup clamped in a universal testing machine. The tubes were displaced in the downward direction at a uniform cross-head speed of 1 mm/min.

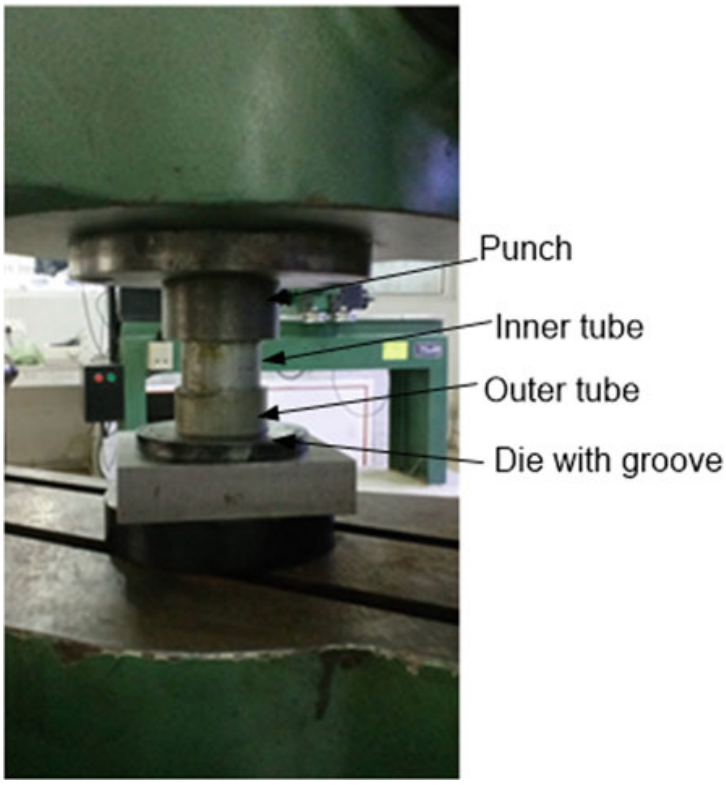


Fig. 6.5 Experimental setup for the proposed joining method

6.4 Material Properties Evaluation

The uniaxial tensile experimental tests were performed in order to study the mechanical properties of the tubes by means of universal testing machine. The specimen used as internal tube is mild steel tube of internal diameter 37.8 mm and thickness of 1.64 mm, the outer tube is of internal diameter 42.8 mm and thickness is 1.92 mm. The mechanical properties obtained for both the tubes from the tensile test are presented in Table 6.2. The internal tube is referred to as tube 1 and the outer tube is referred to as tube 2. The other material properties assumed are density = 7850 kg/m^3, Poisson's ratio = 0.303. These were obtained from tensile tests. The true stress-strain data which is used in the simulation process was fit to the Hollomon's power hardening law defined as, $\sigma = K\varepsilon^n$, where K and n are the strength coefficient and the strain hardening exponent, respectively. For FE simulations, the tubes are considered isotropic, and von-Mises yield function is followed.

Table 6.2 Mechanical properties of the tubes

Tube	Elastic modulus (GPa)	Yield stress (MPa)	Ultimate tensile strength (MPa)	Uniform elongation (%)	n	K (MPa)
Tube 1	84.6	339	406	12.3	0.15	646
Tube 2	106.3	305	342	9.9	0.14	526

6.5 Results and Discussion

Length ratio of both tubes plays a major role in the locking of both the tubes. When the simulation was carried out for tubes having equal length the external tube started to buckle in the mid-region when they were forced to deform along die groove region. Therefore, tubes having different length were introduced in order to get the desired locking of concentric tubes. The length of internal tube is kept fixed and the length of external tube was changed. Figure 6.6 shows the load-displacement curve for different length ratios. The curves have been fit to the data by a polynomial.

It can be seen that for higher and lower length ratios the process is unsuccessful. As the length of external tube increases, the load required for the joint formation also increases. The result of the process for different length ratios, whether it is successful or unsuccessful, is listed in Table 6.3. In case 1, the locking of tubes was not proper. The internal tube strikes the external tube before it moves inside. In cases 2, 3 and 4 the locking was good. In case 5, both the tubes start to buckle before joining.

In forming processes friction plays a key role. The friction between the internal tube and die is considered as frictionless and the friction between both the tubes is

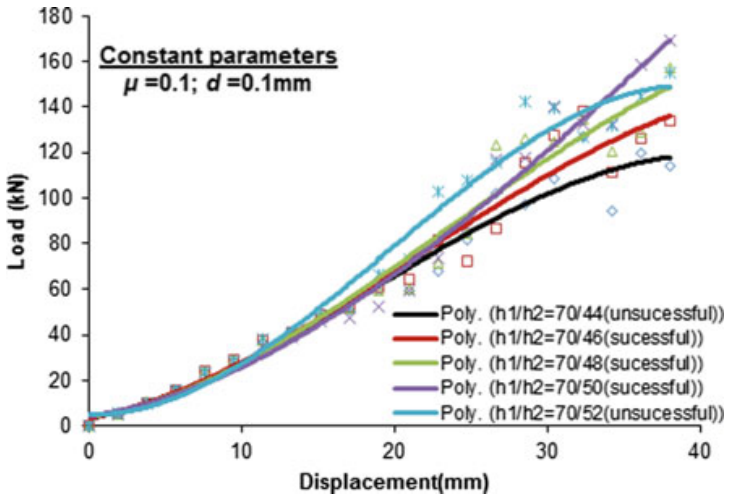


Fig. 6.6 Load evolution for various tubes length ratios

Table 6.3 Effect of length ratio of tubes in concentric tube joining

Cases	h_1/h_2	Results
Case 1	70/44	Unsuccessful
Case 2	70/46	Successful
Case 3	70/48	Successful
Case 4	70/50	Successful
Case 5	70/52	Unsuccessful

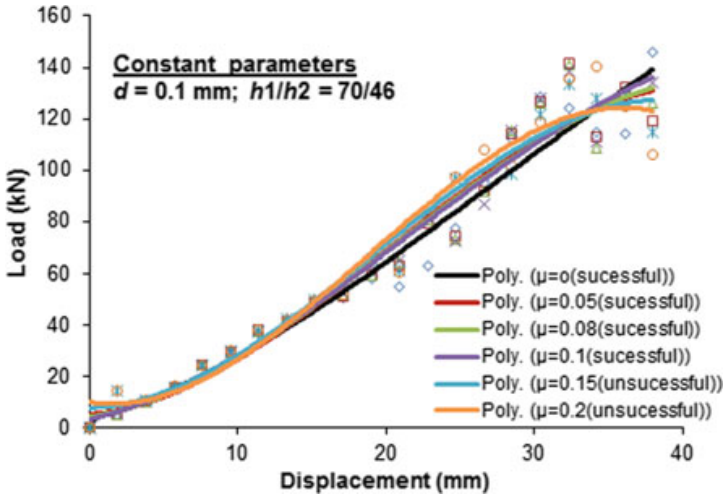


Fig. 6.7 Load evolution for various Coulomb's friction coefficient

Table 6.4 Effect of Coulomb's friction coefficient in concentric tube joining

Cases	μ	Results
Case 1	0	Successful
Case 2	0.05	Successful
Case 3	0.08	Successful
Case 4	0.10	Successful
Case 5	0.15	Unsuccessful
Case 6	0.20	Unsuccessful

varied to observe the effect. As the friction coefficient increases, the load required for the process also increases (Fig. 6.7). For very high value of friction coefficient, the process is unsuccessful. The results for different μ values are listed in Table 6.4. The locking of tubes was acceptable for cases 1–4. In case 5 and 6, tubes started to buckle from the upper portion.

The gap between the tubes was varied in order to evaluate its effect on the forming process. It is observed that for no gap or 0.1 mm gap the process is feasible. For all values other than 0.1 mm, the process is unsuccessful. For all the unsuccessful cases, the load drops when the tubes start buckling (Fig. 6.8). Moreover, in all the unsuccessful cases, the tube buckled when the external tube made contact with the internal tube at die groove region.

Some experimental attempts were made as shown in Fig. 6.9 to join concentric tubes. Though tubes underwent forming in the die groove as observed in simulations, the inner tubes failed before joining is fully accomplished. Such a failure pattern was not observed in FE simulations as instability criterion was not implemented.

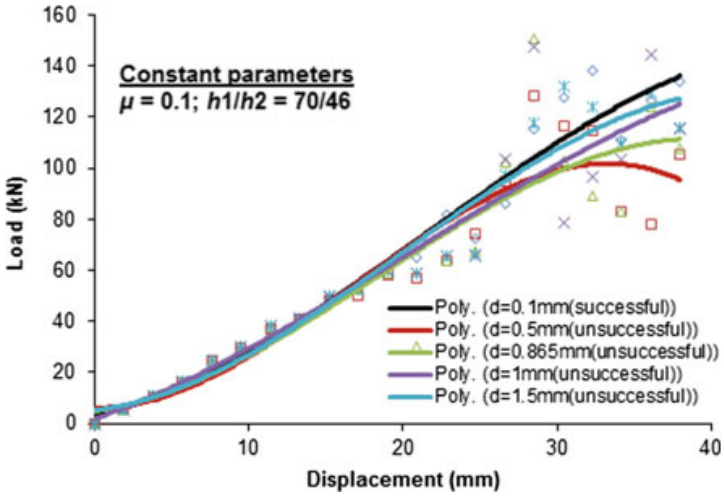


Fig. 6.8 Load evolution for different tube gaps

Fig. 6.9 Joints made by end forming

6.6 Conclusions

An attempt has been made to join concentric tubes by means of end forming operation. The input parameters such as length ratio of tubes, Coulomb's friction coefficient and gap between the tubes have significant effects on the feasibility of the process. The entire analyses have been carried out by FE simulations. The parameters are optimized and experimental demonstration is partially successful. During experimentation, the inner tubes failed due to circumferential stretching before joint formation occurred.

References

1. Agrawal, A.K., Narayanan, R.G.: Joining of a tube to a sheet through end curling. J. Mater. Process. Technol. **246**, 291–304 (2017)
2. Alves, L.M., Dias, E.J., Martins, P.A.F.: Joining sheet panels to thin-walled tubular profiles by tube end forming. J. Clean. Prod. **19**, 712–719 (2011)

3. Alves, L.M., Silva, C.M.A., Martins, P.A.F.: End-to-end joining of tubes by plastic instability. J. Mater. Process. Technol. **214**, 1954–1961 (2014)
4. Silva, C.M.A., Nielsen, C.V., Alves, L.M., Martins, P.A.F.: Environmentally friendly joining of tubes by their ends. J. Clean. Prod. **87**, 777–786 (2015)
5. Alves, L.M., Pardal, T.C.D., Martins, P.A.F.: Nosing thin-walled tubes into axisymmetric seamless reservoirs using recyclable mandrels. J. Clean. Prod. **18**, 1740–1749 (2010)
6. Zhang, Q., Lin, K., Mu, D.: Tube/tube joining technology by using rotary swaging forming method. J. Mater. Process. Technol. **214**, 2085–2094 (2014)
7. Zhang, Q., Zhang, Y., Cao, M., Ben, N., Ma, X., Ma, H.: Joining process for copper and aluminum tubes by rotary swaging method. Int. J. Adv. Manuf. Technol. **89**, 163–173 (2017)

Chapter 7
Optimization of the Forming Parameters in U-Bending for Punch Force and Springback Using Taguchi Method

P. Venkateshwar Reddy, D. Mohana Krishnudu, U. Pranavi and P. Janaki Ramulu

Abstract Springback is a crucial factor that influences the feature of sheet metal in the sheet metal forming (SMF). In SMF operations, springback of the component during unloading mostly determines whether the component confirms to the design dimensions and tolerances. The aim of the current work is to analyze the importance of forming parameters on the responses: punch force and springback in U-Bending of SS 304. Strip length, punch speed and lubricant with three levels each have been considered in the current work as the forming parameters. The effects of different process parameters on U-bending of sheet metal have been investigated by conducting experiments on SS 304. Experiments have been conducted as per Taguchi's L_9 orthogonal array. The optimum conditions have been determined based on their effect on punch force and springback of the sheet metal.

Keywords Springback · U-bending · Optimization · Taguchi · Forming

7.1 Introduction

The sheet metal forming method includes a mixture of elastic–plastic bending and stretch deformation of the sheet. These deformations might prompt a lot of springback of the formed sheet metal part. Springback represents a challenge for

P. Venkateshwar Reddy · D. Mohana Krishnudu
Department of Mechanical Engineering, G Pulla Reddy Engineering College, Kurnool 518007, AP, India

U. Pranavi
Department of Mechanical Engineering, Vardhaman College of Engineering, Hyderabad 501218, T.S, India

P. Janaki Ramulu (✉)
Program of Mechanical Design and Manufacturing Engineering & CoEAME, School of Mechanical, Chemical and Materials Engineering, Adama Science and Technology University, Adama, Ethiopia
e-mail: perumalla.janaki@astu.edu.et

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_7

manufacturers who desires to meet specific dimensions. It is wanted to anticipate and diminish springback with the goal that the last part dimensions can be controlled however much as could reasonably be expected. Significance of the springback calculation is noteworthy from the economical point [1]. Ayres [2] proposed the utilization of various advance procedure to diminish springback in stamping tasks. Liu [3] proposed to shift the holding pressure amid the forming procedure in the way of giving tensile pre-loading to lessen the springback in the formed component. Chou and Hung [4] considered the impact of process factors of the strategies utilized to lessen springback by an optimization procedure for U channel parts. Chan et al. [5] investigated on the springback effect on sheet metal bending by considering various die-punch parameters like punch angle, punch corner radius and die-shoulder radius by performing FEM analysis. Reche et al. [6] studied on air bending tests by considering the effect of bending angle, sheet thickness and loading conditions on spring back. A lot of efforts have been carried out to eliminate the springback, most of them mainly focused on the optimization technique by considering draw-beads, bend curvature radius, punch forces and blank holder force [7, 8]. Maia et al. [9] investigated the U-rail to compare various optimization techniques and prediction tools, with an aim of finding capable strategies for springback compensation. Kakandikar and Nandedkar [10] investigated the process parameters which influences the springback by using optimization technique on automotive component. Vasudevan et al. [11] investigated on the springback of electrogalvanised steel sheet by considering many process parameters and also coating of the materials along with its geometrical parameters and concluded that the punch speed, die opening along with sheet orientation has high influence on the springback. Buang et al. [12] investigated on the stainless steel sheet by considering the effect of the die and punch radii on the springback of V-die bending process and concluded that the springback decreased by decreasing the die and punch radii. Maske and Sawale [13] studied the optimized parameters on the springback effect on aluminum sheet considering the effect of temperature and gauge of the sheet and concluded that the temperature affects the springback phenomenon and their optimized condition is found out. Gautam et al. [14] compared both the numerical and experimental results of springback in 3 ply cladded sheet metal by considering orientation of the sheet metal and concluded that the specimens oriented along the transverse direction have a greater springback when compared to the perpendicular direction.

Only very few of the earlier works have focused on the optimization of forming parameters in sheet metal U-Bending of SS 304. Thus, the aim of the current work is to optimize the forming parameters so as to minimize the punch force and springback using Taguchi Method.

7.2 Materials and Method

7.2.1 Tool and Workpiece Material

Strips made up of SS 304 with thickness of 1 mm were used for the current study. Tables 7.1 and 7.2 shows the chemical composition and mechanical properties of SS 304 respectively. Tools such as punch and die are made of tool steel and their dimensions are shown in Table 7.3. The schematic illustration of tools used in the present study is shown in Fig. 7.1. The fabricated experimental setup is attached to the Universal Testing Machine (UTM) is as shown in Fig. 7.2. Figure 7.3 represents the deformed images of the sheet metal after measuring of springback angle.

7.2.2 Design of Experiments Using Taguchi Method

Taguchi method is used for various problems with multiple parameters due to its practicality and robustness [15, 16]. In Taguchi method, optimization refers to the purpose of best levels of control factors; these best levels are that which maximizes the Signal-to-Noise (S/N) ratios. In this method, the experiments are conducted based on Orthogonal Arrays to conclude the best levels. Strip length, punch speed, and lubricant are the three factors were considered for the experimental outline.

Table 7.1 Chemical composition of SS 304 in wt%

C	Mn	Si	Cr	Ni	P	S	N
0.08	2.0	1.0	18–20	8–10.05	0.045	0.03	0.1

Table 7.2 Mechanical properties of SS304 sheet metal [20]

Property	E (GPa)	v	σ (MPa)	K (MPa)	n	R_0	R_{45}	R_{90}
Value	210	0.3	420	1069	0.263	1.09	0.95	1.05

Table 7.3 U-bending tool dimensions

Tool	Geometry	Value
Punch	Width	35 mm
	Nose radius	4 mm
	Die width	35 mm
Draw die	Die opening length	37 mm
	Die corner radius	6 mm
Strip	Length	75, 95 and 115
	Width	35 mm

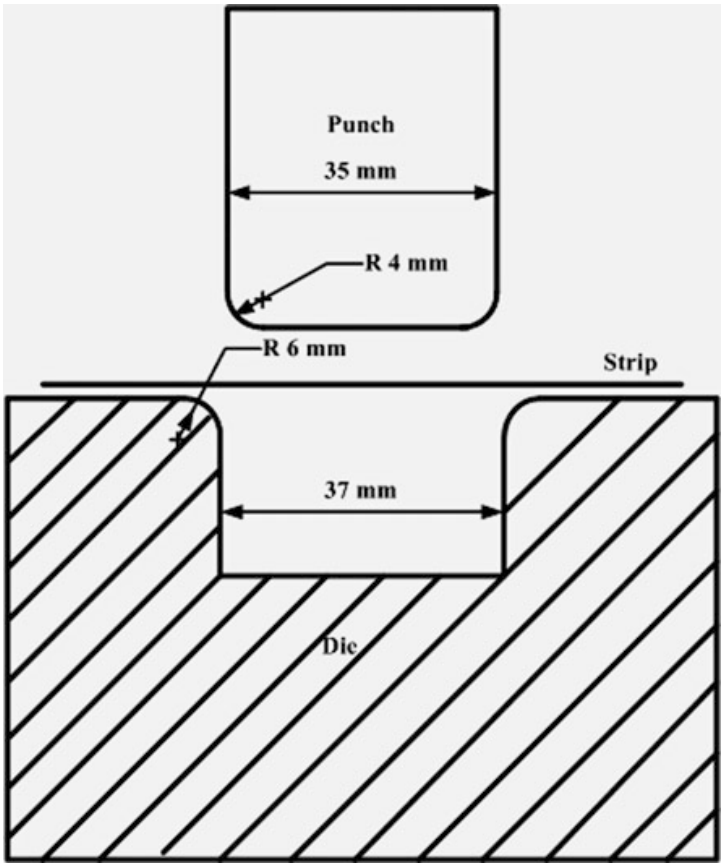


Fig. 7.1 Schematic illustration used in the study for U-bending

Fig. 7.2 Experimental setup for U-bending

Each factor has three levels. So L_9 Orthogonal Array (OA) was shaped in light of the three factors three-level design. The decision of a suitable OA relies upon the measure of control parameters and its levels. To appraise the principle parameter impacts, the design should be Orthogonal and Balanced. Taguchi has arranged some orthogonal arrays that fulfills the over two properties for various number of parameter-level combinations [17]. Inferable from cost and other test confinements, in the present case, a fractional factorial plan with nine runs has been chosen out of

Fig. 7.3 Deformed images after forming the sheet metal

Table 7.4 Considered parameters and their levels

Parameters	Level 1	Level 2	Level 3
Strip length (mm)	75	95	115
Punch speed (mm/min)	5	10	15
Lubricant	Engine oil	Mineral oil	Grease

the full factorial design, $3^3 = 27$ runs. The combination of factors and their levels are recorded in Table 7.4. The recorded methodology of the Taguchi appears in Fig. 7.4. However, this technique helps in deciding the feasible combinations of factors and recognizing the best combination. This, in turn, implies that the resources (materials and time) required for the experiments are also minimum.

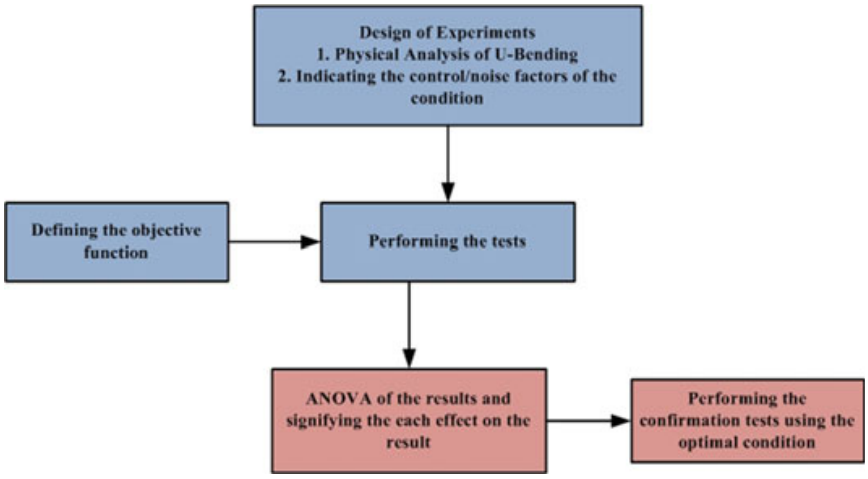


Fig. 7.4 Robust design flowchart for Taguchi method

7.3 Results and Discussion

The main aim of the present study is to optimize the forming parameters like strip length, punch speed and lubricant to achieve low punch force and springback. U-Bending test is analyzed to get a better combination of forming parameters using Taguchi technique for a different set of combinations. Table 7.5 shows the experimental test conditions and its responses.

7.3.1 Effect of Forming Parameters on Punch Force

The influence of forming parameters on Punch force is shown in Fig. 7.5. The punch force increases with increase in strip length since the punch has to travel more distance along with strip into the die. For both punch and die the contact area is increased as the length of the strip increased, thereby the punch force increased as the length of the strip increased. Whereas there is only a minute variation with punch speed and lubrication on punch force. Similar results were observed from the literature work carried out by the authors [18]. In their work lubricant does not affect much on the punch force as the same results were observed in the present study.

Table 7.6 represents the response for mean of means on punch force and it is observed that the most influencing factor on the punch force is the strip length followed by the punch speed and the lubricant. Since the punch travel is more for longer strip length with more contact time between the tools, so the punch force required is more for longer strips. Works carried out by different authors in the literature obtained almost the same results.

Table 7.5 Experimental test conditions and their response values

Test No.	Strip length	Punch speed	Lubricant	Punch force (kN)	Springback angle
1	75	5	Engine oil	10.5	4.6
2	75	10	Mineral oil	11.5	4.5
3	75	15	Grease	10.9	4.9
4	95	5	Mineral oil	17.5	5.1
5	95	10	Grease	16.9	5.5
6	95	15	Engine oil	18.2	5.2
7	115	5	Grease	21.6	5.4
8	115	10	Engine oil	20.6	5.1
9	115	15	Mineral oil	22.1	4.8

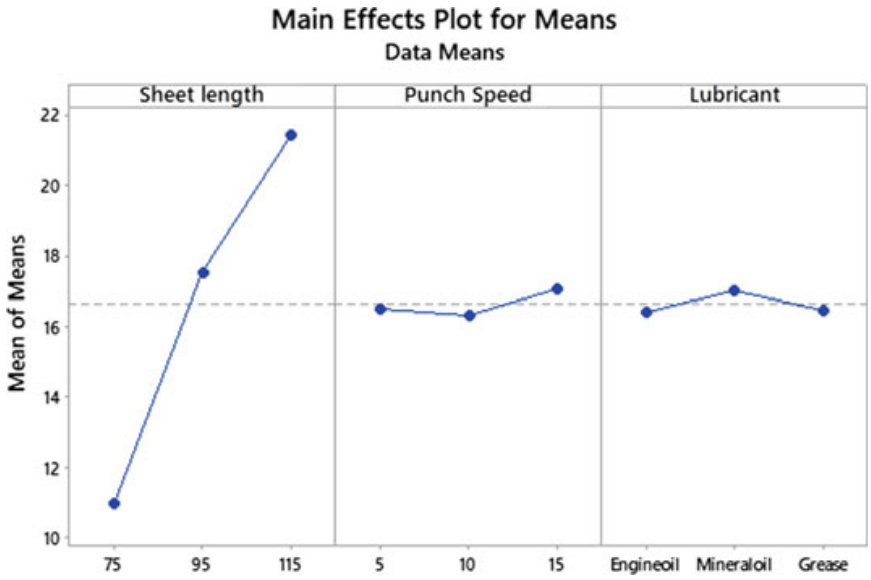


Fig. 7.5 Parameters affect plots on punch force

Table 7.6 Response for mean of means on cutting force

Level	Strip length	Punch speed	Lubricant
1	10.97	16.53	16.43
2	17.53	16.33	17.03
3	21.43	17.07	16.47
Delta	10.47	0.73	0.60
Rank	1	2	3

Table 7.7 represents the ANOVA results for punch force. It is observed that the strip length is the most influencing factor that affects the punch force and the percentage of contribution is about 99%. Meanwhile, the percentage contribution of punch speed and lubricant together is 1%.

Table 7.7 Table representing the ANOVA of punch force

Source	DF	Adj SS	Adj MS	*F*-value	*P*-value
Strip length	2	167.882	83.9411	172.09	0.006
Punch speed	2	0.862	0.4311	0.88	0.531
Lubricant	2	0.682	0.3411	0.70	0.588
Error	2	0.976	0.4878		
Total	8	170.402			
Model summary					
S	*R*-sq	*R*-sq(adj)	*R*-sq(pred)		
0.698411	99.43%	97.71%	88.41%		

7.3.2 *Effect of Forming Parameters on Springback*

The influence of forming parameters on springback is shown in Fig. 7.6. The springback increases with increase in strip length (18.5%) upon a length of 95 mm and then decreases a little (7.2%) for a strip length of 115 mm. Whereas there is only a minute variation with punch speed and also there are some interesting facts observed with the lubricants on springback with variations for different lubricants. Similar results have been observed by the authors in the literature [18]. Lubricants of different viscosity had an impact on the springback in their work similarly same results observed in the present work. The effect of punch speed on springback is not much observed by the literature studies also [19]. Only a minute variation of springback is observed in the present study also similar to the literature study.

Table 7.8 represents the response for mean of means on Springback and it is observed that the most influencing factor on the springback is the strip length and lubricant and only a minute influence by punch speed is observed. Similar results were also observed by the literature studies.

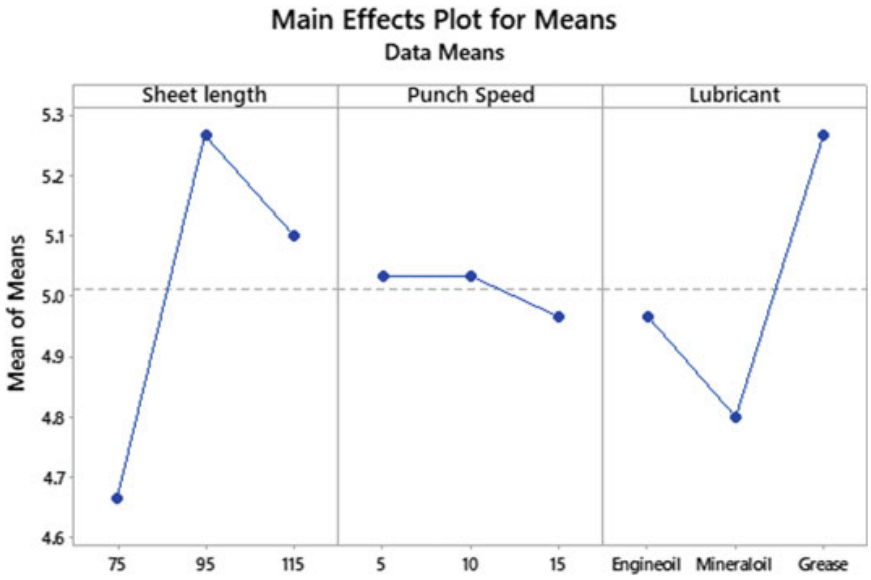


Fig. 7.6 Parameters affect plot on springback

Table 7.8 Response for mean of means on springback

Level	Strip length	Punch speed	Lubricant
1	4.667	5.033	4.967
2	5.267	5.033	4.800
3	5.100	4.967	5.267
Delta	0.600	0.067	0.467
Rank	1	3	2

Table 7.9 ANOVA table of springback angle

Source	DF	Adj SS	Adj MS	*F*-value	*P*-value
Strip length	2	0.575556	0.287778	64.75	0.015
Punch speed	2	0.008889	0.004444	1.00	0.500
Lubricant	2	0.335556	0.167778	37.75	0.026
Error	2	0.008889	0.004444		
Total	8	0.928889			
Model summary					
S	*R*-sq	*R*-sq(adj)	*R*-sq(pred)		
0.0666667	99.04%	96.17%	80.62%		

Table 7.9 represents the ANOVA results for springback. It is observed that the strip length and lubricant are the most influencing factors that affect the springback and the percentage of contributions is about 62.5 and 36.4% respectively. Similar results were observed with the literature studies, all the obtained results are validated with the works carried out by different authors in their works.

7.4 Conclusions

Sheet metal U-Bending was conducted on SS 304 and the forming parameters were analyzed in terms of punch force and springback by varying the strip length, punch speed, and lubricant. Using Taguchi method its optimum machining conditions were determined. Using ANOVA significance parameters were defined. From the present work, the following conclusions were drawn.

- From the statistical analysis, it is observed that strip length is the most influencing parameter that affects both punch force and springback.
- From Taguchi, the optimum forming parameters for punch force were identified as Strip length of 75 mm, Punch speed of 10 mm/min and engine oil as lubricant and are validated by the experiment. The punch force for the optimum conditions is observed as 10.4 kN.
- From Taguchi, the optimum forming parameters for springback were identified as Strip length of 75 mm, Punch speed of 15 mm/min and mineral oil as lubricant and are validated by the experiment. The springback angle for the optimum conditions is observed as 4.5°.
- Obtained results of both punch force and springback are compared to the literature works and a good agreement among the results is observed.

References

1. Gan, W., Wagoner, R.H.: Die design method for sheet springback. Int. J. Mech. Sci. **46**(7), 1097–1113 (2004)
2. Ayres, R.A.: SHAPESET: a process to reduce sidewall curl springback in high-strength steel rails. J. Appl. Metalwork. **3**(2), 127–134 (1984)
3. Liu, Y.C.: The effect of restraining force on shape deviations in flanged channels. J. Eng. Mater. Technol. **110**(4), 389–394 (1988)
4. Chou, I.N., Hung, C.: Finite element analysis and optimization on springback reduction. Int. J. Mach. Tools Manuf. **39**(3), 517–536 (1999)
5. Chan, W.M., Chew, H.I., Lee, H.P., Cheok, B.T.: Finite element analysis of spring-back of V-bending sheet metal forming processes. J. Mater. Process. Technol. **148**(1), 15–24 (2004)
6. Reche, D., Besson, J., Sturel, T., Lemoine, X., Gourgues-Lorenzon, A.F.: Analysis of the air-bending test using finite-element simulation: application to steel sheets. Int. J. Mech. Sci. **57**(1), 43–53 (2012)
7. Yoshihara, S., Manabe, K.I., Nishimura, H.: Effect of blank holder force control in deep-drawing process of magnesium alloy sheet. J. Mater. Process. Technol. **170**(3), 579–585 (2005)
8. Marretta, L., Ingarao, G., Di Lorenzo, R.: Design of sheet stamping operations to control springback and thinning: a multi-objective stochastic optimization approach. Int. J. Mech. Sci. **52**(7), 914–927 (2010)
9. Maia, A., Ferreira, E., Oliveira, M.C., Menezes, L.F., Andrade-Campos, A.: Numerical optimization strategies for springback compensation in sheet metal forming. In: Computational methods and production engineering, pp. 51–82 (2017)
10. Kakandikar, G., Nandedkar, V.: Springback optimization in automotive shock absorber cup with genetic algorithm. Manuf. Review. **5**, 1 (2018)
11. Vasudevan, D., Srinivasan, R., Padmanabhan, P.: Effect of process parameters on springback behaviour during air bending of electrogalvanised steel sheet. J. Zhejiang Univ. Sci. A. **12**(3), 183–189 (2011)
12. Buang, M.S., Abdullah, S.A., Saedon, J.: Effect of die and punch radius on springback of stainless steel sheet metal in the air v-die bending process. J. Mech. Eng. Sci. **8**, 1322–1331 (2015)
13. Maske, N.A., Sawale, J.K.: Taguchi approach for investigation of springback effect in aluminum sheet. Int. J. Mech. Eng. Rob. Res. **2** (2013)
14. Gautam, V., Sharma, P., Kumar, D.R.: Experimental and numerical studies on spring back in U-bending of 3-ply cladded sheet metal. Mater. Today Proc. **5**(2), 4421–4430 (2018)
15. Fei, N.C., Mehat, N.M., Kamaruddin, S.: Practical applications of Taguchi method for optimization of processing parameters for plastic injection moulding: a retrospective review. In: ISRN Industrial Engineering (2013)
16. Reddy, P.V., Reddy, B.V., Rao, P.S.: A numerical study on tube hydroforming process to optimize the process parameters by Taguchi method. Mater. Today Proc **5**(11), 25376–25381 (2018)
17. Logothetis, N.: Managing for total quality: from Deming to Taguchi and SPC. Prentice Hall (1992)
18. Narayanasamy, R., Padmanabhan, P.: Influence of lubrication on springback in air bending process of interstitial free steel sheet. J. Mater. Eng. Perform. **19**(2), 246–251 (2010)
19. Choi, M.K., Huh, H.: Effect of punch speed on the amount of springback in U-bending process of auto-body steel sheets. In: ICTP (2014)
20. Reddy, P.V., Ramulu, P.J., Madhuri, G.S., Govardhan, D., Prasad, P.R.: Design and analysis of deep drawing process on angular deep drawing dies for different anisotropic materials. IOP Conf. Ser. Mater. Sci. Eng. **149**, 012 (2012)

Chapter 8
Experimental Analysis on Springback of JIS 3302 Grade Steel Sheet Material Under Various Heat-Treated Conditions

Kemal Ramato, Perumalla Janaki Ramulu and N. R. R. Anbusagar

Abstract JIS 3302 grade steel sheet material has a wide range of applications in automotive industries. Any kind of sheet metal has to undergo different forming test before a real-time application. There are many forming processes are existed to check the formability of sheet materials. Among all the forming process, bending process is one of the formability tests had a drawback interms of springback. The springback behavior is a measurable quantity for any material after the test. Based on the springback nature of sheet material end product shape finalization will be done. This has a dictum for the present work on springback effect analysis of JIS 3302 steel sheet material is focused. And also, the effect of various heat treatments on springback characteristics of JIS 3302 steel sheet metal studied using experimental and numerical analysis. For the experimental work, sheet was cut with the dimensions of specimen 150 mm × 40 mm × 2 mm in three different rolling directions such as 0°, 45°, and 90° and bending test conducted on V-bending die with 90°. For the tensile testing specimens were cut as per ASTM standard in three rolling directions. The results showed that before heat treatment, the workpiece has taken more load for bending and recorded higher value of springback, whereas, the annealed workpiece has the lowest load consumption to bend and lower springback value.

Keywords Springback · Sheet metal forming · Heat treatment · V-bending · JIS 3302 grade steel · Annealing · Normalizing

K. Ramato · N. R. R. Anbusagar
Program of Mechanical Design and Manufacturing Engineering, School of Mechanical, Chemical and Materials Engineering, Adama Science and Technology University, Adama, Ethiopia
e-mail: kemalramato63@gmail.com

N. R. R. Anbusagar
e-mail: asr2775@gmail.com

P. J. Ramulu (✉)
Program of Mechanical Design and Manufacturing Engineering & CoEAME, School of Mechanical, Chemical and Materials Engineering, Adama Science and Technology University, Adama, Ethiopia
e-mail: perumalla.janaki@astu.edu.et

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_8

8.1 Introduction

Plastic deformation is concerning about metal forming process which has huge applications in various industries. Their main attribute is high efficiency, repeatability, and economy; simultaneously keeping the proper surface condition in mass production. The most frequently used techniques for sheet metal forming are bending, stretching and stamping [1]. Bending is one of the most commonly used methods beyond the sheet metal forming process. The common products obtained with the aid of bending are automobile panels, supermarket shelves, automotive component, aerospace vehicles, pressure vessels, medical equipment, electrical component, drums and housing utensils, etc. [2]. During the forming process force applied when the force is removed, the metal tries to coming back to its original shape and results in a phenomenon named "springback". This is due to the elastic deflection that occurs in the material after the bending process takes place [3]. Ramulu et al. [4] studied springback analysis on AA 6061 Aluminum alloy sheets with seven different rolling directions of sheet with the combination of U and V die shape. In order to make bending at a desired angle the process parameters considered were different load, displacement, and die angle. Finally, concluded that, the effect of rolling direction on springback effects was very sharp and prominent. Main factors that change the springback effect were indicated as sheet thickness, material properties and tool geometry [5–12]. Based on these literatures, the present work aimed to investigate the effect of springback on JIS 3302 steel sheet metal of 2 mm thickness by experimental and numerical work with different heat-treating conditions of the sheet.

8.2 Methodology and Materials

8.2.1 Materials

JIS 3302 grade steel sheet of 2 mm thickness material was used for the present work. This grade material has a wide-ranging of applications and economic considerations in the industries. Table 8.1 shows the chemical composition of JIS 3302 steel material before heat treatment.

Table 8.1 JIS 3302 grade steel chemical composition (in wt%)

C	Si	Mn	P	S	Cr	Mo	Ni	Cu	W	Fe
0.05	0.0065	0.19	0.078	0.035	0.041	0.003	0.014	0.02	0.04	99.5

8.2.2 Experiment: Bending Process

The specimens were prepared in the form strips at different rolling directions (0°, 45°, and 90°) of sheet. This specimen was cut by hydraulic shearing machine from 1000 mm × 2500 mm × 2 mm sheet metal with specimen dimensions of 150 mm × 40 mm × 2 mm as shown in Fig. 8.1. In total, 9 experiments were planned in which 3 tests at ambient condition, 3 tests at full annealing condition and other three tests at normalizing condition in three rolling directions. In a whole, 27 specimens were prepared for three repetitions for each. The full annealing and normalizing heat treatment methods were performed at 900 °C. This 900 °C temperature was considered due to virtually near to recrystallization temperature where phase changes. When phage changes, change in mechanical properties is inevitable.

The V-bending process was used for the bending tests. The experiment set-up was composed of a punch, a die and guide pins which are given in Fig. 8.2. This experiment was performed on hydraulic press brakes machine. The movement of the punch was hydraulically powered and applied enough force on the blank. Experimental setup was not disturbed for all experiments and the tests were conducted under the same operating conditions. The punch stroke was fixed and not moved up to the bottom of the die. A gap of 2 mm, which was equal to sheet thickness was maintained between the final position of punch and the die to avoid crushing of the sheet material at the bending profile. The angle of the die was 90°.

8.2.3 Tensile Test Procedure

The tensile specimens were made from sheet at rolling direction of 0°, 45° and 90° as per ASTM standard E-8 M as shown in Fig. 8.3. Tensile tests were performed at

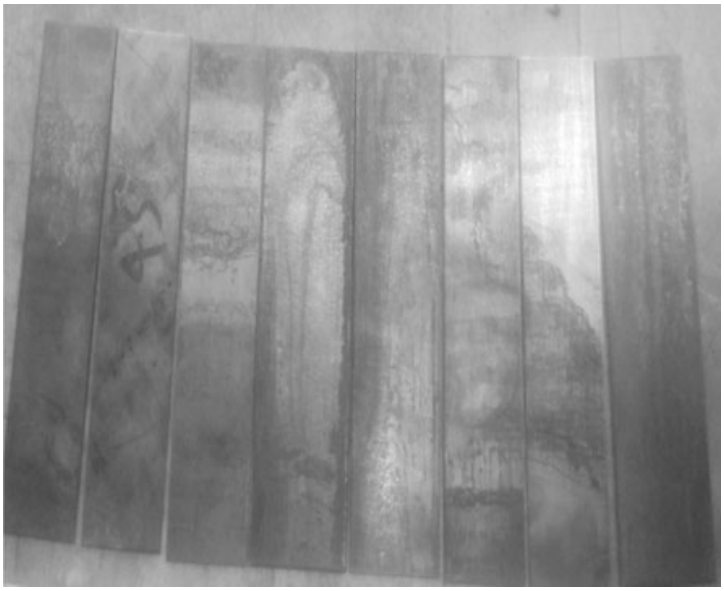


Fig. 8.1 Specimen for bending test with 150 mm × 40 mm × 2 mm dimensions

Fig. 8.2 Experimental setup for V-bending

Fig. 8.3 Specimen prepared for tensile test

a nominal cross-head speed of 1 mm/min, of samples at room temperature and annealed at 900 °C conditions on an ISO 68982-1 universal testing machine. Tests were repeated three times for each condition. The mechanical properties of JIS 3302 grade steel sheet material like yield strength, ultimate tensile strength, total elongation, uniform elongation, strain hardening coefficient (n) and strength coefficient (K) were evaluated as per the standard procedure after testing the samples till failure. The obtained load-stroke behavior during testing was converted into engineering stress–strain and true stress–strain plot for evaluating the mechanical properties. The plastic strain ratios at different rolling directions were obtained after testing the base material to 10% plastic strain.

8.3 Results and Discussion

8.3.1 Tensile Test Result

The results of tensile tests are shown in Table 8.2 such as ultimate stress (Tensile stress), yield stress, total elongation, strain hardening coefficient (n) and material strength (K). The ultimate stress and yielding stress for annealed specimens are the lowest when it is compared with the results from normalized and before heat treating specimen as presented in Table 8.2.

8.3.2 Springback Results from the Experiments

The experiments are conducted on JIS 3302 grade steel sheet of 2 mm thickness with different rolling directions. For illustration few of the deformed sheets are shown in Fig. 8.4. For each test, three repetitions are followed. In each repetition, springback angles are measured with equal time interval. All the experimental data is tabulated in Table 8.3. Among all the specimens, the major springback has seen without heat-treated specimen as 3.5° in 90° rolling direction and minimum is in annealed sample as 0.5° in 0° rolling direction (Fig. 8.5).

The effect of springback and resulting properties of mechanically formed steel plates have been presented and discussed. It is found that the 2 mm V-bended sheet metal exhibited the largest value of springback 3.060 for experimental. Due to the springback errors, the desired angle of 90° was not achieved. From this of point, it can be concluded that springback causes geometric inaccuracies in sheet metal bending.

Table 8.2 Mechanical properties of JIS 3302 grade steel sheet material

Condition of specimen	Rolling direction	Trials	Mechanical properties of the specimens				
			Tensile stress (MPa)	Yield stress (MPa)	% elongation	*n*	*K* (MPa)
Annealing	0°	1	350.3	278.4	52	0.35	723.3
		2	361	214.3	60	0.855	869.3
		3	408.4	330.2	60	0.53	971.0
		Average	373.2	274.3	57.33	0.578	854.5
	45°	1	300	212	58.375	0.32	720.5
		2	299.6	174.1	56	0.4	680.4
		3	320.3	275	54	0.52	856.1
		Average	386.5	220.4	56.125	0.413	763.9
	90°	1	395.5	132.4	55.6	0.49	824.4
		2	397.1	152.4	48.1	0.376	801.7
		3	405.4	289.8	60.7	0.4354	885.3
		Average	399.3	191.5	54.8	0.4338	837.1
Normalizing	0°	1	382.8	376	84.6	0.49	877.4
		2	424.9	352.7	61.5	0.491	977.4
		3	451.9	395.1	61.5	0.3758	883.6
		Average	419.8	308.2	69.2	0.452	912.8
	45°	1	485.4	267.8	73.1	0.5	945.1
		2	402.9	348.1	60	0.51	956.2
		3	444.7	380.2	55.2	0.47	934.6
		Average	431.3	332.0	62.77	0.493	945.9
	90°	1	492.2	421.8	53.8	0.51	1154.5
		2	395.5	334	61.5	0.5325	942.0
		3	440.1	372.3	53.8	0.5412	1051.4
		Average	442.6	176.9	56.37	0.5279	1049.3
Without heat treatment	0°	1	572.5	508.5	44	0.258	1049
		2	562.6	501.8	60	0.0038	363.5
		3	572	529.5	48	0.246	1030.8
		Average	569.0	513.3	50.7	0.169	814.4
	45°	1	474.7	387.3	46.7	0.2864	903.6
		2	534.7	447.8	48.2	0.34	765
		3	584.7	476.65	48	0.53	876
		Average	531.4	437.25	46.3	0.385	848.2
	90°	1	562.8	501.2	46.2	0.153	912.1
		2	549.6	482.3	38.5	0.262	1014.0
		3	521.5	124.3	38.5	0.3	1024.4
		Average	544.6	369.27	41.07	0.238	983.5

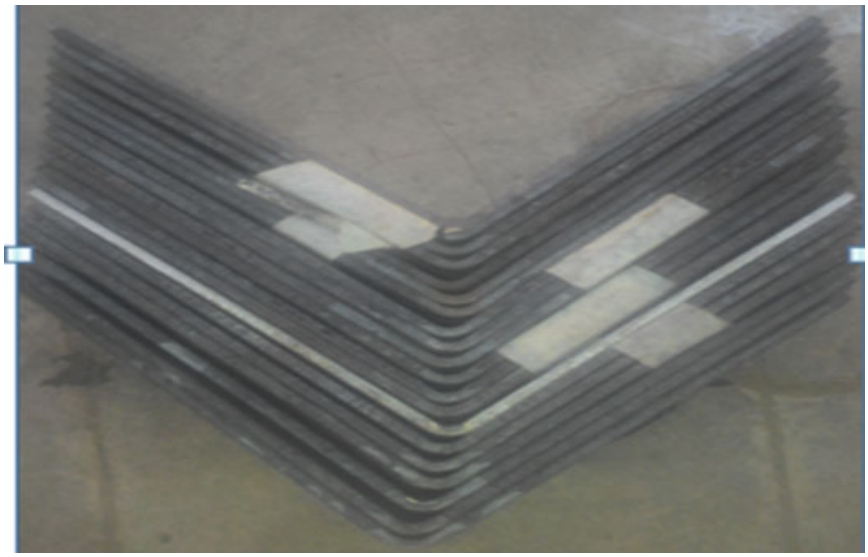


Fig. 8.4 Specimens after V-bending at different heating conditions

Table 8.3 Experimental results of measure springback

Heat treat and with heat treat specimen	Cutting angle of specimen	Experimental result
Annealing	0°	1
	45°	1.43
	90°	1.4
Normalizing	0°	1.5
	45°	1.4
	90°	1.43
Without heat treat	0°	2.6
	45°	3.06
	90°	2.067

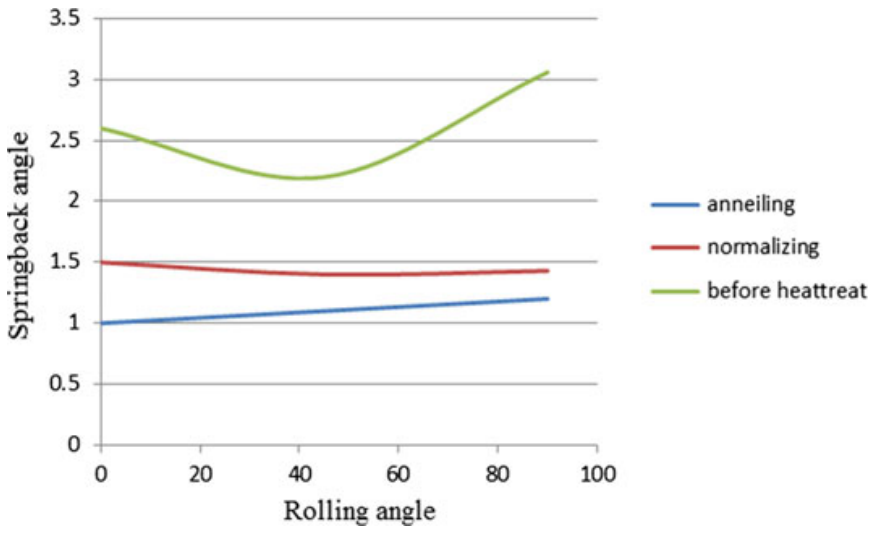


Fig. 8.5 Springback comparison with rolling angle

8.4 Conclusions

In this study, several bending tests are conducted to determine the effect of heat treatment on the springback. For this work, two different types of heat treatment were applied to the JIS 3302 grade steel and the amount of springback is determined under three circumstances. Moreover, the effect of rolling angle to the springback is also studied individually. Hence, the general conclusions attained in this study are as follows:

- The specimen used have the highest yield strength without heat treatment condition compared to annealing and normalizing and it also observed that higher springback obtained in the specimen compared to others.
- It observed that when rolling direction increase the elongation of specimen is decreases.
- In heat treatment condition, required punch force to bend the sheet metal was higher than the material in normalizing condition where the lowest in annealing condition.

References

1. Alghtani, A.H.: Analysis and optimization of springback in sheet metal forming. Ph.D. Diss., University of Leeds (2015)
2. Zulkafli, N.: Experimental and analytical evaluation of bending for aluminum. PhD diss., UMP (2012)
3. Wan-Nawang, W.A., Qin, Y., Liu, X.: An experimental study on the springback in bending of w-shaped micro sheet-metal parts. In MATEC Web of Conferences 2 (2015)
4. Ramulu, P.J., Rao, P.S., Yimer, W.: Springback analysis on AA 6061 aluminum alloy sheets. AIP Conf. Proc. **1769**(1), 200023 (2016)
5. Saikiran Reddy, P., Janaki Ramulu, P., Durga Srinivas, A., Viswa Teja, Y., Pavani, S.: Springback studies on different grades of steel sheets. Appl. Mech. Mater. **813**, 629–633 (2015)
6. Krinninger, M., Opritescu, D., Golle, R., Volk, W.: Experimental investigation of the influence of punch velocity on the springback behavior and the flat length in free bending. Procedia Cirp. **41**, 1066–1071 (2016)
7. Mustafa, O.: Mathematical modeling of the effect of different parameters on spring back in sheet metal formability process. Am. J. Eng. Res. **6**(10), 198–205 (2017)
8. Osman, M.A., Shazly, M., El-Mokaddem, A., Wifi, A.S.: Springback prediction in V-die bending: modelling and experimentation. J. Achie. Mater. Manuf. Eng. **38–2**, 179–186 (2010)
9. Wagoner, R.H., Lim, H., Lee, H.G.: Advanced issues in springback. Int. J. Plast **45**, 3–20 (2013)

10. Toros, S., Alkan, M., Ece, R.E., Ozturk, F.: Effect of prestraining on the springback behavior of the AA5754-O alloy. Mater. Tehnol. **45**(6), 613–618 (2011)
11. Soualem, A.: A detailed experimental study and evaluation of springback under stretch bending process. World Acad. Sci. Eng. Technol. Int. J. Mech. Aerosp. Ind. Mechatron. Manuf. Eng. **8**(6), 1128–1131 (2014)
12. Aso, T., Iizuka, T.: FEM analysis of springback control with lump-punch penetration after V-bending. J. Phys. Conf. Ser. **734**, 032068 (2016)

Chapter 9
Experimental and Numerical Assessment of Residual Stresses in AA6061 After Surface Treatment by Deep Cold Rolling

Rakesh Kumar, Gulshan Kumar, Vijay Gautam and Subhajit Konar

Abstract Deep cold rolling is a mechanical method of surface treatment used to adapt the desired surface properties in machine components for aerospace industries. In this process, localized plastic deformation is rendered on the surface by rolling a form tool wheel on the specimen to condense high dislocation density which imparts high compressive residual stresses resulting in increased tensile and fatigue strength. The main focus of the present study is to characterize the distribution of residual stresses on the surface and subsurface of an 8-mm-thick plate made of an aluminium alloy AA6061-O, after deep cold rolling operation to different depths of 0.5, 1, 1.5 and 2.0 mm. The observed residual stresses are correlated with the microstructural features in the deformed depths of the specimens. The residual stress and grain size in the deformed specimens are examined by using grazing incidence X-ray diffraction and electron backscatter diffraction technique, respectively. It is observed that deep cold rolling leads to significant grain refinement, resulting in high values of residual stresses through thickness. The compressive residual stresses are found to be higher on the surface of the specimen and reduce gradually with the penetration depth. The experimental results are compared with the results obtained by FE analysis using ABAQUS. A good agreement is observed between the experimental and the simulated values of residual stresses and the trends of stress distribution pattern.

Keywords AA6061 · Deep cold rolling · Residual stress · Grazing incidence X-ray diffraction · Electron backscatter diffraction · Finite element modelling

R. Kumar (✉) · G. Kumar
Thapar Institute of Engineering and Technology, Patiala, Punjab 147004, India
e-mail: rmnp71@gmail.com

V. Gautam · S. Konar
Delhi Technological University, New Delhi 110042, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_9

9.1 Introduction

Residual stresses are the stresses preserved in a material in the absence of any external force. These stresses can originate from phase transformation, thermal origin and defects like vacancies, dislocation, voids, cracks, etc. [1]. Residual stresses are mainly categorized into micro and macro stresses. Micro stresses are developed in grains due to existence of inclusions, dislocations, faults, etc. The macro stress is the average of stresses generated in each grain [2]. The stresses, compressive in nature, are beneficial in fatigue designing of components under dynamic load, whereas the tensile residual stresses are detrimental for the components subjected to fatigue and must be avoided. Undesirable residual stresses can lead to increased chances of crack propagation, distortion and premature failures in the components [3]. There are several methods to impart useful compressive residual stresses like deep cold rolling (DCR), shot peening, burnishing, surface hardening, etc. on the surface of the component to improve its fatigue strength [4]. Although shot peening is generally adopted for most components subjected to fatigue, DCR extends several benefits over the shot peening process. DCR imparts a deeper layer of compressive residual stress with remarkably good surface finish, but this process has a limitation in applying it for different geometrical profiles of the components [5]. This technique is widely used in automotive parts (e.g. crankshafts and shaft fillets), turbine blades, wheel flange, turbine disk, rotors, compressor fan blade, railway axles, etc. [6].

Subsequent paragraphs, however, are indented. DCR improves the fatigue strength, wear resistance and surface finish and decreases crack propagation and stress corrosion cracking by inducing a compressive residual stress on the surface [7]. This process is carried out using a roller shape or a ball-type deforming tool. As the rigid blunt roller or ball tool deforms the material underneath to a predetermined depth, a longitudinal groove is created along the length of the tool path resulting in an elasto-plastic zone. On removal of the deforming forces, the recovery of elastic region develops a zone of large compressive residual stress on the surface [5].

Wong et al. [5] explored the feasibility of DCR on titanium test specimens using three different cold rolling tool designs. The effects of process variables and design variables on yielding depth and residual stresses set-up were investigated. It was concluded that the rolling pressure affects the residual stress profile the most. Bozdana et al. [8] developed a new ultrasonic-assisted DCR, using static and dynamic forces of ultrasonic vibrations and investigated the effects of process parameters on titanium alloy test coupons. The process proved to be advantageous in reduced deforming pressure treating thin components without deteriorating the shape and a magnitude of 900 MPa of residual stress could be set-up at a depth of 0.2 mm. Zhu et al. [9] investigated the fatigue strength of titanium alloy specimens treated and untreated by ultrasonic-assisted DCR. Scanning electron microscopy of

the fractured specimens was also done. It was concluded that the new process was able to generate a deep layer of residual stress with high surface finish and resulted in enhanced fatigue strength of the fatigue test specimens. Xie et al. [14] studied the effects of ultrasonic-aided DCR process on microstructure of 30CrMnSiNi2A steel using nano-hardness tester, x-ray diffraction (XRD), scanning electron microscope (SEM), electron backscatter diffraction (EBSD) and transmission electron microscope (TEM). It was observed that severe plastic deformation by ultrasonic-aided DCR produced nano-crystallization, grain refinement, strain-induced martensite transformation, shear bands, intense shear texture and rolling texture in the surface and subsurface of the test specimens. Liu et al. [10] studied the mechanical properties of the surface prepared by ultrasonic surface rolling process, and nano-indentation tests were adopted to evaluate the force and depth of penetration. The magnitude of yield strength and strain hardening exponent were evaluated and the results were compared with the results obtained by FE simulations and found to be in good agreement. Juijerm et al. [11], surface treated the aluminium alloys AA5083 and 6110 using deep rolling at room temperature. It was observed that deep rolling enhanced the fatigue strength of aluminium alloys as compared to the polished condition due to near surface compressive residual stresses as well as work hardening and increased hardness induced by deep rolling.

Aluminium alloys due to its light weight, non-magnetic in nature, resistance to corrosion, high specific strength and better ductility are widely used in various industrial applications (aerospace, automotive, marine, building structure, etc.) [12], and offer lot of potentials to be surface treated by DCR. The development of a surface with high magnitude of compressive residual stress in high strength aluminium alloys depends on many factors like feed rate, roller speed, roller feed, applying force by roller, etc. Feed of roller affects the upper surface of the component, but does not affect the depth of penetration. The optimum combination of rolling force and rolling feed affects the surface hardness the most, whereas, if rolling force is maximum and feed is minimum then maximum surface hardness is achieved. The set-up for the DCR can be easily installed with conventional milling and lathe machines [13].

In the present work, the feasibility study of DCR on aluminium alloy AA6061 is carried out. The main focus of the study is to characterize the distribution of residual stresses on the surface and subsurface of 8-mm-thick plate made of an aluminium alloy AA6061-O after deep cold rolling operation to different depths of 0.5, 1, 1.5 and 2.0 mm. The observed residual stresses are correlated with the microstructural features in the deformed depths of the specimens. The residual stress and grain size in the deformed specimens are examined by using grazing incidence X-ray diffraction and electron backscatter diffraction technique, respectively.

9.2 Methodology

9.2.1 Experimental Work

Material Selection. In the present study, a wrought aluminium alloy AA6061 in annealed state is selected on the basis of its wide applications and high potential to be worked at room temperature by DCR. AA6061 is a precipitation hardening alloy, containing magnesium and silicon as its major alloying elements. The alloying elements enable the material to possess excellent tensile strength coupled with ductility and corrosion resistance and offer high resistance to fracture toughness. Some of the typical examples of applications of this alloy are heavy duty structures, marine structures, rail road, pipe lines, etc. The material is procured in the form of a plate with dimension 300 mm × 100 mm × 8 mm. The chemical composition of the alloy obtained by spark-based spectrometer is given in Table 9.1.

Tensile Properties. The tensile samples are prepared as per the ASTM standard by wire-cut electric discharge machining (WEDM) as shown in Fig. 9.1. A minimum of four samples of AA6061 is tested to obtain the scatter in results. The tensile specimens are tested on a 50 kN UTM, and load versus elongation data are procured using a dedicated Horizon software. The engineering stress–engineering strain curves are plotted to obtain the various tensile properties of the material. The engineering stress–strain plots are converted into true stress–true strain data and ln true stress versus ln true strain data between yield strength and UTS is used to determine the strain hardening exponent and strength coefficients.

Experimental set-up for DCR and sample preparation. Deep rolling set-up is installed on a horizontal milling machine and shown in Fig. 9.2. The milling cutter is replaced with a properly designed roller wheel of 75 mm in diameter, 5 mm thick

Table 9.1 Chemical composition of AA6061 in wt%

AA6061	Mg	Si	Cu	Cr	Fe	Ti	Zn	Co
	0.98	0.61	0.23	0.18	0.21	0.15	0.02	0.15

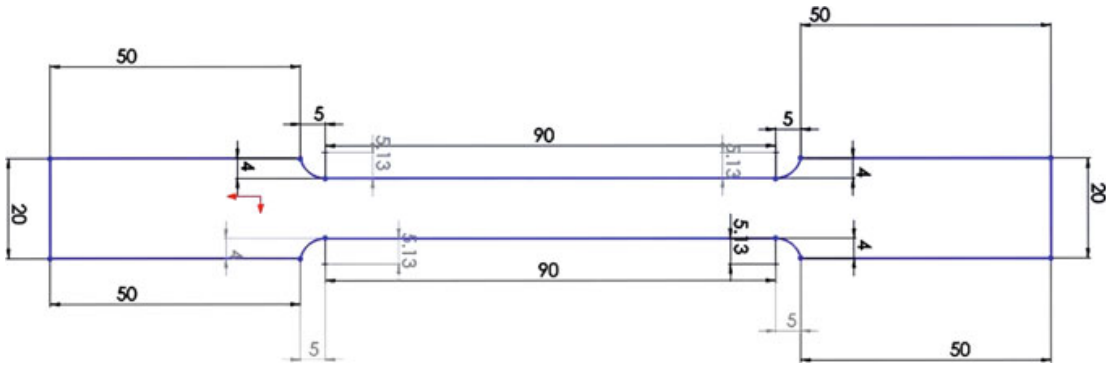


Fig. 9.1 Tensile sample details as per ASTM-E8M (all dimensions are in mm)

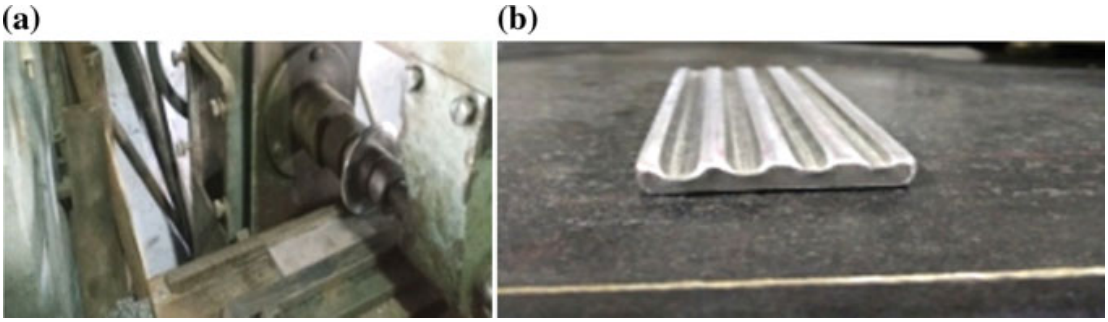


Fig. 9.2 **a** Deep rolling set-up on a milling machine (*Courtesy* Thapar University, Patiala) and **b** Deep cold-rolled aluminium specimen

and a nose radius of 2.5 mm made of D2 steel, hardened and tempered to deep roll the aluminium alloy at room temperature. The deep rolling process is carried out on the samples rigidly secured to the table using a fixture and the spindle is rotated at 10RPM to achieve the desired plastic deformation to different depths of 0.5, 1.0, 1.5 and 2 mm.

DCR-processed samples are carefully partitioned longitudinally by a diamond abrasive wheel cutter and polished in different stages from finishing with sub-micron SiC 1200 mesh to electropolishing to perform the microstructure characterization of different depth-rolled samples. The electropolishing is done on prepared samples so as to remove dust and scratches and to get a mirror surface grove of the samples. Electropolishing is done with perchloric acid 20% and methanol 80% at 27 V and 4.5 A current for a soaking time of 10 s.

Electron back scatter diffraction (EBSD), as shown in Fig. 9.3, is performed to measure the grain orientation, texture, grain size and grain boundary in plastically deformed regions of the processed samples to correlate the microstructure with residual stress.

Grazing Incidence X-ray diffraction. The grazing incidence X-ray diffraction (XRD) is a rapid Panalytical make, non-destructive system for analysing a large range of alloys, including metals, polymers and semiconductors as shown in Fig. 9.4. This technique is used to determine the residual stresses on a surface at an omega angle. In this case, the longitudinal stress is measured (σ_{11}) which is residual stress. Residual stress measurements on four samples processed by DCR to different depths of 0.5, 1.0, 1.5 and 2 mm are diagnosed on the surfaces at an omega angle of 0.1°, 0.5° and 0.8°.

9.2.2 Finite Element Modelling (FEM)

A three-dimensional, elasto-plastic finite element model is developed for DCR. The FE simulations were conducted using ABAQUS 6.13-1 software. In the simulations, an elastic plastic deformable rectangular aluminium alloy (Al6061) plate

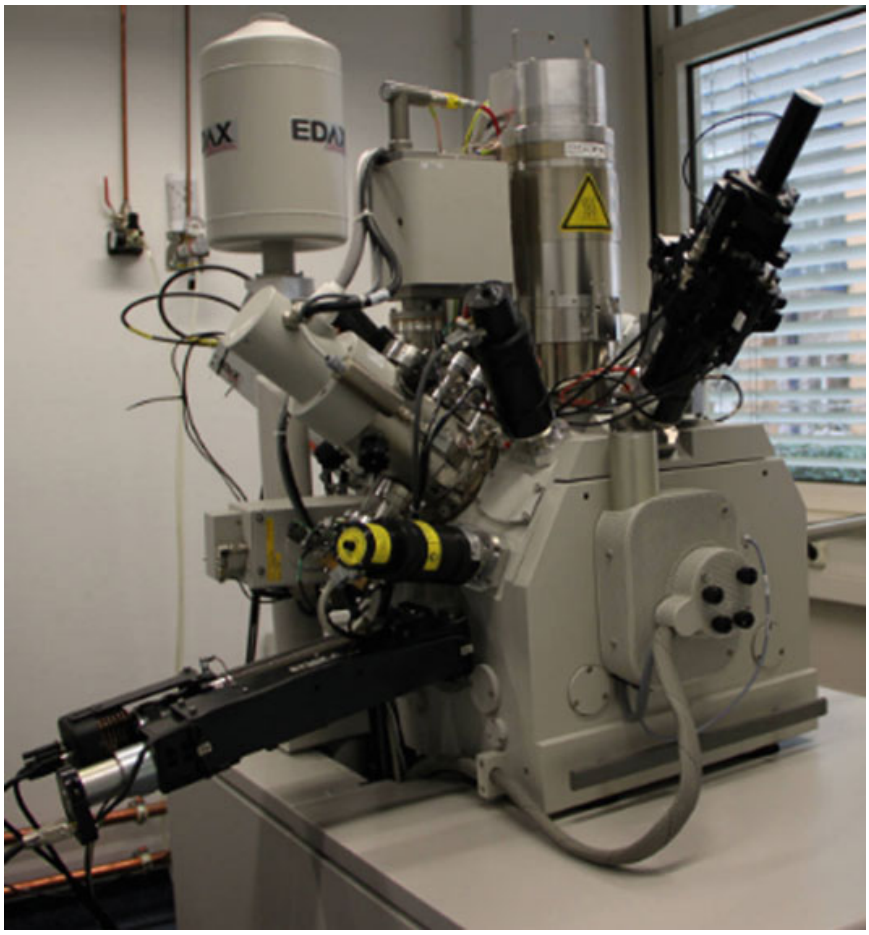


Fig. 9.3 Electron backscatter diffraction instrument to characterize grain structure (*Courtesy* IIT, Bombay)

of dimension 150 mm in length, 100 mm in width and 8 mm thickness is modelled and encastered along the edges to restrict the movement in any direction. The material model is designed to be isotropic in nature with modulus of elasticity (E) and Poisson's ratio (v) as 68.9 GPa and 0.33, respectively. The plastic data between 0.2% offset yield strength and ultimate tensile strength are selected from true stress–true strain plots to simulate the annealed behaviour of the material. The geometric model is using continuum stress/displacement and 8-nodded linear brick C3D8R elements. Element size plays an important role in attaining accurate results. A large number of elements can increase the computational time and a very small number of elements can lead to the inaccurate results. A structured mesh is used in this model. ABAQUS/Explicit has three different options for C3D8R elements, i.e. average strain kinematic, orthogonal and centroid formulations. The average strain kinematic formulation has been used in this study. Though, average strain kinematic formulation required more computational time compared to other two formulation techniques. The Lagrangian approach has been used in the present paper. In this approach, the initial configuration for the time $t = 0$ and the final state for the time $t + \Delta t$ are considered. Dynamic explicit method is used for the simulations.

Fig. 9.4 GIXRD Panalytical^{TM} set-up (*Courtesy* IIT, Bombay)

Mesh density is a significant metric used to control accuracy and a high-density mesh will produce results with high accuracy. However, if a mesh is too dense, it will require a large amount of computer memory and long runtimes, especially for multiple-iteration runs that are typical of nonlinear analyses. The basic fundamental and accurate method for evaluating mesh quality is to refine the mesh until von Mises stress value doesn't change significantly with each refinement.

In this study, 56,250 elements and 58,420 nodes with 15 elements in thickness direction and 250 elements are used in the width direction. To save the time, a biased mesh is generated in the deformation region using double biasing with a bias ratio of 3.

A roller wheel with 75 mm in diameter, 5 mm in thickness, 2.5 mm as nose radius is modelled as discrete rigid. The wheel is provided to have a uniform rotational speed of 10RPM. The simulations were conducted for penetration depths of 0.5, 1.0, 1.5 and 2.0 mm to deform the material longitudinally along the length of the specimen as shown in Fig. 9.5.

Coulomb friction was given to the contact surfaces and it is important to note that adequate friction is required to contact between the sheet and deep roller. Frictional forces at the interface of the roller and the sheet are responsible for pulling the sheet through the roller. A residual stress gradient has been observed at the surface of the deep-rolled sheet.

Boundary conditions are divided into two steps: an initial step and the analysis step. In the initial step, the sheet was given an initial velocity of 10 mm/s in

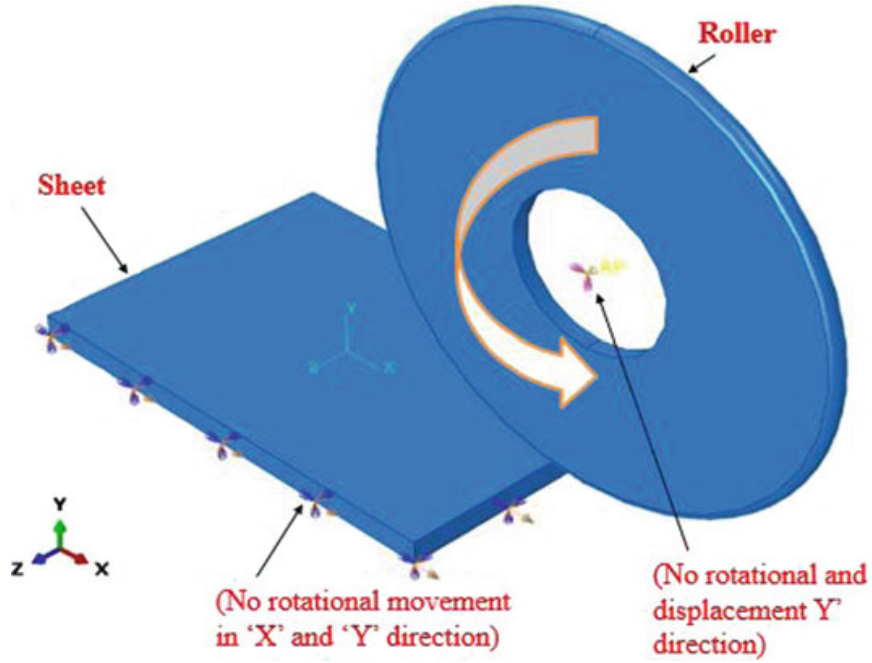


Fig. 9.5 A 3D finite element model DCR

'*x*'-direction. The interaction between the roller and the upper surface of the sheet is defined in this step. A displacement/rotation boundary condition is defined to constrain the motion of the sheet. The side faces of the sheet are made zero in '*y*'- and '*z*'-direction. Similarly, the boundary condition of the roller was made such that the roller could rotate only about their respective axis of rotation. All other translational and rotational degrees of freedom were constrained in '*x*'- and '*z*'-directions. In the analysis step, the rollers were given with an equal and opposite angular velocity. The friction model used in this study, allows the relative motion between the two contact surfaces. On the basis of literature survey on DCR, a frictional coefficient value of 0.2 is used in all the simulations.

9.3 Result and Discussion

9.3.1 Tensile Properties

The results obtained from the uniaxial tension tests conducted at room temperature are plotted and depicted in Fig. 9.6a and ln(true stress) versus ln(true strain) are shown in Fig. 9.6b. It is observed that the yield strength of annealed aluminium alloy AA6061 is 69 MPa and tensile strength is 132 MPa. The ductility of the alloy is excellent and found to be 42%. The strain hardening exponent is 0.3 on an average basis indicating good uniform stretchability.

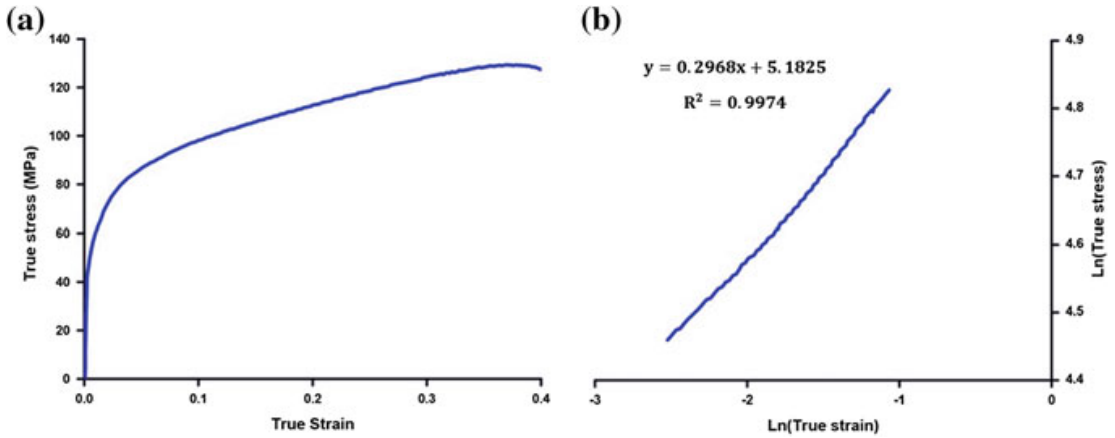


Fig. 9.6 **a** True stress versus true strain plot, **b** ln (true stress) versus ln (true strain) plot for tested tensile specimens of AA6061

9.3.2 Microstructure Characterization

Electron backscatter diffraction (EBSD) is used to characterize the microstructure of the deformed zones of AA6061 specimens. EBSD analysis revealed elongated grains in as-received parent material with a coarse grain size of 135.89 µm as shown in Fig. 9.7a. The various results of EBSD scan obtained at different DCR depths of 0.5, 1.0, 1.5 and 2.0 mm are shown in Fig. 9.7b–e, respectively. It is observed that the grain size with a deformation depth of 0.5 mm is 3.12 µm but the fine grains are segregated in a zigzag manner as shown in Fig. 9.7b. A fairly homogeneous ultrafine-grained microstructure is revealed at a deformation depth of 1.0 mm and the grain size is 1.95 µm. A similar ultrafine homogeneous grain size and distribution is observed with penetration depth of 1.5 and 2.0 mm without any remarkable difference in the average grain size.

9.3.3 Residual Stresses by Experiments and Simulations

The GIXRD method is used to determine the residual stress generated by DCR at the surface and subsurface of the specimens and the results obtained are plotted and shown in Fig. 9.8. It is interesting to find that the specimens of AA6061-O as procured has the range of residual stress ranging from −160 MPa(compressive) to 20 MPa (tensile) at the surface and at a depth of 0.6 mm below surface, respectively. After the first surface treatment by DCR with a tool penetration depth of 0.5 mm, there is a drastic increase in the values of residual stresses ranging from −230 to −160 MPa at surface and 0.6 mm below surface, respectively. All the generated residual stresses are compressive in nature. The similar trend is observed with tool penetration depth of 1, 1.5 and 2.0 mm but the magnitude of residual

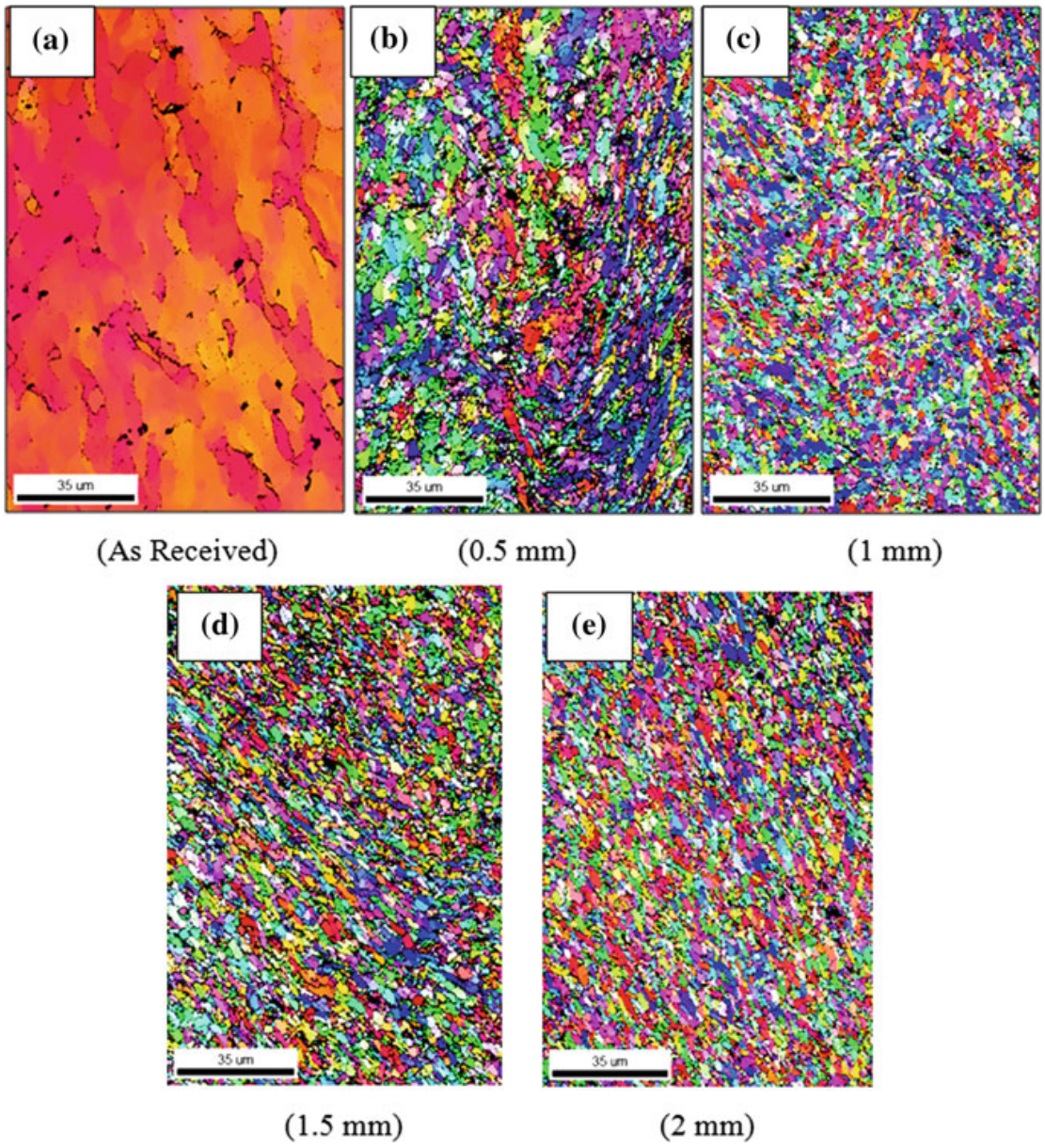


Fig. 9.7 EBSD scans after surface treatment by DCR

stresses is highest with higher depth of penetration by DCR. The values of residual stress on the surface of the material deformed is comparatively higher than the values generated over the subsurface, and the values of compressive residual stress is a function of depth of penetration by DCR, i.e. the stresses are higher on the surface and starts decreasing with increasing distance in depth. This could be attributed to the higher localized deformation at the surface directly under tool and lower concentration of deformation underneath the surface.

The results of residual stresses obtained by FE simulations are captured and plotted between longitudinal stress versus penetration depth for different tool depths of 0.5, 1.0, 1.5 and 2.0 mm and are shown in Fig. 9.9. It is clearly observed that the FE simulation results are in good agreement with the results obtained by

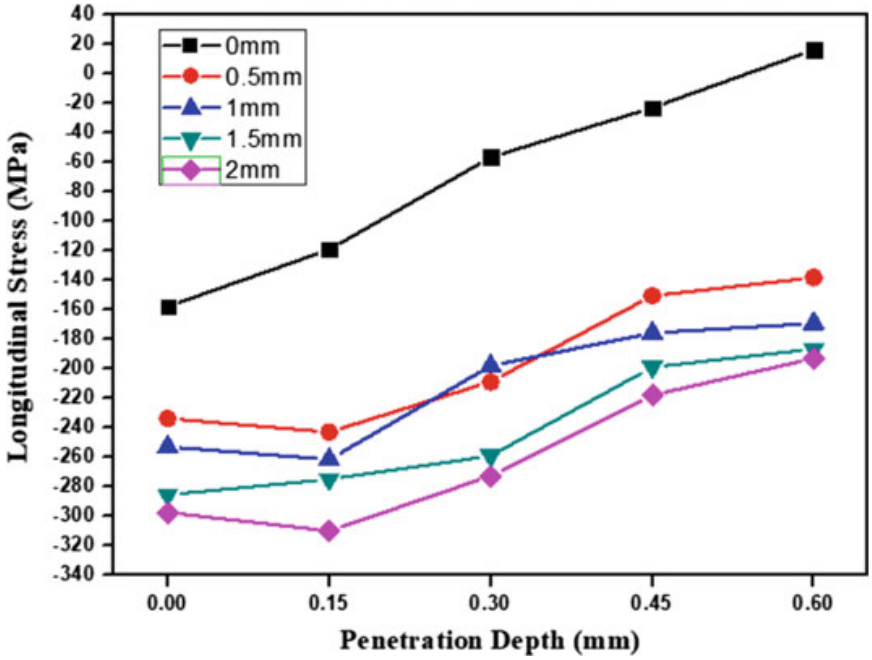


Fig. 9.8 Experimental results of residual stress versus penetration depth

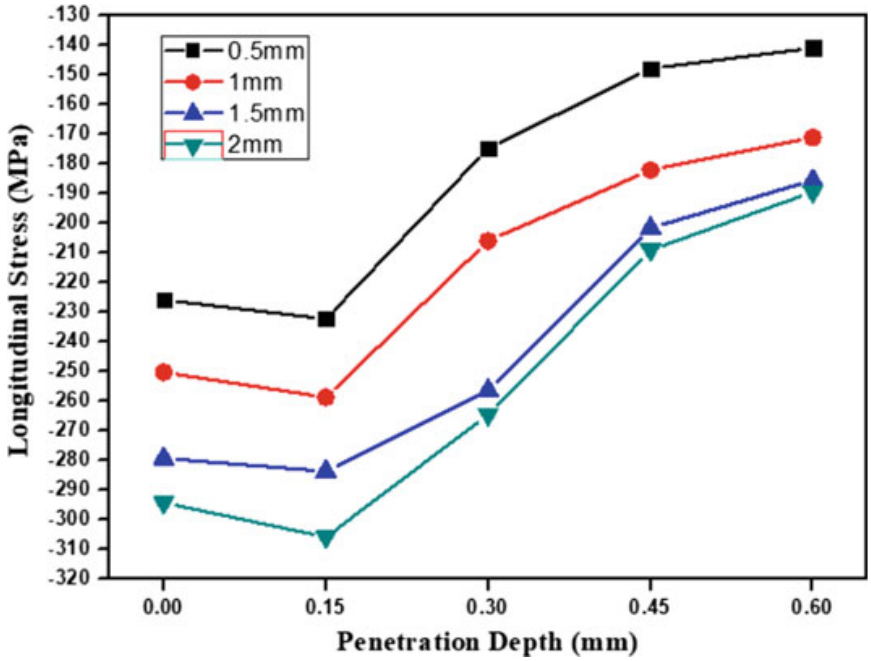


Fig. 9.9 FE simulation results of residual stress gradient in longitudinal direction (LD)

experiments. It is depicted from Table 9.2 that the percentage error in all cases between experimental and simulation for various DCR depth surfaces with their respective penetration depth are below 5%.

It is also observed that up to the depth of 0.15 mm, the values of residual stress increase and then reduce gradually with increasing distance beneath the surface.

Table 9.2 Chemical composition of AA6061 in wt%

DCR depth (mm)	% error				
	Penetration depth (mm)				
	0.00 (%)	0.15 (%)	0.30 (%)	0.45 (%)	0.60 (%)
0.50	3.29	4.36	4.98	1.72	2.10
1.00	1.02	1.03	4.04	3.70	1.20
1.50	2.13	3.20	1.00	1.46	0.48
2.00	1.07	1.29	2.93	3.99	1.76

This could be attributed to the higher plastic deformation resulting in higher dislocation density at a distance of 0.15 mm below the surface, and the material is getting plasticised between the tool and the material itself due to resistance to deformation from the strength offered by the material below the subsurface. A similar trend in the results of residual stresses is observed with different tool penetration depths by DCR of 0.5, 1.0, 1.5 and 2.0 mm. The contour of longitudinal residual stress obtained from simulations is shown in Fig. 9.10. It is found that the stresses on the surface of the specimens in the un-deformed region are tensile in nature whereas the stresses on the deformed region are compressive. This could be attributed to self-equilibrating nature of residual stress, i.e. the net sum of residual stress is zero.

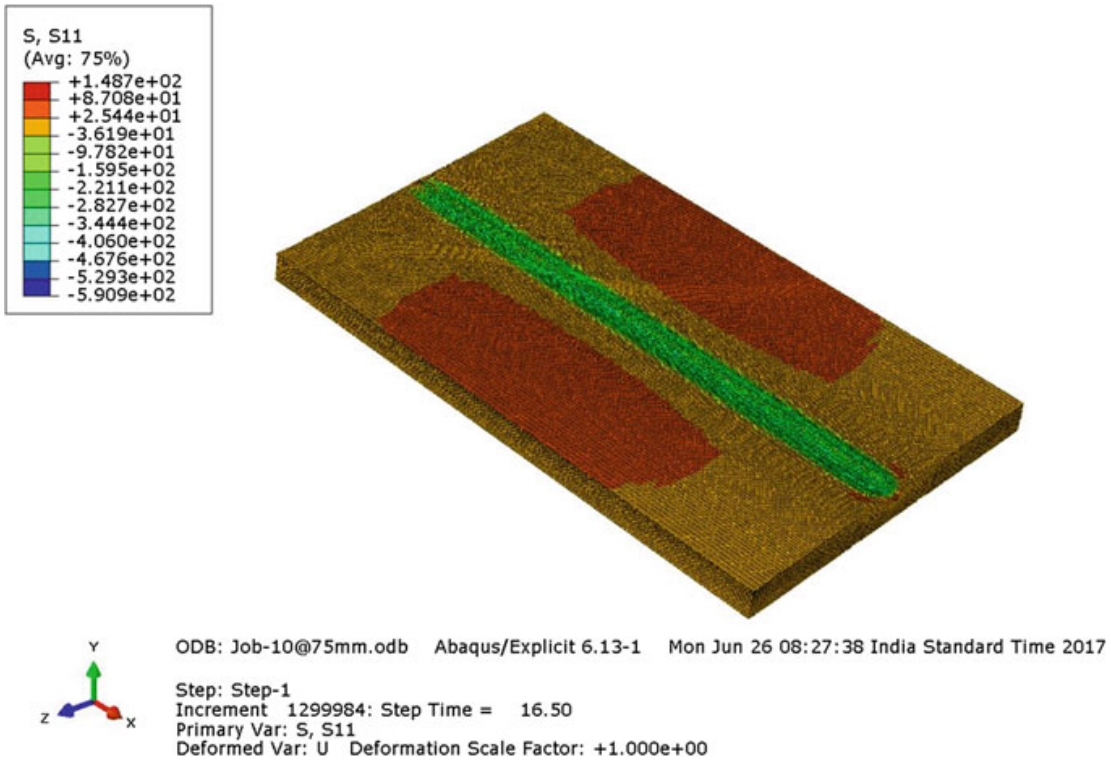


Fig. 9.10 FE simulation results of residual stress by DCR to 0.5 mm tool depth

9.4 Conclusions

Aluminium plate of AA6061 with 8 mm thickness is deep-rolled to different depths of 0.5, 1.0, 1.5 and 2.0 mm. On the basis of experimental work and FE simulations, the following conclusions can be drawn:

- The yield strength of aluminium alloy AA6061 in annealed state is 69 MPa and tensile strength is 132 MPa. The ductility of the alloy is excellent and found to be 42%.
- EBSD analysis revealed elongated grains in as-received parent material with a coarse grain size of 135.89 µm. The grain size with a deformation depth of 0.5 mm is 3.12 µm but the fine grains are segregated in a zigzag manner. A fairly homogeneous ultrafine-grained microstructure is revealed at a deformation depth of 1.0 mm and the grain size is 1.95 µm. A similar ultrafine homogeneous grain size and distribution are observed with penetration depth of 1.5 and 2.0 mm without any remarkable difference in the average grain size.
- The values of residual stress on the surface of the material deformed are comparatively higher than the values generated over the subsurface and is a function of depth of penetration by DCR, i.e. the stresses are higher on the surface and starts decreasing with increasing distance in depth. This could be attributed to the higher localized deformation at the surface directly under tool and lower concentration of deformation underneath the surface.
- The FE simulation results are in good agreement with the results obtained by experiments.

References

1. Understanding stress and strain. Material science and Engineering, http://home.iitk.ac.in/~anandh/E-book.htm. Last accessed 2017/05/23
2. Chatterjee, K., Venkataraman, A., Garbaciak, T., Rotella, J., Sangid, M.D., Beaudoin, A.J., Pilchak, A.L.: Study of grain-level deformation and residual stresses in Ti-7Al under combined bending and tension using high energy diffraction microscopy (HEDM). Int. J. Solids Struct. **94**, 35–49 (2016)
3. Resudual stress info. Proto manufacturing, http://www.protoxrd.com/residual-stress-info.html . Last accessed 2016/11/23
4. Majzoobi, G.H., Azadikhah, K., Nemati, J.: The effects of deep rolling and shot peening on fretting fatigue resistance of Aluminum-7075-T6. Mater. Sci. Eng. A **516**(1), 235–247 (2009)
5. Wong, C.C., Hartawan, A., Teo, W.K.: Deep cold rolling of features on aero-engine components. Procedia CIRP **13**, 350–354 (2014)
6. Liouand, T.I., Wardany, E.I.: Finite element analysis of residual stress in Ti-6Al-4V alloy plate induced by deep rolling process under complex roller path. Int. J. Manuf. Eng. http://dx.doi.org/10.1155/2014/786354
7. Altenberger, I.: deep rolling—the past, the present and the future (2005)

8. Bozdana, A.T., Gindy, N.N., Li, H.: Deep cold rolling with ultrasonic vibrations-a new mechanical surface enhancement technique. Int. J. Mach. Tools Manuf. **45**(6), 713–718 (2005)
9. Zhu, Y.L., et al.: Evaluation of an ultrasound-aided deep rolling process for anti-fatigue applications. J. Mater. Eng. Perform. **18**(8), 1036 (2009)
10. Liu, Y., Zhao, X., Wang, D.: Determination of the plastic properties of materials treated by ultrasonic surface rolling process through instrumented indentation. Mater. Sci. Eng. 21–31 (2014)
11. Juijerm, P., et al.: Fatigue lifetime improvement of aluminium alloys by mechanical surface treatment. In: Proceeding of Thailand Materials Science and Technology Conference
12. http://www.constellium.com/technology-center/aluminium-alloy-properties. Last accessed 2017/07/13
13. Beghini, M., Bertini, L., Monelli, B.D., Santus, C., Bandini, M.: Experimental parameter sensitivity analysis of residual stresses induced by deep rolling on 7075-T6 aluminium alloy. Surf. Coat. Technol. **254**, 175–186 (2014)
14. Xie, J., Zhu, Y., Huang, Y., Bai, C., Ye, X.: Microstructure characteristics of 30CrMnSiNi2A steel after ultrasound-aided deep rolling. J. Mater. Eng. Perform. **22**(6), 1642–1648 (2013)

Chapter 10
Analysis of Hot Workability of Inconel Alloys Using Processing Maps

Gauri Mahalle, Nitin Kotkunde, Amit Kumar Gupta and Swadesh Kumar Singh

Abstract Understanding deformation behavior and workability is key prerequisite to optimize process parameters to improve the material processing conditions and confirming safe performance during forming. The deformation characteristics of Inconel 718 and 625 alloys have been investigated using hot tensile tests at temperature range (room temperature to 973 K) at interval of 100 K and at quasi-static strain rate of 0.0001, 0.001, and 0.01 s^{-1}. The experimental findings confirmed that tensile flow stresses are considerably subtle to strain deformation temperature and strain rate. From tensile flow stress behavior, processing maps have been developed using dynamic material model approach at different value of strains. It was observed that the temperature range ($T < 550$ K) for Inconel 718 and 600–900 K for Inconel 625 for lower strain rate (0.0001 s^{-1}) efficiency were maximum at different considered strain values. The optimum working parameters for both the alloys have been identified.

Keywords Inconel alloys · Deformation behavior · Workability · Hot tensile test · Processing maps

10.1 Introduction

High-strength Inconel alloys are special grade of precipitation-hardenable Ni-Cr-based superalloys with higher concentration (by wt%) of nickel (51–61%), chromium (18–21%), and iron (18.5–4%) in addition with relative higher contents of niobium, molybdenum, titanium, aluminum and carbon [1]. Because of the excellent combination of alloying elements, Inconel alloy possesses superior

G. Mahalle · N. Kotkunde (✉) · A. K. Gupta
Department of Mechanical Engineering, BITS Pilani Hyderabad Campus, Hyderabad 500078, Telangana, India
e-mail: nitink@hyderabad.bits-pilani.ac.in

S. K. Singh
Department of Mechanical Engineering, GRIET, Hyderabad 500072, Telangana, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_10

mechanical properties like high yield (700–1250 MPa) and ultimate tensile strengths (950–1650 MPa), good creep and rupture strengths, and high resistance to fatigue and corrosion at elevated temperature. This results in broad applicability in aerospace, nuclear, marine, and chemical industries [2, 3]. Because of constricted forming temperature assortment, more deformation resistance and intricate microstructures, extensive investigation of flow behavior at high temperature of Inconel alloy is very important [4, 5]. Moreover, Inconel alloys are difficult to machine to aimed shape due to extreme toughness, poor thermal conductivity, work hardening characteristics, and presence of hard abrasive carbide at elevated temperature [6].

Understanding deformation behavior is key necessity to optimize process parameters for improving the material processing conditions and confirming safe performance during various manufacturing process. Significant investigations attempted to improve constitutive expression to define complex deformation characteristics by strain, temperature, and strain-rate effects [7, 8]. Tan et al. [9] examined the deformation characteristics for Inconel 718 superalloy tested in hot compression experiments and fitted Arrhenius constitutive equation. Further, optimal processing parameters were determined. Cai et al. [10] stated that strain does not influence the processing map in the test temperature (1050–1180 °C) and developed the deformation equation to describe necessity of peak stress on tested temperature and strain rate. Zhang et al. [11, 12] observed the flow behavior, microstructure evolution, and dislocation behavior with processing maps of the Ni-based superalloy. Peak efficiency in the process of deformation is highly influenced by dislocation density in instability domain with low DRX. Authors investigated effect of the early aging time and microstructure on the processing map of precipitation-hardenable Ni-based superalloy [13, 14]. Though, no substantial effort has been reported in studied literature on the processing plots of Inconel alloys at warm temperatures and quasi-static range strain rates. This paper focuses on the processing map for Inconel alloys at warm temperatures and quasi-static strain rate.

10.2 Materials and Experimental Procedures

In present experimental work, Inconel 718 and 625 alloy sheets, of thickness 1 mm with nominal composition (wt%) as mentioned in Table 10.1, are used. Sub-sized *ASTM E08/E8M-11* standard dimensions are used as flat tensile test specimens, as shown in Fig. 10.1a. Normally, strain-rate scale of the static tension test, either with screw driven or hydraulic machine, is measured in the middle of 10^{-5} and 10^{-1} s^{-1} [15]. It is valuable for studying hot forming performed at lower strain rates as Inconel alloys are strain-rate sensitive [16]. Tensile tests are conducted on a UTM with maximum weight carrying capacity of 50 kN. Two-zone split furnace is having heating capacity up to 1500 K with ±3 K accuracy. Tensile experiments are accomplished from RT to 973 K intervals of 100 K and at 0.01, 0.001, and

Table 10.1 Chemical composition (wt%) of received Inconel 718 and 625 alloy sheet

Element	Ni	Fe	Mo	Cr	Nb	Al	Mn	Ti	Si	Others
For Inconel 718	51.463	20.44	2.871	18.279	5.012	0.561	0.062	1.092	0.051	0.169
For Inconel 625	61.495	3.304	9.479	21.739	3.271	0.067	0.123	0.166	0.101	0.104

0.0001 s^{-1} quasi-static strain rates. First, samples are heated at 20 °C/min to their deformation temperature and then heat preservation time is 10 min to achieve uniform temperature. The Experimental procedure for Quasi-static isothermal hot tensile tests is shown in Fig. 10.1b.

Figure 10.2a–d gives representative true stress–strain curves of alloys measured along rolling direction with different temperatures. Decrease in the flow stress and increased fracture strain show that flow stresses are significantly influenced by test temperature and strain rates. First, flow stress rises linearly to yield stress value for small strain values, due to effect of strain hardening. Noticeable observation from true stress–strain curve at lower strain rate is the serrated yielding phenomenon reported over the temperature range 673–973 K. The serration behavior is described by the Portevin-Le Chatelier effect specially, when homogeneous deformation occurs. The serration observed in Inconel alloys is combination of type A and B, the rise in strain rate and test temperature, according to Rodriguez classification of serrated flow [17]. Similar observations were reported previous studies [18, 19]. Differently, confirming the serrated yielding phenomena is by strain-rate sensitivity (m) parameter. The negative m value confirms claim of PLC or DSA effect in alloys.

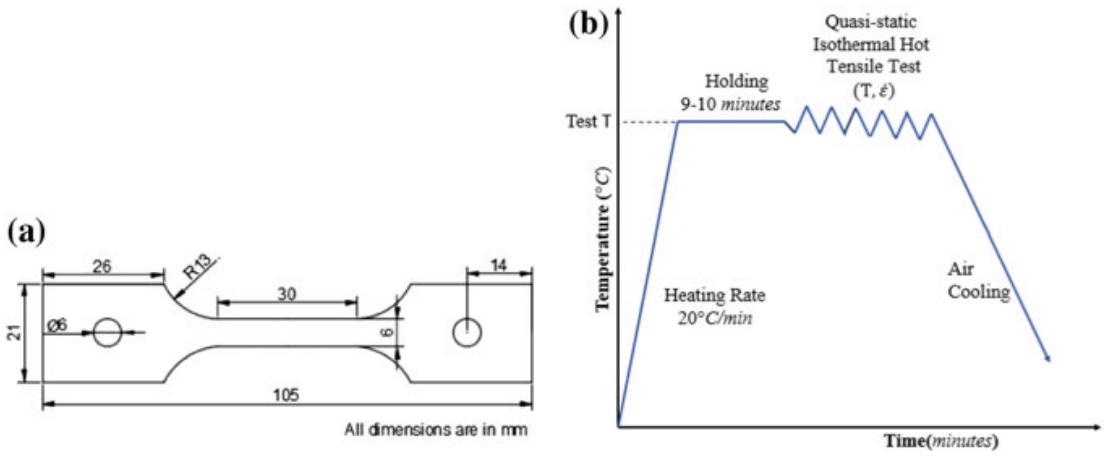


Fig. 10.1 **a** Schematic of used sub-sized ASTM E08/E8M-11 tensile test specimen, **b** Experimental procedure for Quasi-static isothermal hot tensile tests

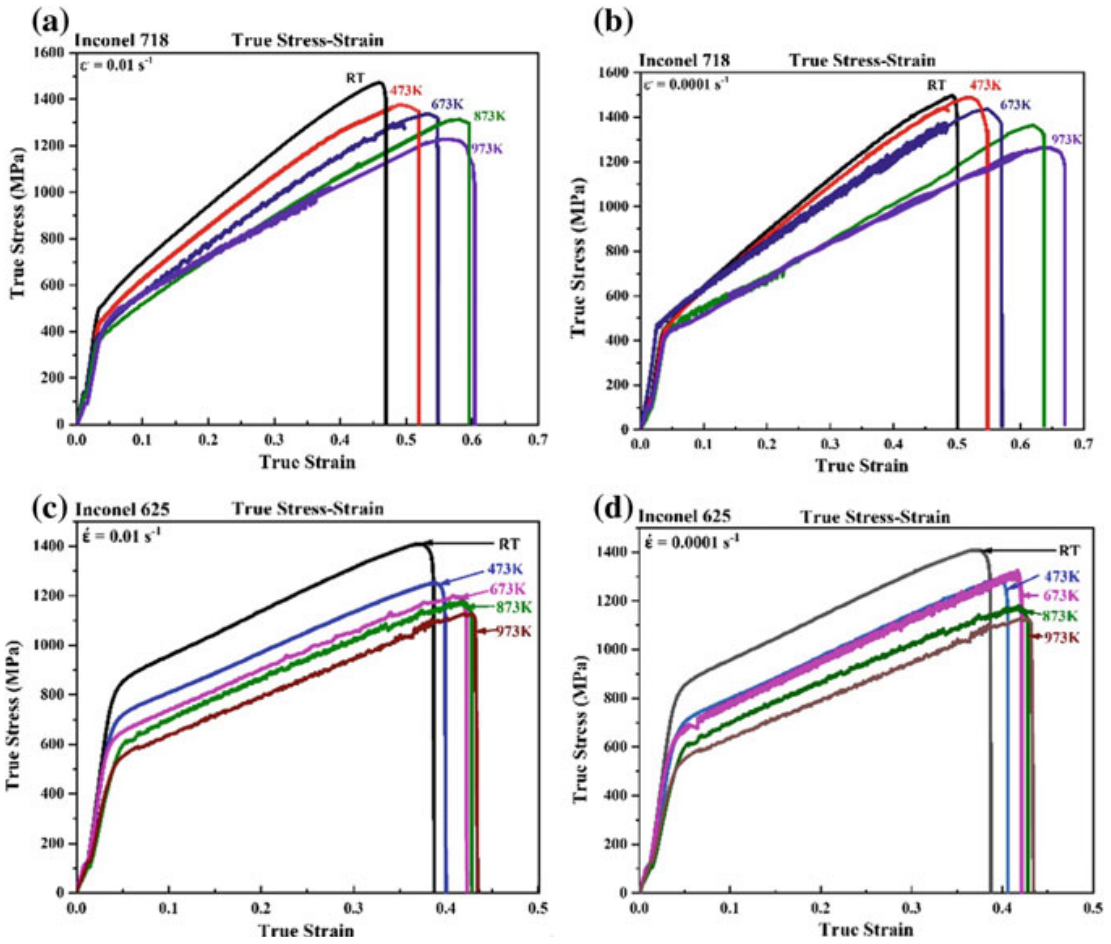


Fig. 10.2 True stress–strain curve of Inconel alloys at various test temperature with strain rate of **a–c** 0.01 s^{-1} and **b–d** 0.0001 s^{-1}

10.3 Processing Map

Processing map is basically combination of the power dissipation efficiency plot and flow instability plot for various temperatures and strain rates. It shows occurrences of damage as a function of processing variable (temperature, strain and strain-rate sensitivity). Processing map is very useful for the selection of deformation processing conditions. Fracture boundaries in processing map are very sensitive to prior thermomechanical history, microstructure, composition of metal, and numerous material parameters (like diffusivity).

At constant temperature and strain, the test metal piece has been considered as energy dissipation system. Considering material as energy dissipation, total input power (P) is [20]

$$P = \sigma\dot{\varepsilon} = G + J = \int_0^{\dot{\varepsilon}} \sigma \, \mathrm{d}\dot{\varepsilon} + \int_0^{\sigma} \dot{\varepsilon} \, \mathrm{d}\sigma \tag{10.1}$$

where σ and $\dot{\varepsilon}$ are flow stress and strain rate. G content defines the heat dissipation during plastic deformation and J co-content defines internal energy variation because of the microstructural evolution. While in plastic deformation process, heat dissipated mostly by G content (due to temperature rise) and less amount by J co-content (due to microstructural evolution). Strain-rate sensitivity (m) from flow stress expression is used to determine the power distribution among G content and J co-content. For plastically deforming materials, power law can be represented as

$$\sigma = k(T, \varepsilon, \dot{\varepsilon})\dot{\varepsilon}^{m(T,\varepsilon,\dot{\epsilon})} \tag{10.2}$$

when strain and deformation temperature are considered as constants. The flow stress expression is given as [20, 21]

$$\sigma = k\dot{\varepsilon}^{m} \tag{10.3}$$

where K is strain hardening constant and m is the strain-rate sensitivity, calculated as [20, 21]

$$m = \frac{\partial J}{\partial \sigma} = \frac{\dot{\varepsilon}\mathrm{d}\sigma}{\sigma \mathrm{d}\dot{\varepsilon}} = \frac{\partial(\ln \sigma)}{\partial(\ln \dot{\varepsilon})} \tag{10.4}$$

Figure 10.3 plotted with ln σ and ln $\dot{\varepsilon}$ values for tested temperatures and fitted with polynomial quadratic fit. m values are determined from slope of these lines. This value varies from 0 to 1 (m = 0 represents no power dissipation, m = 1 represents ideal linear dissipator, and $m > 0$ represents material with positive strain-rate sensitive).

Diversions of m at 0.15 strain with respect to strain rate and deformation temperature are plotted in Fig. 10.4. From the figures, m value first drops and then rises as deformation temperature increases. It is because of microstructure evolution which needs more heat dissipation for particular test temperature. Even when m is

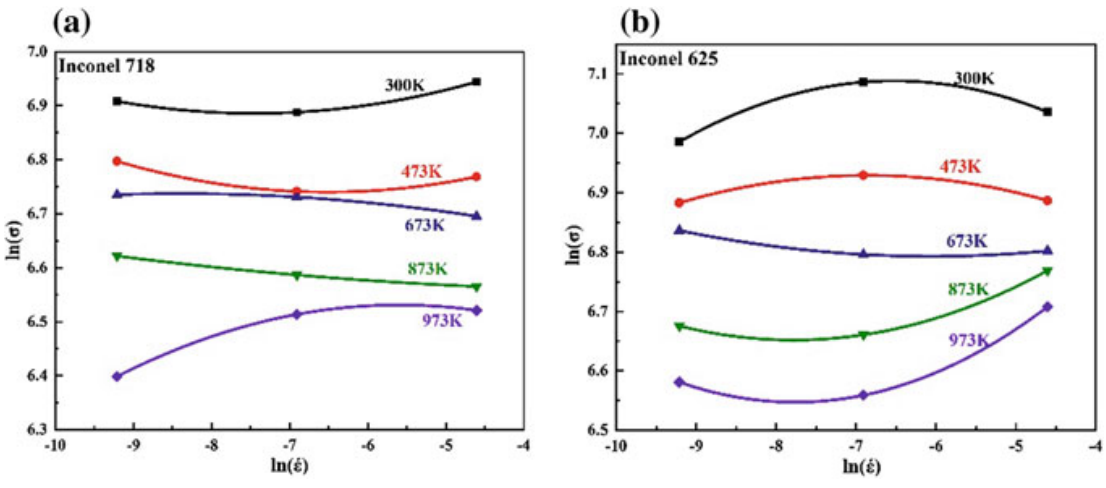


Fig. 10.3 Correlation between the flow stress and strain rate at 0.15 strain of Inconel alloys (**a**, **b**)

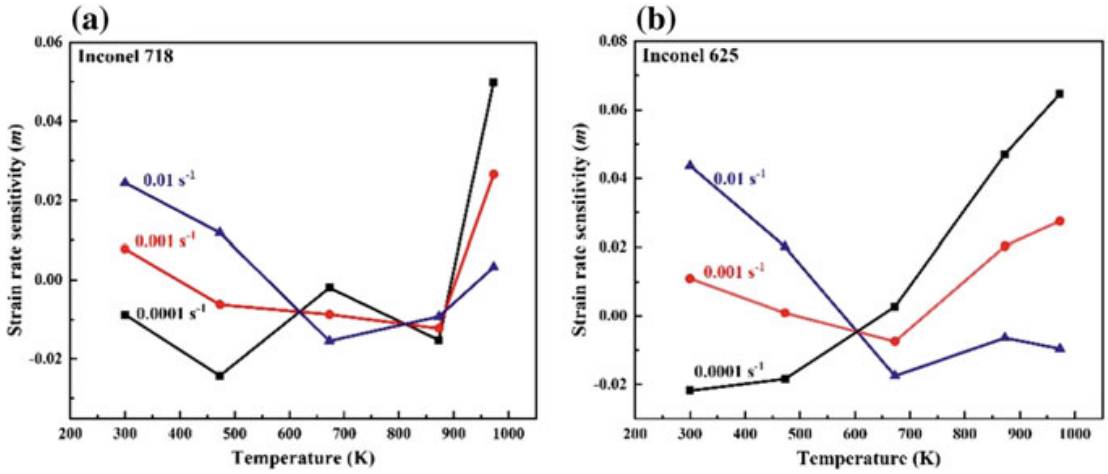


Fig. 10.4 Changes in strain-rate sensitivity (m) exponent at 0.15 strain with temperature for various strain rates of Inconel alloy (**a**, **b**)

negative or zero then also microstructural evolution can occur, this must be evaluated by joined analysis of both efficiency plot and instability plot.

Power dissipation map is one of the representations of the power dissipation efficiency (η) in the microstructural evolution. Dissipation efficiency (η) is expressed as

$$\eta = \frac{J}{J_{\max}} = \frac{2m}{m+1} \tag{10.5}$$

Here,

$$J = \int_0^{\sigma} \dot{\varepsilon} \mathrm{d}\sigma = \frac{m}{m+1}\sigma\dot{\varepsilon}$$
$$J_{\max} = \frac{\sigma\dot{\varepsilon}}{2} \tag{10.6}$$

and $J_{\max}$ represents the ideal linear dissipation (m = 1).

During deformation process, to identify the microstructural instability, number of flow instability criterion has to be reported by researchers [22, 23]. Most commonly used criteria for creation of the processing maps are Prasad instability criterion. Researchers have reported that Prasad instability criterion is more conventional to use than other [24, 25]. Therefore, Conventional Prasad's criterion is centered on maximum rate of entropy creation in the metallurgical system and considered to trace instable regions in processing map. It is represents as

$$\xi = \frac{\partial \ln[m/(m+1)]}{\partial \ln \dot{\varepsilon}} + m \tag{10.7}$$

here, ξ is the dimensionless instability parameter.

Instability plot is divided generally into two regions by instability parameter ξ: region with $\xi < 0$ considered as the instability domain, whereas $\xi > 0$ considered as the stable domain. Contour plot constructed to conclude difference of η at various strain rate and deformation temperature with a fixed strain value. Figure 10.5 signifies contour plot viewing changes of η value considered for Inconel alloys at different strain rate and deformation temperature at strain of 0.15. This plot is called as power dissipation plot. Contour value represents efficiency of the material's power dissipation. From this figure, it is noted that the η value rises with a drop in the strain rate at low test temperature as well as in high test temperature region. Highest for Inconel 718 alloy is 38%, which occurs at a temperature range of 300–480 K and lower stress rate (0.0001 s^{-1}). While the highest η around 35% for Inconel 625 alloy is occurring at 650–700 K temperature range also at lower strain rate of 0.0001 s^{-1}. Generally, except few unpredictable variations (such as unpredictable fracture), high power dissipation efficiency defines better workability of the material [24]. Hence, it is necessary to consider the other aspects (unreliable variations) while deciding the experimental processing conditions of material. The goal of microstructural optimization is not sufficient only on the basis of the highest efficiency [24, 25].

Figure 10.6 represents processing plots by superimposing efficiency plots on instability plots at 0.07, 0.15, and 0.20 strain. The shaded gray region represents the instability ($\xi < 0$). For Inconel 718 alloy, it is perceived that the instability takes place in lower temperature ($T > 600$ K) for strain rates of 0.0001 s^{-1}. This indicates that Inconel 718 alloy possess very poor workability in this range of temperature and strain rates. Whereas for Inconel 625 alloy, instability occurs in high temperature ($T > 700$ K) for all stain rates. The reason for the stability of Inconel 718 alloy is might be due to the formation of precipitate at test temperature which strengthen the alloy. For both the alloys, higher η values are detected in lower strain rates (0.0001 s^{-1}) region. Further, it can be detected that the influence of strain on unstable region is considerable. The total area of the unstable region increases with increase in strain as almost instability is noticed in high strain rate and low-temperature regions. The highest efficiency in the stable region for 0.15 and

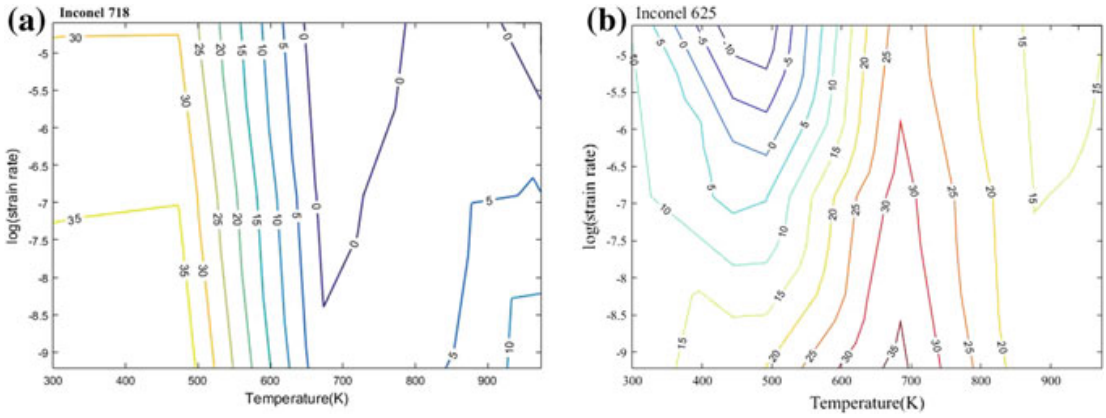


Fig. 10.5 Power dissipation efficiency plot for Inconel alloys (**a**, **b**) at 0.15 strain

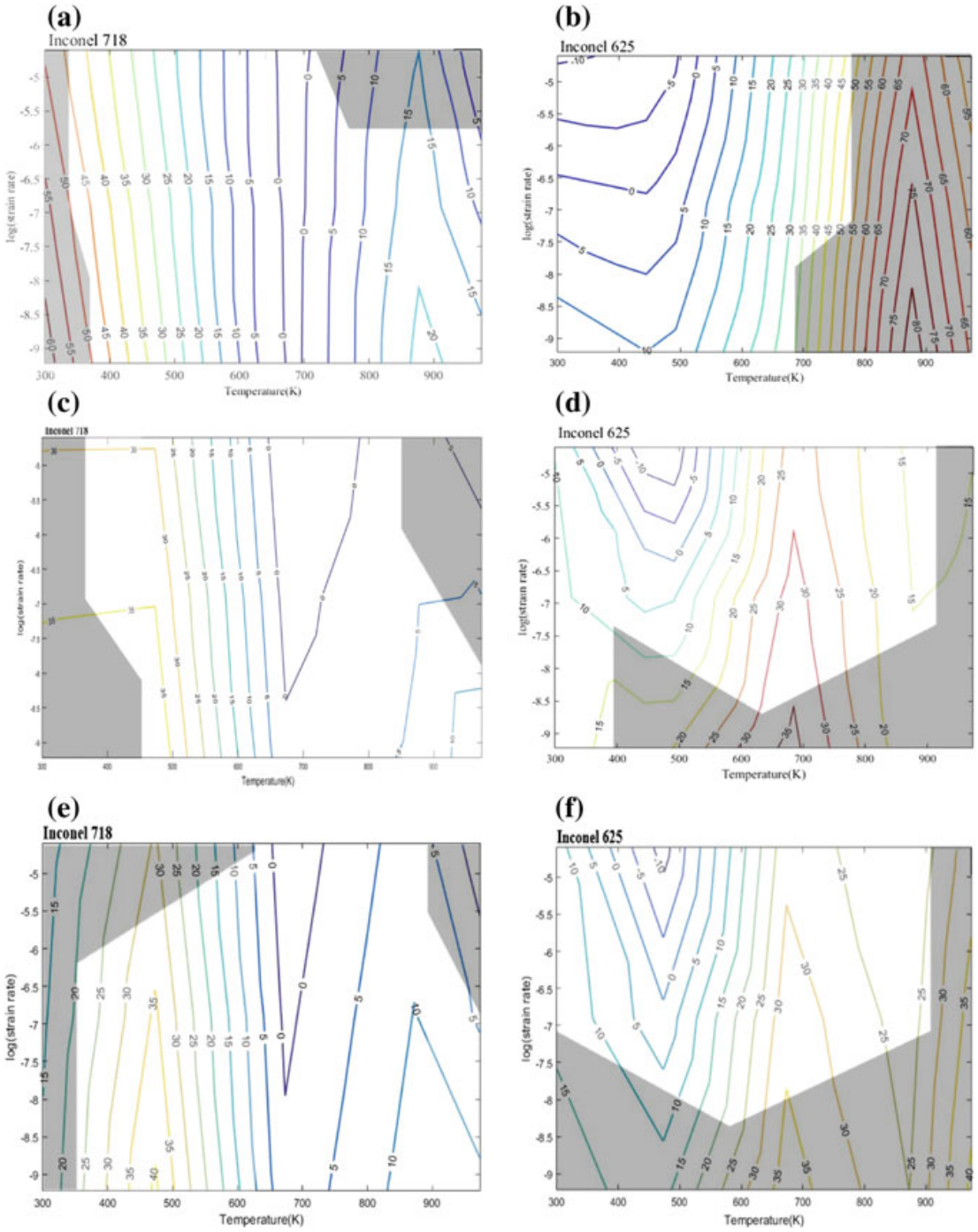


Fig. 10.6 Processing plots for Inconel alloys at **a**, **b** 0.07, **c**, **d** 0.15, and **c** 0.20 strain

0.20 strains is 35 and 38–40%, which is lower than the efficiency at 0.06 strain, i.e. 80 and 60% for Inconel 625 alloy and 718 alloy, respectively. Based on considerations, for the processing of Inconel alloys, 0.06 strain was selected as the optimal parameter condition.

10.4 Conclusions

This study involves workability analysis of Inconel alloy, namely 625 and 718, by processing plots for temperature range of 300–973 K at the strain rates of 0.01, 0.001, and 0.0001 s^{-1}. Following conclusions can be made from the analysis:

- The flow stress and strain hardening behaviors are considerably affected by strain rates and test temperatures in Inconel alloys. The reduction in yield strength observed approx. 31% (From 499.7 to 343.58 MPa) and ultimate strength by approx. 19.8% (from 951.14 to 762.14 MPa), happens with improvement in % elongation from 41.8 to 51.6% from RT to 973 K.
- Processing plots have been established for Inconel alloys at 0.06, 0.15, and 0.20 strains. It is noticed that temperature range ($T < 550$ K) for Inconel 718 and 600–900 K for Inconel 625 for lower strain rate (0.0001 s^{-1}) efficiency were maximum at different considered strain values. Therefore, this condition is considered as optimal parameter processing for the deformation process for Inconel alloys. Further, influence of strain on the instable region is noticed.

Future work includes integration of these established processing plots in finite element analysis of deformation processes of Inconel alloy.

References

1. Special Metals: INCONEL alloy 718. (2007). SMC-066
2. Reed, R.C.: The Superalloys fundamentals and applications. The Superalloys: Fundamentals and Applications. Cambridge University Press (2006). https://doi.org/10.1017/cbo9780511541285
3. Schafrik, R.E., Ward, D.D., Groh, J.R.: Application of alloy 718 in GE aircraft engines: past, present and next five years. In: Superalloys 718, 625, 706 and Various Derivatives (2001). https://doi.org/10.7449/2001/superalloys_2001_1_11
4. Lin, Y.C., Deng, J., Jiang, Y.Q., Wen, D.X., Liu, G.: Effects of initial δ phase on hot tensile deformation behaviors and fracture characteristics of a typical Ni-based superalloy. Mater. Sci. Eng. A **598**, 251–262 (2014). https://doi.org/10.1016/j.msea.2014.01.029
5. Zhang, P., Hu, C., Ding, C.G., Zhu, Q., Qin, H.Y.: Plastic deformation behavior and processing maps of a Ni-based superalloy. Mater. Des. **65**, 575–584 (2015). https://doi.org/10.1016/j.matdes.2014.09.062
6. Rahman, M., Seah, W.K.H., Teo, T.T.: The machinability of Inconel 718. J. Mater. Process. Technol. **63**, 199–204 (1997). https://doi.org/10.1016/S0924-0136(96)02624-6
7. Lin, Y.C., Chen, X.M.: A critical review of experimental results and constitutive descriptions for metals and alloys in hot working. Mater. Des. **32**, 1733–1759 (2011). https://doi.org/10.1016/j.matdes.2010.11.048
8. Liang, R., Khan, A.S.: A critical review of experimental results and constitutive models for BCC and FCC metals over a wide range of strain rates and temperatures. Int. J. Plast. **15**, 963–980 (1999). https://doi.org/10.1016/S0749-6419(99)00021-2
9. Tan, Y.B., Ma, Y.H., Zhao, F.: Hot deformation behavior and constitutive modeling of fine grained Inconel 718 superalloy. J. Alloys Compd. **741**, 85–96 (2018). https://doi.org/10.1016/j.jallcom.2017.12.265

10. Cai, D., Xiong, L., Liu, W., Sun, G., Yao, M.: Development of processing maps for a Ni-based superalloy. Mater. Charact. **58**, 941–946 (2007). https://doi.org/10.1016/j.matchar.2006.09.004
11. Zhang, P., Hu, C., Ding, C., Zhu, Q., Qin, H.: Plastic deformation behavior and processing maps of a Ni-based superalloy. Mater. Des. **65**, 575–584 (2015). https://doi.org/10.1016/j.matdes.2014.09.062
12. Zhang, H., Zhang, K., Lu, Z., Zhao, C., Yang, X.: Hot deformation behavior and processing map of a γ-hardened nickel-based superalloy. Mater. Sci. Eng. A **604**, 1–8 (2014). https://doi.org/10.1016/j.msea.2014.03.015
13. Wen, D.X., et al.: Effects of initial aging time on processing map and microstructures of a nickel-based superalloy. Mater. Sci. Eng. A **620**, 319–332 (2014). https://doi.org/10.1016/j.msea.2014.10.031
14. Sui, F.L., Xu, L.X., Chen, L.Q., Liu, X.H.: Processing map for hot working of Inconel 718 alloy. J. Mater. Process. Technol. **211**, 433–440 (2011). https://doi.org/0.1016/j.jmatprotec.2010.10.015
15. Dieter, G.E., Kuhn, H.A., Semiatin, S.L.: Handbook of Workability and Process Design (2003)
16. Iturbe, A., et al.: Mechanical characterization and modelling of Inconel 718 material behavior for machining process assessment. Mater. Sci. Eng. A **682**, 441–453 (2017). https://doi.org/10.1016/j.msea.2016.11.054
17. Rodriguez, P.: Serrated plastic flow. Bull. Mater. Sci. **6**, 653–663 (1984). https://doi.org/10.1007/BF02743993
18. Caliari, F.R., et al.: Study of the Mechanical Behavior of an Inconel 718 Aged Superalloy Submitted to Hot Tensile Tests. SAE Technical Paper 2011-36-0328, (2011). https://doi.org/10.4271/2011-36-0328
19. Maj, P., et al.: Microstructure and strain-stress analysis of the dynamic strain aging in Inconel 625 at high temperature. Met. Mater. Int. **23**, 54–67 (2017). https://doi.org/10.1007/s12540-017-6264-1
20. Prasad, Y.V.R.K., Rao, K.P., Sasidhara, S.: Hot Working Guide—A Compendium of Processing Maps. ASM International (2015). https://doi.org/10.1016/b978-0-08-033454-7.50019-x
21. Liu, J., Cui, Z., Li, C.: Analysis of metal workability by integration of FEM and 3-D processing maps. J. Mater. Process. Technol. **205**, 497–505 (2008). https://doi.org/10.1016/j.jmatprotec.2007.11.308
22. Prasad, Y.V.R.K.: Processing maps: a status report. J. Mater. Eng. Perform. **12**, 638–645 (2013). https://doi.org/10.1361/105994903322692420
23. Murty, S.V.S.N., Rao, B.N.: On the development of instability criteria during hotworking with reference to IN 718. Mater. Sci. Eng. A **254**, 76–82 (1998). https://doi.org/10.1016/S0921-5093(98)00764-3
24. Sun, Y., Wan, Z., Hu, L., Ren, J.: Characterization of hot processing parameters of powder metallurgy TiAl-based alloy based on the activation energy map and processing map. Mater. Des. **86**, 922–932 (2015). https://doi.org/10.1016/j.matdes.2015.07.140
25. Zhou, G., Ding, H., Cao, F., Zhang, B.: A comparative study of various flow instability criteria in processing map of superalloy GH4742. J. Mater. Sci. Technol. **30**, 217–222 (2014). https://doi.org/10.1016/j.jmst.2013.07.008

Part II
Welding and Joining Technologies

Chapter 11
Investigation on Metallographic Analysis of Electron Beam Ti6Al4V Alloy Welds

Sohini Chowdhury, Yadaiah Nirsanametla and Manapuram Muralidhar

Abstract Electron beam weld quality is evaluated by careful selection of weld parameters and by control of the temporal distribution of electron beam on surface of work piece. Moreover, the integral effect of the electron beam current and weld velocity on weld joint is significant during electron beam welding (EBW) process. In the present work, bead-on-plate EBW of Ti6Al4V alloy plates of 5 mm thickness is carried out to examine the influences of various process parameters with respect to microstructure evolution and weld bead shapes and dimensions. Also, a relationship among input thermal energy, macro- and microstructural characteristics of electron beam Ti6Al4V alloy weldments is established. The experiments were carried out with beam current ranging from 15 to 30 mA and welding velocity from 1200 to 1400 mm min^{-1}. Moreover, a constant focus current and accelerating voltage of 2030 mA and 60 kV are considered. Furthermore, welding modes such as melt-in and keyhole modes are examined with respect to the linear energy and welding process efficiency along with microstructure evolution. In addition, the magnitude of temperature in the heat-affected zone of electron beam welded Ti6Al4V alloy samples is determined using Rosenthal's analytical model at distinct weld regimes. Furthermore, Marangoni and Fourier numbers are employed to investigate weld pool characteristics at distinct weld regimes. Based upon the results, it is observed that electron beam power has a significant influence on weldment profile and penetration level while the welding speed has a noteworthy impact on the solidified structure of fusion zone. Also, three weld regimes were identified under the specified welding conditions. Keyhole mode welding is achieved for linear energy above 77 J mm^{-1}, and melt-in mode welding is attained when linear energy is below 45 J mm^{-1}. Moreover, transition mode welding is identified when the linear energy is in between 56.57 and 75 J mm^{-1}. Furthermore, the efficiency of welding process was assessed by relating electron beam power with fusion zone area. It was recognised that a higher linear energy promoted a higher process efficiency comparative to a lower linear energy. A lower linear

S. Chowdhury · Y. Nirsanametla (✉) · M. Muralidhar
Department of Mechanical Engineering, North Eastern Regional Institute of Science and Technology (NERIST), Nirjuli, Arunachal Pradesh 791109, India
e-mail: ny@nerist.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_11

energy promoted a rich domain of acicular α^1 solidified structure in the fusion zone when compared with higher linear energy. Moreover, the integrity of bead-on-plate welds of Ti6Al4V alloy is of sound quality and without any internal weld defects.

Keywords Electron beam welding · Fourier number · Ti6Al4V alloy · Marangoni number · Microstructure evolution · Rosenthal model

11.1 Introduction

The necessity for lightweight materials has substantially escalated in the fabrication industries on account of critical service aspects. Specifically, titanium alloys are currently implemented in all ranges of applications such as automotive, aerospace and nuclear applications. However, the extensive employment of titanium alloys in critical manufacturing domain can be attributed to excellent properties at high temperature, superior fatigue behaviour, corrosion resistance features and high specific strength [1–3]. In addition, the usage of titanium alloys has now been extended to bio-medical devices because of high bio-compatibility property [4]. α-β two-phase Ti6Al4V alloy is the most commonly used titanium-based alloys and is widely implemented in major domain of manufacturing sector. Earlier, several researchers attempted to join titanium assemblies with numerous joining techniques including fusion welding and solid state joining [5–10]. However, an elemental issue of joining titanium and its alloys is its reaction with atmospheric gases at an elevated temperature which results into embrittlement in the welded specimen [11]. Therefore, the EBW process is the most desirable welding technique to join titanium metal assemblies as it provides a controlled atmosphere to operate. Moreover, EBW offers production of wide span of complicated weld geometries along with an enhanced quality level, flexibility and processing capability. Besides, high energy density electron beam produces welds with minimum heat-affected zone (HAZ) and deep penetration which accounts for low level of distortion in weldments. However, the deep penetration mechanism is associated with vapour cavity formation also referred to as keyhole which is enclosed by high-temperature molten weld pool.

Although, the output of welding operation is prominently established upon keyhole characteristics whose integrity is governed by an equilibrium between recoil pressure in the vapour cavity and surface tension of the weld pool. Hence, weld penetration and weld defects are directly related to keyhole stability that implies; if keyhole collapses, a lower penetration is attained. Besides, weld defects in the form of porosities, spiking defects and several others are generated due to keyhole instability and fluctuations. Moreover, several experimental efforts also provided a qualitative understanding of keyhole dynamics as a function of time [12–14]. Also, the keyhole stability is based on optimisation of weld parameters. Akman et al. demonstrated that control of laser output parameters can efficiently control weld penetration and weld bead width in laser welding process

[15]. Squillace et al. [16] reported that the under fill defects in Ti6Al4V alloy laser welds depends on a definite heat input. The authors also determined that the under fill defect is maximum in keyhole mode welds as compared to conduction mode welds. Li et al. studied the influence of electron beam weld parameters on weld profile and weld distortion of titanium-based Ti2AlNb alloy plates [17]. Liu et al. conducted experiments to determine the welding stresses induced after electron beam welding and milling operations from a thick titanium metal plate of 50 mm thickness [18]. Also, Junaid et al. studied the effect of beam parameters on titanium-based Ti5Al2.5Sn alloy sheets with respect to weld bead width and mechanical characteristics. However, from the esteemed reports, it is recognised that EBW of Ti6Al4V alloy is limited. Thus, a comprehensive analysis of EBW of Ti6Al4V alloy can lay a good contribution to the prevailing approaches.

In this work, electron beam welding experiments were performed on 5 mm thick Ti6Al4V alloy plates and weld seams are evaluated in terms of weld geometry and microstructure. The aim of the present work is to establish a relationship between input thermal energy, weld morphology and microstructural characteristics for electron beam welded Ti6Al4V alloy at distinct weld regimes. Moreover, different weld regimes are characterised for Ti6Al4V alloy samples under different welding conditions corresponding to Marangoni number and Fourier numbers. Also, the temperature distribution in HAZ of Ti6Al4V alloy is determined by employing Rosenthal's analytical model.

11.2 Experimental Investigation

In the present work, Ti6Al4V alloy in annealing condition is considered for experimental investigation. The electron beam bead-on-plate welds were conducted on 5 mm thick Ti6Al4V alloy, and the plate dimension (length × width) is 100 mm 50 mm. Prior to welding operations, all the sample surfaces were properly cleaned with acetone and clamped carefully by fixtures to prevent distortion during welding operation. However, to ensure high thermal transfer rate during joining process, the backing plate in fixture is assembled with copper material.

The weld tests were carried out for a suitable combination of weld speed and electron beam power where a stable weld seam is acquired. The different welding conditions that were adapted during EBW of Ti6Al4V alloy are listed in Table 11.1. The experiments were conducted by varying weld speed and electron beam current at fixed accelerating voltage and focusing condition. The accelerating voltage level was maintained constant at 60 kV for all the samples, while the beam current was varied between 15 and 30 mA and weld speed was ranged between 1200 and 1400 mm/min. Furthermore, for evaluation of weld properties, the metallographic examination was conducted. The specimens were sectioned at middle segment across the weld length, grounded and chemically etched with Keller's reagent solution. Thereafter, the weld cross sections and microstructure of the polished weld specimen were acquired with optical microscope.

Table 11.1 Electron beam welding conditions of Ti6Al4V alloy

Weld sample	Accelerating voltage (kV)	Beam current (mA)	Focus current (mA)	Weld speed (mm/min)
1	60	15	2030	1400
2	60	15	2030	1200
3	60	22	2030	1400
4	60	22	2030	1200
5	60	25	2030	1400
6	60	25	2030	1200
7	60	30	2030	1400
8	60	30	2030	1200

To characterise the weld geometry and distinguish weld regimes of Ti6Al4V alloy weldment, linear energy is considered. The linear energy (L_e) which is the ratio of power transferred to the weld speed is represented as

$$L_E = VI/u \tag{11.1}$$

where V, I and u indicate accelerating voltage (kV), electron beam current (mA) and weld speed (mm/s), respectively. Moreover, the process efficiency is quantified by fundamental welding process efficiency (W_E) parameter which correlates the transmitted beam power with fusion zone area. It takes into account the energy required for creating the joining or welding phenomena in the material. Hence, the parameter is represented as

$$W_E = A/L_e \tag{11.2}$$

where A indicates area of the fusion zone (mm^2).

In order to demonstrate thermal and fluid flow characteristics of Ti6Al4V alloy welds, Marangoni number and Fourier numbers are employed. Marangoni number is represented as a ratio of surface tension force to viscous force, and it is mathematically represented as

$$M_a = \frac{d\gamma}{dT} \times \frac{w\Delta T}{\mu\alpha} \tag{11.3}$$

Fourier numbers is defined as a ratio of thermal transmission rate to thermal storage rate and mathematically given as

$$F_o = \frac{\alpha}{ul} \tag{11.4}$$

where $d\gamma/dT$, w, μ, α, ΔT represent surface tension gradient (N m^{-1} K^{-1}), width of molten pool (mm), viscosity (kg m^{-1} s^{-1}), thermal diffusivity (m^2 s^{-1}), temperature gradient (K).

11.3 Results and Discussion

In this section, the macro- and microstructural characteristics of Ti6Al4V alloy weld joint are characterised for a wide range of linear energy varying from 38.57 to 90 J mm^{-1}. The estimation of temperature is also carried out using Rosenthal's analytical model. Moreover, weld pool behaviour is identified by Fourier numbers and Marangoni number at distinct weld regimes.

The effect of process variables such as beam current and welding speed on weld morphology and microstructural characteristics is examined. Figure 11.1 demonstrates the appearance of bead-on-plate Ti6Al4V alloy weld joint corresponding to weld sample #1 at 38.57 J/mm linear energy. It is observed that a smooth, bright, clean weld bead is formed with no oxide formation which indicates the effectiveness of electron beam welding process.

Moreover, the seam morphology of Ti6Al4V alloy weldments corresponding to weld sample #2 and #8 at a constant welding velocity of 1200 mm min^{-1} is presented in Fig. 11.2. Different weld profiles and weld regimes are identified under different linear energy values. The macrograph displays an elliptical profile at linear energy of 45 J/mm (Fig. 11.2a) while a nail head profile (Fig. 11.2b) is attained at linear energy of 90 J/mm. The transition from an elliptical profile to nail head profile is attributed to an increase in linear energy which increases the weld

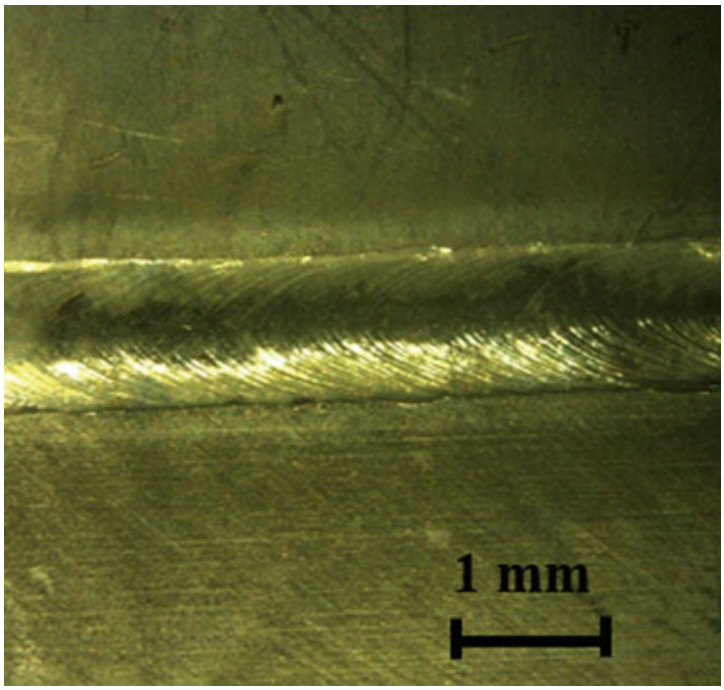


Fig. 11.1 Appearance of the weld surface at linear energy 38.57 J/mm

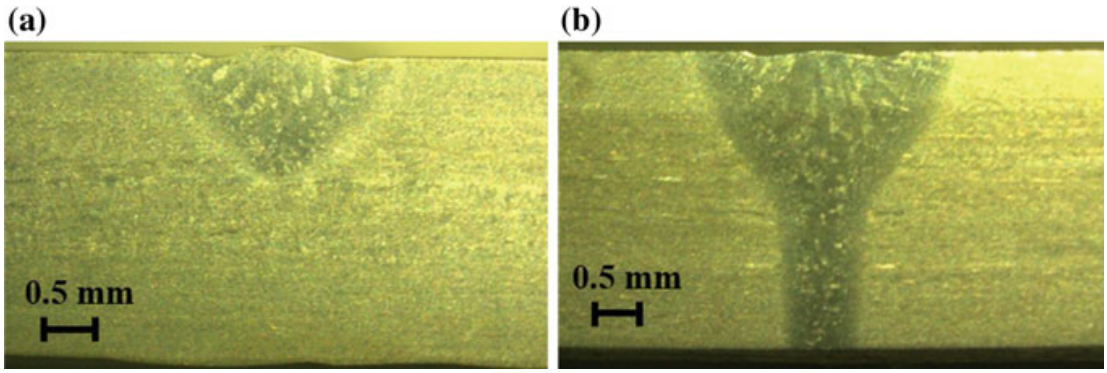


Fig. 11.2 Cross section of weld samples **a** #2 and **b** #8 at constant weld speed of 1200 mm/min

penetration considerably. Also, the cross-sectional macrograph of Ti6Al4V alloy weldments displays a conduction mode welding (Fig. 11.2a) at 45 J/mm and a keyhole mode welding (Fig. 11.2b) at 90 J/mm. This indicates that conduction mode welding is attained for linear energy values at or below 45 J/mm, and keyhole mode welding is attained for linear energy values at or above 77 J/mm. In addition, transition mode welds are also observed between linear energy values of 56.57 and 75 J/mm. Besides, full penetration welds are desirable in most applications due to high joint strength acquired. Moreover, the macrograph reveals no major weld defects such as porosity and spiking defects.

Table 11.2 depicts fusion zone area, linear energy and welding process efficiency values of weld samples. Weld sample 2 reveals a higher welding efficiency than weld sample 1. This is probably due to the fact that the energy spent in the form of heat conduction loss and overheating is higher than energy spent to increase fusion zone area of weld sample 1. Conversely, for weld sample 1, more energy is spent on increase in fusion zone area than thermal losses. Therefore, it can be concluded that lower weld speed or higher linear energy promotes higher process efficiency than higher weld speed or lower linear energy at a constant electron beam power. Figure 11.3 illustrates the cross-sectional macrographs of weld samples under varying linear energy conditions. Figure 11.3a, b corresponds to weld sample

Table 11.2 Efficiency parameters of electron beam welded Ti6Al4V alloy

Weld sample	Weld area (mm^2)	Linear energy ($J\ mm^{-1}$)	Welding process efficiency ($mm^3\ J^{-1}$)
1	3.53	38.57	0.078
2	5.972	45	0.154
3	8.58	56.57	0.135
4	9.43	66	0.166
5	10.422	64.28	0.138
6	11.2	75.75	0.174
7	12.541	77.14	0.139
8	14.25	90	0.184

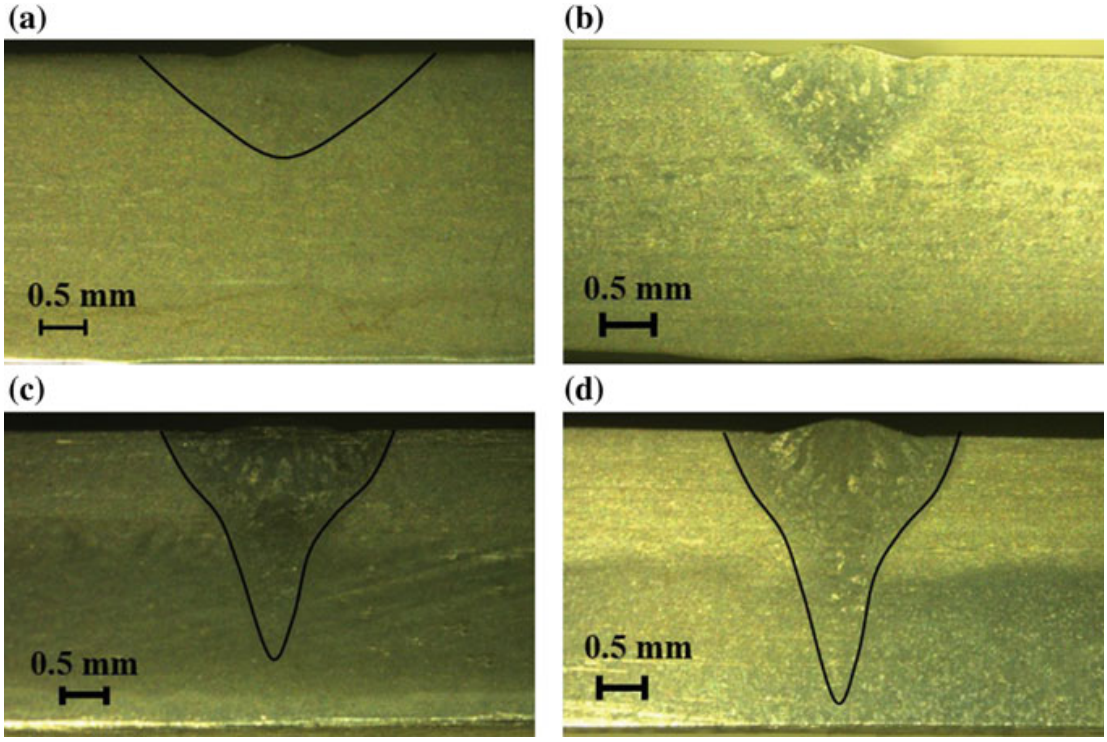


Fig. 11.3 Cross-sectional macrographs of weld samples corresponding to linear energy, **a** 38.57 J mm^{-1}, **b** 45 J mm^{-1}, **c** 75.75 J mm^{-1} and **d** 64.28 J mm^{-1}

1 and 2, respectively, which refers to a decrease in weld speed from 1400 to 1200 mm/min at constant beam current of 15 mA. Similarly, Fig. 11.3c, d corresponds to weld sample 5 and 6, respectively, which refers to a decrease in weld speed from 1400 to 1200 mm/min at a constant beam current of 25 mA. The reduction in weld penetration and an increase in weld width are identified at higher weld speed. Moreover, with an increase in beam current, penetration level in weldment is increased remarkably.

The heat transfer mode and material flow characteristics in the molten metal during welding operation determine the weld joint quality. Moreover, the prediction of temperature using experimental methods is critical during the welding process since molten weld pool width is quite small and there is a possibility of plasma formation. Hence, an analytical model is employed in the present work for estimation of the temperature field in the HAZ analytically. Also, with the purpose of addressing the issue of heat transfer mode in electron beam weldments, Marangoni number [19] and Fourier numbers [20] are employed.

In the present work, Rosenthal's analytical model is employed, since it takes into account the integral effect of processing parameters and thermal properties of Ti6Al4V alloy for prediction of temperature analytically [21, 22]. The positions for estimation of peak temperature in transverse direction of welding line are shown in Fig. 11.4a. The estimated temperature values corresponding to welding conditions are given in Table 11.1 and presented in Fig. 11.4b. It can be identified that the magnitude of temperature in HAZ of weld plates is much lower than liquidus

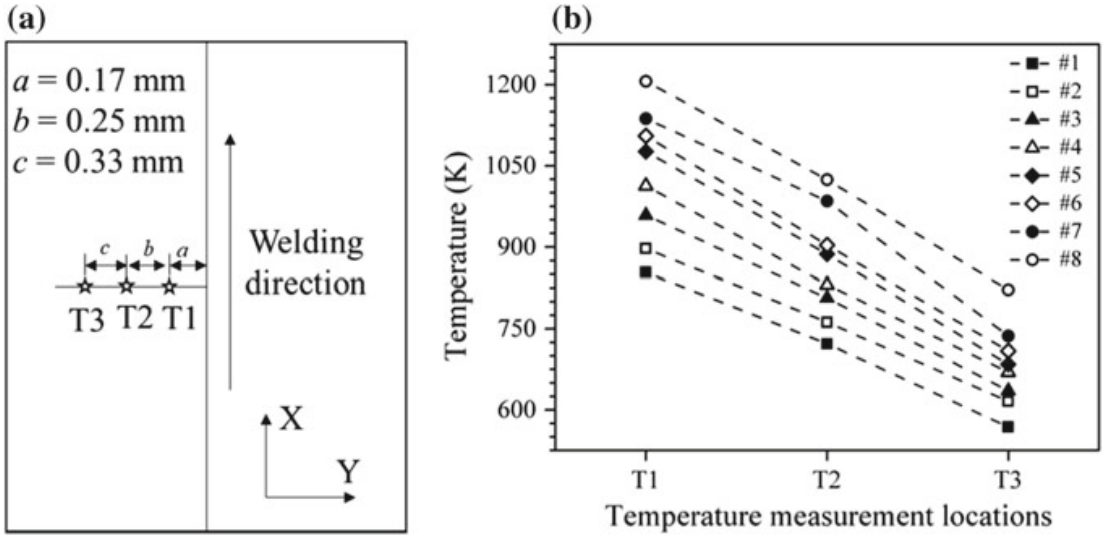


Fig. 11.4 **a** Temperature measurement locations along the welding line and **b** estimated temperature values corresponding to processing conditions of Table 11.2

temperature. Also, temperature at location '*T*1' is determined to be higher than '*T*2' and '*T*3' locations. Since '*T*1' corresponds to a location near FZ while, '*T*2' and '*T*3' correspond to the locations away from the melt zone. Moreover, it is determined that conduction weld (#2) exhibited a lower temperature value when compared with keyhole weld (#7). Furthermore, it is determined that the percentage difference in magnitude of temperature at '*T*1' location for conduction and keyhole mode is 26.0.

Figure 11.5 presents the variation of Marangoni number (Ma) and Fourier numbers as a function of linear energy corresponding to electron beam Ti6Al4V alloy welds. It is observed that *Ma* and *Fo* increases and reduces linearly with increases in magnitude of linear energy, respectively. Moreover, three distinct weld regimes are separated with respect to *Ma* and *Fo* values. Fundamentally, higher *Ma* value indicates higher molten metal velocity at the surface and convective heat

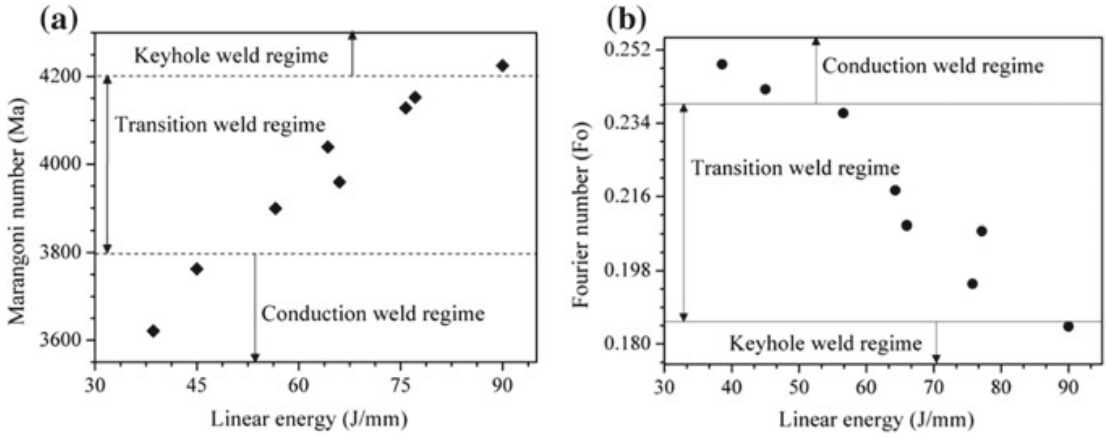


Fig. 11.5 Influence of linear energy on **a** Marangoni number and **b** Fourier number

transmission in the weld pool dictates the process. From Fig. 11.5a, it can be observed that the conduction weld regime exhibits lower Marangoni number when compared with transition and keyhole weld regime. This reflects that conduction welds are subjected to low molten metal velocity at the surface and produces a more stable weld bead with lower aspect ratio. The Fourier numbers is associated with solidification rate of crystals at the FZ which eventually accounts for thermo-mechanical damage in the material. Higher Fourier numbers is always preferable in weldments, since it indicates lower thermal storage rate and lower characteristics length of the molten pool on account of high solidification rate. From Fig. 11.5b, it is observed that conduction weld regime exhibited higher Fourier numbers than transition and keyhole welds. Thus, conduction welds are subjected to high solidification rate and low thermo-mechanical damage relative to transition and keyhole welds.

The microstructure of parent metal of Ti6Al4V alloy primarily consists of body-centred cubic (BCC) β phase distributed in an area of hexagonal close-packed α phases. However, during welding operation, due to rapid heating and cooling process, transformation from α to β phase and decomposition of β to α phase occurs at the fusion zone, respectively, in a short duration of time. This entire heating and cooling operation depends on heating and cooling rate of Ti6Al4V alloy. In case cooling rate is greater than critical cooling rate of Ti6Al4V alloy (around 623 K/s), then microstructure evolution from β to α^1 martensite structure is promoted in the weld zone. Similarly, at lower cooling rate, secondary lamella growth is promoted due to diffusion-controlled nucleation [23]. However, β phases can be traced out in the fusion zone irrespective of cooling rate of Ti6Al4V alloy weldments. The microstructure of fusion zone region of weld samples #1 (Fig. 11.6a) and #8 (Fig. 11.7a) displays similar microstructure at the weld zone comprising of

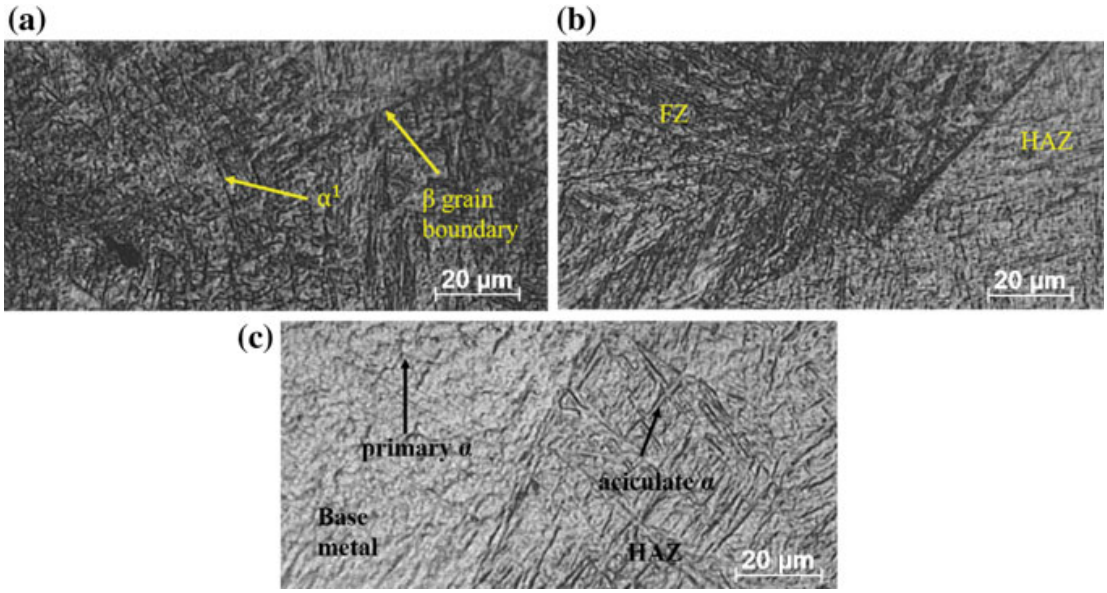


Fig. 11.6 Illustartion of microstructure at different zones for Ti6Al4V alloy weldment at 38.57 J/mm

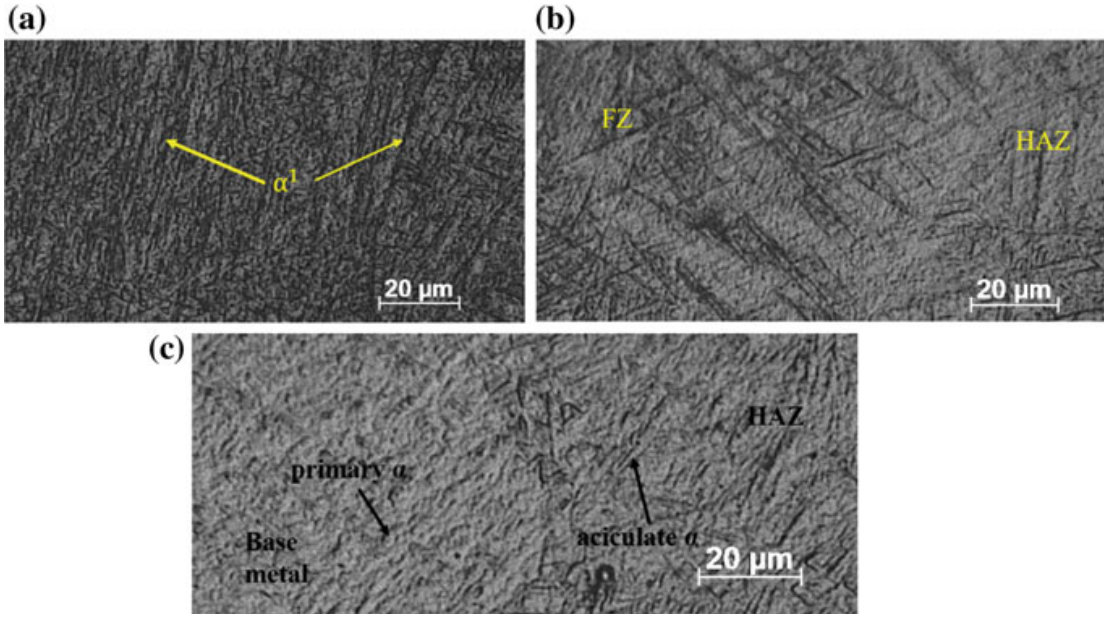


Fig. 11.7 Illustartion of microstructure at different zones for Ti6Al4V alloy weldment at 90 J/mm

martensitic α^1 solidified structure within the β grains. The solidified microstructure α^1 is evolved when β phase is quenched above beta transus temperature at a cooling rate above the critical cooling rate of Ti6Al4V alloy [24]. Moreover, the length and distribution of martensitic structures in the FZ vary in accordance with the solidification rate. It has been determined that the size of α^1 grains is higher in keyhole weld regime relative to the conduction weld regime. However, the percentage content of martensitic structure reduces with increase in linear energy. Therefore, it can be inferred that the fusion zone of weld samples #1 and #8 does not differ much by cooling rate. Although, the amount, distribution and size of acicular α^1 solidified structure at each weld zone differ comparatively corresponding to weld speed. Also, the cooling rate of Ti6Al4V alloy weldment is comparatively enhanced due to implementation of copper as backing plate in fixture [25].

Figures 11.6c and 11.7c display the different weld zones of Ti6Al4V alloy corresponding to weld samples #1 and #8 at 1400 mm/min and 1200 mm/min weld speed, respectively. The heat-affected zone (HAZ) displays a much more complex microstructure as it consists of primary α grains, martensitic α^1 and β grains. Moreover, due to temperature gradient in the Ti6Al4V alloy specimen, the microstructure in the HAZ varied from a rich α^1 (near fusion zone) to a poor α^1 (far from fusion zone) [26]. However, the microstructure in distinct weld zones is directly proportional to the linear energy. A higher linear energy leads to lower cooling rate which induces diffusion-controlled phase transformation. Conversely, a lower linear energy promotes a rich domain of acicular α^1 solidified structure in the fusion zone.

11.4 Conclusion

This paper has presented detailed experimental results of electron beam welding process of 5 mm thick Ti6Al4V alloy specimen. Based on the present investigation, the following conclusions are derived.

- The linear energy which is expressed as power transmitted to weld speed is an influential parameter to determine weld profiles and weld regimes of Ti6Al4V alloy specimen.
- A lower weld speed promotes higher process efficiency than higher weld speed.
- The temperature distribution in HAZ of Ti6Al4V alloy weldments is lower in conduction weld regime than keyhole weld regime.
- The fusion zone of electron beam Ti6Al4V alloy weldments primarily consists of martensitic α^1 structures with small amount of β grains.
- The distribution, amount and size of martensitic α^1 and β grains are influenced by heat input and cooling rate.
- Also, a higher linear energy leads to lower cooling rate which induces diffusion-controlled phase transformation.
- The conduction weld regime exhibited lower martensitic size than keyhole weld regime. However, the density of martensitic structures in the FZ is determined to be more in the conduction weld regime.
- The magnitude of *Ma* number is determined to be lower in conduction weld regime. However, *Fo* number value is relatively higher in conduction weld regime than keyhole weld regime.

Acknowledgements This work was financially supported by second phase of Technical Education Quality Improvement Programme (TEQIP), India and North Eastern Regional Institute of Science and Technology (NERIST). We also thank Hindustan Aeronautics Limited, Koraput, Odisha, India for providing experimental facility.

References

1. Veiga, C., Davim, J.P., Loureiro, A.J.R.: Properties and applications of titanium alloys: a brief review. Adv. Mater. Sci. **32**, 133–148 (2012)
2. Henriques, V.A.R.: Titanium production for aerospace applications. J. Aerosp. Technol. Manag. **1**(1), 7–17 (2009)
3. Peters, M., Kumpfert, J., Ward, C.H., Leyens, C.: Titanium alloys for aerospace applications. Adv. Eng. Mater. **5**(6), 419–427 (2003)
4. Elias, C.N., Lima, J.H.C., Valiev, R., Meyers, M.A.: Biomedical applications of titanium and its alloys. Biol. Mater. Sci. Overv. **60**(3), 46–49 (2008)
5. Saresh, N., Gopalakrishna, P.M., Mathew, J.: Investigation into the effects of electron beam welding on thick Ti-6Al-4V titanium alloy. J. Mater. Process. Technol. **192–193**, 83–88 (2007)
6. Lathabai, S., Jarvis, B.L., Barton, K.J.: Comparison of keyhole and conventional gas tungsten arc welds in commercially pure titanium. Mater. Sci. Eng.: A **299**, 81–93 (2001)

7. Greenwell, T., Ramulu, M., Labossiere, P.: Characterization of tensile behavior in friction stir-welded titanium alloy, Ti-6Al-4V. In: Proceedings of the SEM Annual Conference and Exposition on Experimental and Applied Mechanics, pp. 2659–2668 (2009)
8. Mironov, S., Sato, Y.S., Kokawa, H.: Development of grain structure during friction stir welding of pure titanium. Acta Mater. **57**(45), 19–28 (2009)
9. Balasubramanian, M., Jayabalan, V., Balasubramanian, V.: Effect of pulsed gas tungsten arc welding on corrosion behavior of Ti–6Al–4V titanium alloy. Mater. Des. **29**, 1359–1363 (2008)
10. Li, R., Li, Z., Zhua, Y., Ronga, L.: A comparative study of laser beam welding and laser–MIG hybrid welding of Ti–Al–Zr–Fe titanium alloy. Mater. Sci. Eng.: A **528**, 1138–1142 (2011)
11. Donachie. M.J.: Titanium, A. Technical Guide, 2nd edn. ASM International, Metals Park (2000)
12. Tong, H., Giedt, W.H.: Radiographs of the electron beam welding cavity. Rev. Sci. Instrum. **40**(10), 1283–1285 (1969)
13. Mara, G. L.: A study of the penetration mechanism of electron beam welding and the spiking phenomenon. Ohio State University, Master Thesis (1971)
14. Arata, Y., Matsuda, F., Murakami, T.: Some dynamic aspects of weld molten metal in electron beam welding. Trans. Join. Weld. Res. Inst. 152–161 (1973)
15. Akman, E., Demir, A., Canel, T., Sınmazcelik, T.: Laser welding of Ti6Al4V titanium alloys. J. Mater. Process. Technol. **209**, 3705–3713 (2009)
16. Squillace, A., Prisco, U., Ciliberto, S., Astarita, A.L.: Effect of welding parameters on morphology and mechanical properties of Ti-6Al-4V laser beam welded butt joints. J. Mater. Process. Technol. **212**, 427–436 (2012)
17. Li, Y., Zhao, Y., Li, Q., Wu, A., Zhu, R., Wang, G.: Effects of welding condition on weld shape and distortion in electron beam welded Ti2AlNb alloy joints. Mater. Design **114**, 226–233 (2017)
18. Chuan, L., Jianxun, Z., Bing, W., Shuili, G.: Numerical investigation on the variation of welding stresses after material removal from a thick titanium alloy plate joined by electron beam welding. Mater. Des. **34**, 609–617 (2012)
19. Robert, A., Debroy, T.: Geometry of laser spot welds from dimensionless numbers. Metall. Mater. Trans. B **32**, 941–947 (2001)
20. Mukherjee, T., Manvatkar, V., De, A., DebRoy, T.: Dimensionless numbers in additive manufacturing. J. Appl. Phys. **121**, 064904 (2017)
21. Rosenthal, D.: Mathematical theory of heat distribution during welding and cutting. Weld. J. **20**(5), 220–234 (1941)
22. Rosenthal, D.: The theory of moving sources of heat and its application to metal treatments. Trans. Am. Soc. Mech. Eng. **68**, 849–866 (1946)
23. Mayuri, B., Swarup, B.: Influence of heat Input in microwelding of titanium alloy by microplasma arc welding. J. Mater. Process. Technol. **231**, 100–112 (2016)
24. Fan, Y., Yang, Z., Cheng, P., Egland, K., Yao, Y.L.: Effects of phase transformations on laser forming of Ti-6Al-4V Alloy. J. Appl. Phys. **98**(1), 013518 (2005)
25. Ahmed, T., Rack, H.J.: Phase transformations during cooling in $\alpha + \beta$ titanium alloys. Mater. Sci. Eng.: A **243**, 206–211 (1998)
26. Giuseppe, C., Michelangelo, M., Sabina, L.C.: Ytterbium fiber laser welding of Ti6Al4V alloy. J. Manuf. Process. **20**(1), 250–256 (2015)

Chapter 12
Multi-objective Optimization of FSW Process Variables of Aluminium Matrix Composites Using Taguchi-Based Grey Relational Analysis

Subramanya R. B. Prabhu, Arun Kumar Shettigar, Mervin A. Herbert and Shrikantha S. Rao

Abstract Successful joining of aluminium alloys using friction stir welding (FSW) opens a new window research in extending this technique to join aluminium matrix composites (AMCs). Current research is focused on optimization of process variables for multiple responses simultaneously. Experiments were performed using tool pin profile, rotational speed (RS) and welding speed (WS) as ideal process variables for multi-objective optimization in FSW of AMCs. Tensile strength, macro-hardness and elongation are considered as multi-response behaviour. Grey relational grade for the chosen multiple responses are obtained using grey analysis. Analysis of variance was utilized to understand the influence of process variables on the grey relational grade. Analysis reveals that RS and WS were the most influencing process variables on the output responses. Confirmation experiments were performed at optimized process variables to validate the present study. Predicted values were in good agreement with the experimental results.

Keywords Friction stir welding · Grey relational analysis · Tensile strength · Macro-hardness · Elongation

S. R. B. Prabhu (✉) · M. A. Herbert · S. S. Rao
Department of Mechanical Engineering, National Institute of Technology Karnataka, Surathkal, India
e-mail: subramanya.prabhu@gmail.com

S. R. B. Prabhu · A. K. Shettigar
Department of Mechatronics Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_12

12.1 Introduction

FSW is a novel technique invented by The Welding Institute (TWI, UK) in 1991 to join difficult to weld materials like aluminium and its alloys [1]. FSW is a solid-state welding process where materials are joined by plasticizing and forging action rather than melting and recasting [2]. FSW offers several benefits over conventional fusion welding process. A non-consuming rotating tool alters the material flow and creates a defect-free weld during FSW. The nature of material flow is affected by the tool pin profile as well as process variables [3]. Cavaliere et al. [4] performed FSW of AA6082 alloys and studied the effect of process variables on microstructure and mechanical properties. Rajakumar et al. [5] analysed the influence of process variables such as rotational speed (RS), welding speed (WS), axial load, tool pin profiles and hardness of tool on the mechanical properties of FS-welded AA 7075 alloy. Malopheyev et al. [6] worked on optimization of FSW of 6061-T6 aluminium alloy processing variables and stated that increasing the welding speed hindered precipitation coarsening which improved after ageing. Literature study reveals that most of the study focused on the influence of tool profile and process variables on microstructure and tensile properties. Studies indicate that the quality of the weld mainly depends on various process variables such as shoulder diameter, axial load, RS and WS. To predict the mechanical properties of FS-welded AA 6061, Elangovan et al. [7] created a mathematical model by utilizing process variables like WS, RS, tool geometry and axial load. Murugan and Ashok Kumar [8] also developed similar model to predict tensile strength of friction stir welded stir cast AA6061-T6/AlNp composite incorporating RS, WS, axial load and tool profile as process variables. Artificial neural network (ANN) was used by Ukuyucu et al. [9] to get correlation among mechanical properties of FS-welded aluminium plates and FSW process variables. Subramanya et al. [10] studied multi-response optimization of friction stir welding process variables, using TOPSIS approach to optimize FSW of aluminium matrix composites. Multi-response base on orthogonal array (OA) coupled with grey relational analysis was used by Vijayan et al. [11] to optimize the FSW process variables to join aluminium alloy AA5083 by FSW. Deepandurai and Parameshwaran [12] performed multi-response optimization of FSW variable for AA7075/SiCp using response surface methodology and fuzzy-based grey relational analysis. Comprehensive literature findings indicate that most of the reports allied to the optimization of FSW variables have utilized one response at a time. However, meagre study had been performed on the influence of variables on the multiple responses simultaneously. The present work is focused on optimization of variables while FSW of AMCs for macro-hardness, tensile strength and % elongation using grey relational analysis.

12.2 Experiments

12.2.1 Process Variables Selection

FSW was performed on a conventional vertical milling machine (Make: Bharat Fritz Werner). The material used in this study was AA6061-4.5(wt%) Cu-5(wt%) SiC composite. These composites were prepared by stir casting technique. Composite plates having size of 100 × 50 × 6 mm were prepared from the stir cast composite by machining. FSW tool was prepared using M2 steel with a hardness of 60 HRC. Three types of tool geometries were used in the present study, namely square (SQ), threaded cylindrical (TC) and combined threaded cylindrical and square (CTS). Tool pin geometries are shown in Fig. 12.1.

Set-up used for FSW process is shown in Fig. 12.2. Process was carried out by altering the process variable values. RS, WS and tool pin geometries are chosen as process variables to optimize the FSW of AMCs. Other variables such as tool shoulder diameter of 18 mm and tool tilt angle of 2.50 are maintained constant in order to limit the study. Working range of process variables is fixed by conducting the trial experiments which gives defectless joints. It has been observed that process variables outside the working range yields defects in the joints, such as crack and tunnel hole. The number of experiments for the present study was considered based on Taguchi's design of experiment method. As per Taguchi's method, the degree of

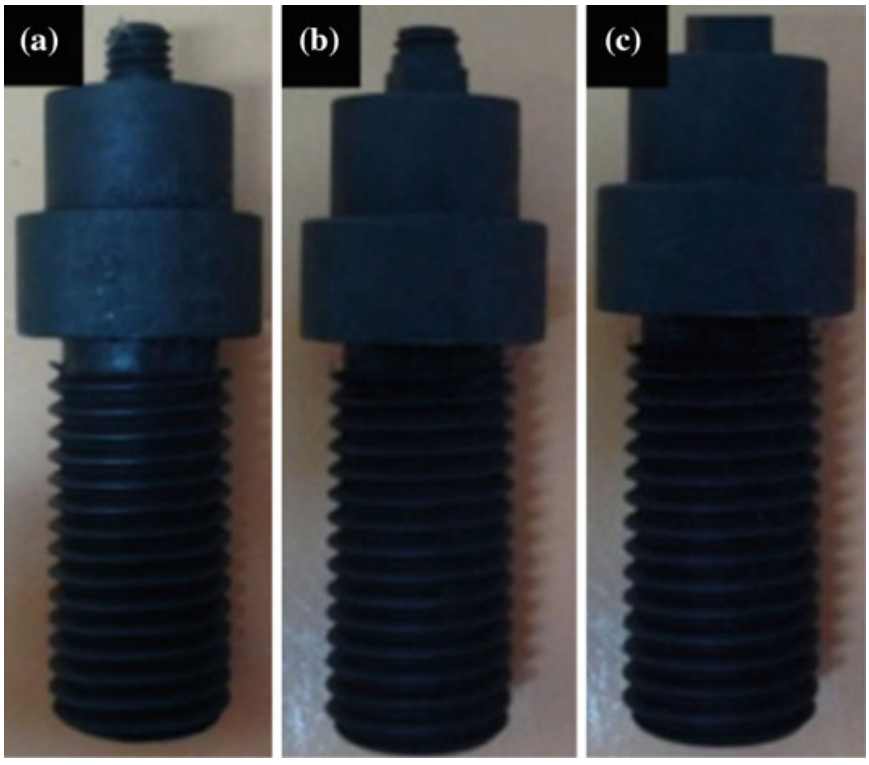


Fig. 12.1 FSW tools with **a** threaded cylindrical (TC), **b** combined threaded cylindrical and square (CTS) and **c** square (SQ) profiled pin

Fig. 12.2 FSW set-up

Table 12.1 FSW process variables with their levels

Factors	Process variables	Level 1	Level 2	Level 3
A	Tool geometry	TC	SQ	CTS
B	Rotational speed (RS)	700	1000	1300
C	Traverse speed (WS)	60	70	80

freedom (DOF) of selected OA must be equal to or higher than the total DOF required for the experiment. In the present work, L9 OA was used to carry out the experiments. Process variables used in the present study and their three levels were shown in Table 12.1.

12.2.2 Evaluation of FSW Performances

Specimens for tensile tests were prepared as per ASTM E8 standard from both welded and the base material. Specimens were cut from the workpiece in a direction normal to the weld path. Three specimens were prepared from each of the welded parts and average of these values was considered as parametric values. Ultimate tensile strength and % elongation values were measured by performing test on universal testing machine.

The hardness test was performed across the weld at the middle region at 3 mm gap on both sides of the weld path. Test was performed on Vickers macro-hardness tester by choosing load of indentation of 5 kg for 15 s. Table 12.2 lists the FSW process variables used in the study and corresponding output responses.

Table 12.2 FSW process variable values and output responses

Expt. No.	Process variables			Output responses		
	Tool profile	RS (RPM)	WS (mm/min)	UTS (MPa)	Vickers hardness (Hv)	% Elongation
1	TC	700	60	118	102	8.6
2	TC	1000	70	158	118	7.4
3	TC	1300	80	135	112	8.1
4	SQ	700	70	144	114	8.8
5	SQ	1000	80	166	123	7.9
6	SQ	1300	60	129	107	9.1
7	CTS	700	80	145	116	8.4
8	CTS	1000	60	150	118	8.2
9	CTS	1300	70	146	126	8.6

12.3 Grey Relational Analysis

The Taguchi method is an important tool for enhancing the productivity and quality during research and development. Taguchi offers meticulous way of design and analysis of experiment to improve the productivity. Most of the Taguchi techniques were on optimization of single responses. However, combining grey relational analysis (GRA) with the Taguchi methods enables the optimization of process variables for two or more than two responses simultaneously [13]. In GRA, grey relational coefficients for various responses are computed and average of these are called grey relational grade (GRG). These GRGs are considered as single-output response for the Taguchi experiment. In the present study, GRA-based Taguchi method is used to optimize the FSW variables for the output responses namely macro-hardness, tensile strength and % elongation.

12.3.1 Response Pre-processing

Since the unit and range of one response differ from the others, pre-processing of responses were necessary in GRA. When the scatter range of response is too large or when target direction of the response is different, then also response pre-processing is to be carried out in GRA. Response pre-processing is a process of converting original responses into comparable values. Responses are to be normalized in the range of zero to one to make them comparable values. In the present study, all the three responses namely tensile strength, macro-hardness and % elongation, 'larger the better' criteria were selected, as quality of the weld was represented by higher values of these responses. The original values of responses were normalized as follows for the 'larger the better' criteria.

$$p_i^*(k) = \frac{p_i(k) - \min p_i(k)}{\max p_i(k) - \min p_i(k)} \tag{12.1}$$

where $p_i(k)$ and $p_i^*(k)$ are the values before and after response pre-processing, where i represents the experiment number and k represents the response. Normalized values after pre-processing were listed in Table 12.3.

Further, $\Delta_{oi}(k)$ represents the deviation of the normalized value $p_i^*(k)$ from the reference value $p_0^*(k)$. Calculation was done for all the experiments and responses and listed in Table 12.4.

$$\Delta_{oi}(k) = \left|p_0^*(k) - p_i^*(k)\right| \tag{12.2}$$

Table 12.3 Normalized value of responses after pre-processing

Expt. No.	UTS (MPa)	Vickers hardness (Hv)	% Elongation
1	0.0000	0.0000	0.7059
2	0.8333	0.6667	0.0000
3	0.3542	0.4167	0.4118
4	0.5417	0.5000	0.8235
5	1.0000	0.8750	0.2941
6	0.2292	0.2083	1.0000
7	0.5625	0.5833	0.5882
8	0.6667	0.6667	0.4706
9	0.5833	1.0000	0.7059

Table 12.4 Deviation sequence of each response

Expt. No.	UTS (MPa)	Vickers hardness (Hv)	% Elongation
1	1.0000	1.0000	0.2941
2	0.1667	0.3333	1.0000
3	0.6458	0.5833	0.5882
4	0.4583	0.5000	0.1765
5	0.0000	0.1250	0.7059
6	0.7708	0.7917	0.0000
7	0.4375	0.4167	0.4118
8	0.3333	0.3333	0.5294
9	0.4167	0.0000	0.2941

12.3.2 Calculating Grey Relational Coefficient (GRC) and Grey Relational Grade (GRG)

A grey relational coefficient (GRC) was calculated using the values obtained after response pre-processing. GRC formulates the relationship among the actual normalized response values with the ideal values. The GRC can be calculated as follows.

$$\delta_i(k) = \frac{\Delta_{\min} + \delta \cdot \Delta_{\max}}{\Delta_{oi}(k) + \delta \cdot \Delta_{\max}} \tag{12.3}$$

where $\Delta_{oi}(k)$ is the deviation of the normalized value $p_i^*(k)$, and from the reference value $p_0^*(k)$, δ is a coefficient of identification. The value of δ is taken as 0.33 by considering the equal preference for all the variables. The GRG for each experiment were calculated and tabulated in Table 12.5.

After calculating GRC, grey relational grade was obtained by taking the average of GRC corresponding to each experiment. The multiple response optimization of the process variables are based on the GRG.

$$\beta_i = \frac{1}{n}\sum_{k=1}^{n} \delta_i(k) \tag{12.4}$$

where $\delta_i(k)$ is the GRG for the *i*th experiment which vary from 1 to 9 and n is the number of responses. Table 12.4 lists the GRG for each experiment. The experiment which is nearer to the ideal normalized value shows higher GRG. From the Table 12.4, Experiment 5 has best multi-response value, since it has the highest GRG. From the present work, it could be inferred that the optimization of multiple responses of FSW of AMCs can be done by converting multiple responses into a grey relational grade.

Table 12.5 Grey relational coefficient (GRC) and grey relational grade (GRG)

Expt. No.	Grey relational coefficient (GRC)			Grey relational grade (GRG)
	UTS	Vickers hardness	Elongation	
1	0.2481	0.2481	0.5287	0.3417
2	0.6644	0.4975	0.2481	0.4700
3	0.3382	0.3613	0.3594	0.3530
4	0.4186	0.3976	0.6516	0.4893
5	1.0000	0.7253	0.3186	0.6813
6	0.2998	0.2942	1.0000	0.5313
7	0.4300	0.4420	0.4449	0.4389
8	0.4975	0.4975	0.3840	0.4597
9	0.4420	1.0000	0.5287	0.6569

Table 12.6 Grey relational grade (GRG) response table

Factors	Process variable	Grey relational grade			Main effect
		Level 1	Level 2	Level 3	
A	Tool profile	0.3882	0.5673[a]	0.5185	0.1791
B	Rotational speed (RS)	0.4233	0.5370[a]	0.5137	0.1137
C	Welding speed (WS)	0.4442	0.5387[a]	0.4911	0.0945

Total mean value of the GRG = 0.4913
[a]Optimum GRG levels

The orthogonal experimental design enables to separate out the influence of each process variables on the GRG at different level. For instance, the hardness at levels 1, 2 and 3 can be computed by averaging the GRG for the Experiments 1–3, 4–6 and 7–9, respectively. The average of GRG for each level of the other responses namely tensile strength and elongation were calculated and listed in Table 12.6. The average of GRG for the nine experiments was computed and shown in Table 12.6. When the GRG is larger, the quality of the product will be nearer to the ideal value. Hence, higher GRG is required for optimum performance.

Based on higher GRG, the optimal process variables for better hardness, tensile strength and % elongation is Level 2 of each process variables as given in Table 12.6.

From the analysis, the optimal process variables are RS of 1000 RPM, WS of 70 mm/min and tool pin profile at Level 2, i.e. square pin profiled tool.

12.3.3 Validation Test

To confirm the enhancement of output responses while FS welding of aluminium composite, validation test was performed. Validation experiments were performed by choosing the process variables at optimum levels. Composites were welded using tool with square profiled pin, choosing RS as 1000RPM and WS as 70 mm/min. Welded samples were tested for tensile strength, macro-hardness and % elongation. Samples show better mechanical properties as given in Table 12.7 in comparison with the sample obtained from 5th experiment which has higher GRG. This implies that Taguchi-based GRA can be successfully employed for optimization of multiple responses simultaneously.

Table 12.7 Comparison of responses for experiment No. 5 and experiment with optimum process variables

Expt. No.	Process variables			Output responses		
	Tool profile	RS (RPM)	WS (mm/min)	UTS (MPa)	Vickers hardness (Hv)	% Elongation
5	SQ	1000	80	166	123	7.9
Optimum variables	SQ	1000	70	170	126	8.4

12.4 Result and Discussion

Amount of heat generated at this region depends on the RS, WS and type of tool, which controls the plasticization and flow of material. Frictional heat generated at the weld region and stirring action of the FSW tool influences the recrystallization of the grain, grain growth and grain orientation. All these factors control the mechanical properties of the joint [14–20]. From the microstructural study, it is found that SiC particles distributed within the matrix dendrites with no definite grain orientations as reported in literature [14]. Figure 12.3 depicts the various zones of weld area. Based on the grain size, the weld area is classified into four zones. (a) Stir zone (SZ) is a dynamically recrystallized zone due to the severe stirring action of the tool. The frictional heat produced at the tool workpiece interface is sufficient to plasticize the material and develops a new set of fine equiaxed grains. (b) Thermo-mechanically affected zone (TMAZ), next to SZ, is not in direct contact with the FSW tool. The heat produced at the SZ and the stirring action of the tool in the SZ alters the size of the grain and its orientation in this region. Material plastically deforms due to the shear stress resulted from the material flow. The grins are elongated and inclined in the direction of shear stress.

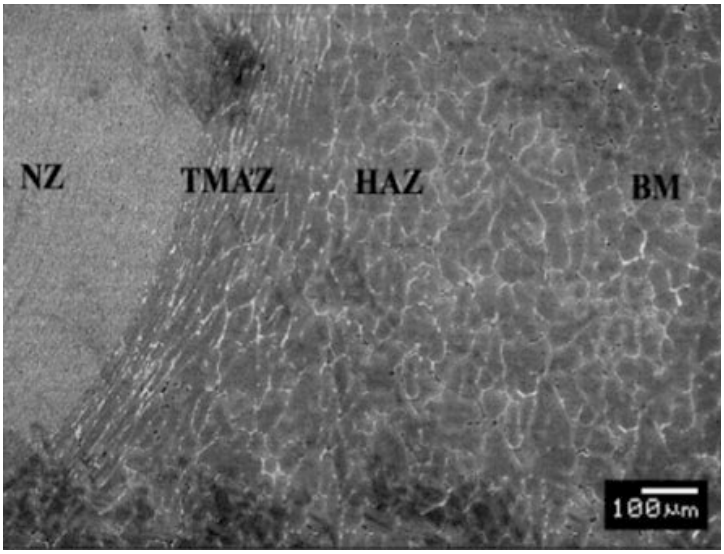


Fig. 12.3 Various zones of friction stir welded joint

(c) Heat affected zone (HAZ), is a zone next to TMAZ, does not undergo plastic deformation but, heavily influenced by the thermal cycle. The heat generated at the weld zone is sufficient to growth of the grains at HAZ. The grain structure at the HAZ is similar to the base material but of bigger size due to grain growth. (d) Base material (BM), region which is neither affected by the thermal cycle nor with the mechanical action of the tool.

Plasticization of material and the material flow in the weld region largely depends on the heat generated at the tool–workpiece interface and the amount of heat supplied to the weld region. Amount of heat generated at the tool–workpiece interface was dependent on the RS and heat supplied to workpiece was on WS [14–16]. As the RS increases, the quantity of heat generates at the weld region increases. So, at low speed, the heat generated at the region is less which affects the plasticization of material which in terms affects the weld quality. RS higher than the optimum speed generates more heat at the weld region, thereby creating turbulent material flow in the region [17]. Optimum heat generation is must for plasticization and proper material flow in the weld region. WS controls the transfer of generated heat at tool–workpiece to the workpiece [18]. Lower WS leads to higher heat input and results in the grain growth and larger heat-affected zone (HAZ). This phenomenon reduces the tensile strength and hardness. Higher WS leads to lower heat input to the weld region, thereby affecting the plasticization of material. Also, higher WS increases the cooling rate which affects the solidification of the plasticized materials in the weld region. Optimum WS is necessary to control the cooling rate which in term controls the grain growth in the weld region [19]. Mechanical properties of the welded samples were dependent on the grain structure and thus optimum RS and WS are necessary to obtain better mechanical properties.

Material flow around the tool pin is affected by the type of pin profile [15, 21]. The factor which varies with different type of tools was the ratio of swept volume to static volume. Conventional unthreaded cylindrical tool has this ratio as one, where both volumes are equal. Presence of thread on the cylindrical pin slightly increases this ratio and it has value of 1.02 in the present study. Polygon-shaped tools have higher swept volume compared to the static volume, results in higher ratio values. Square-shaped pin used in the present study has ratio of 1.57 and combined threaded cylindrical and square pin has ratio of 1.30. As the ratio increases, the pulsating effect of the tool also increases which results in improved material flow around the pin. Proper flow of material results in uniform distribution of particles, finer grain structure, thereby improving the mechanical properties of the welded parts. The present study supports the above phenomenon, indicating FSW with square profiled pin exhibits better mechanical properties.

12.5 Conclusions

The GRA has been used successfully to optimize the FSW process variables in joining AMCs for multiple responses namely macro-hardness, tensile strength and % elongation. Taguchi's L9 orthogonal array has been utilized for the experimental work. The FSW process variables are optimized for multiple responses, namely macro-hardness, tensile strength and % elongation. The present study indicates that complicated multiple response optimizations simultaneously can be significantly simplified by the Taguchi and GRA approach. From the study, it is inferred that to obtain better responses simultaneously, the optimum level of process variables to weld AMCs are square profiled pin with rotational speed of 1000 rpm and WS of 70 mm/min.

References

1. Thomas, W., Nicholas, E., Needham, J., Murch, M., Temple-Smith, P., Dawes, C.: Friction Stir Butt Welding. International Patent No. PCT/GB92/02203, GB Patent No. 9125978.8, 1991, U.S. Patent No. 5,460,317, 1995 (1991)
2. Mishra, R.S., Ma, Z.Y.: Friction stir welding and processing. Mater. Sci. Eng. R **50**, 1–78 (2005)
3. Subramanya, P., Amar, M., Arun, S., Mervin, H., Shrikantha, R.: Friction stir welding of Aluminium matrix composites—a review. In: MATEC Web of Conferences, vol. 144, p. 03002 (2018)
4. Cavaliere, P., De Santis, A., Panella, F., Squillace, A.: Effect of welding parameters on mechanical and microstructural properties of dissimilar AA6082–AA2024 joints produced by friction stir welding. Mater. Des. **30**, 609–616 (2009)
5. Rajakumar, S., Muralidharan, C., Balasubramanian, V.: Influence of friction stir welding process and tool parameters on strength properties of AA7075-T6 aluminium alloy joints. Mater. Des. **32**, 535–549 (2011)
6. Malopheyev, S., Vysotskiy, I., Kulitskiy, V., Mironov, S., Kaibyshev, R.: Optimization of processing-microstructure-properties relationship in friction-stir welded 6061-T6 aluminum alloy. Mater. Sci. Eng. A **662**, 136–143 (2016)
7. Elangovan, K., Balasubramanian, V., Babu, S.: Predicting tensile strength of friction stir welded AA6061 aluminium alloy joints by a mathematical model. Mater. Des. **30**, 188–193 (2009)
8. Murugan, N., Ashok Kumar, B.: Prediction of tensile strength of friction stir welded stir cast AA6061-T6/AlNp composite. Mater. Des. **51**, 998–1007 (2013)
9. Ukuyucu, H., Kurt, A., Areakliuglu, E.: ANN application to FSW of Al plates. Mater. Des. **28**, 78–84 (2007)
10. Prabhu, S.R., Shettigar, A., Herbert, M., Rao, S.: Multi response optimization of friction stir welding process variables using TOPSIS approach. In: IOP Conference Series: Materials Science Engineering, vol. 376, p. 012134 (2018)
11. Vijayan, S., Raju, R., Rao, S.R.K.: Multiobjective optimization of friction stir welding process parameters on aluminum alloy AA 5083 using Taguchi-based grey relation analysis. Mater. Manuf. Process. **25**(11), 1206–1212 (2010)
12. Deepandurai, K., Parameshwaran, R.: Multi response optimization of FSW parameters for cast AA7075/SiCp composite. Mater. Manuf. Process. **31**(10), 1333–1341 (2016)

13. Ghetiya, N.D., Patel, K.M., Kavar, A.J.: Multi-objective optimization of FSW process parameters of aluminium alloy using Taguchi-based grey relational analysis. Trans. Indian Inst. Met. **69**, 917 (2016)
14. Shettigar, A.K., Prabhu, S.R., Malghan, R., Rao, S., Herbert, M.A.: Application of neural network for the prediction of tensile properties of friction stir welded composites. Mater. Sci. Forum **880**, 128–131 (2017)
15. Prabhu, S.R., Shettigar, A.K., Rao, K., Rao, S., Herbert, M.A.: Influence of welding process parameters on microstructure and mechanical properties of friction stir welded aluminium matrix composite. Mater. Sci. Forum **880**, 50–53 (2017)
16. Li, Y.Z., Wang, Q.Z., Xiao, B.L., Ma, Z.Y.: Effect of welding parameters and B_4C contents on the microstructure and mechanical properties of friction stir welded B4C/6061Al joints. J. Mater. Process. Technol. **251**, 305–316 (2018)
17. Vijay, S.J., Murugan, N.: Influence of tool pin profile on the metallurgical and mechanical properties of friction stir welded Al–10 wt% TiB2 metal matrix composite. Mater. Des. **31**, 3585–3589 (2010)
18. Bozkurt, Y., Uzun, H., Salman, S.: Microstructure and mechanical properties of friction stir welded particulate reinforced AA2124/SiC/25p–T4 composite. J. Compos. Mater. **45**, 2237–2245 (2011)
19. Dinaharan, I., Murugan, N.: Effect of friction stir welding on microstructure, mechanical and wear properties of AA6061/ZrB2 in situ cast composites. Mater. Sci. Eng. A **543**, 257–266 (2012)
20. Kalaiselvan, K., Dinaharan, I., Murugan, N.: Characterization of friction stir welded boron carbide particulate reinforced AA6061 aluminum alloy stir cast composite. Mater. Des. **55**, 176–182 (2014)
21. Bist, A., Saini, J.S., Sharma, B.: A review of tool wear prediction during friction stir welding of aluminium matrix composite. Trans. Nonferr. Met. Soc. China **26**(8), 2003–2018 (2016)

Chapter 13
Tool Durability and Weldability in Hybrid Friction Stir Welding of High-Strength Materials

Amit Chaudhary, Chiranthan Ramesh, Viswanath Chinthapenta and Murshid Imam

Abstract Challenges that prevent broad industry adoption for friction stir welding of high-strength materials are (i) premature tool failure, (ii) selection of tool material and design, and (iii) weldability. To address these typical issues, a finite element-based study for laser-assisted friction stir welding process is carried out in ABAQUS, commercially available FE software. Experimentally, it has been shown that for welding of thick plate, laser assistance ahead of the tool is sought to be beneficial. However, the introduction of laser heating adds another complexity in modeling aspects. It is carried out in twofold. Firstly, the laser preheating is modeled to obtain the thermal map on the sample. Next, the regular FSW process is carried out on the already preheated sample. The important findings achieved are improving weld productivity and satisfactory tool life.

Keywords Friction stir welding · High-strength material · Laser heating · Thermal diffusivity · Finite element modeling

13.1 Introduction

Friction stir welding is a solid-state joining process which is developed by W. M. Thomas from TWI, Cambridge, England in 1991 [1]. This process was particularly used for the material which was unweldable by traditional welding technologies. In FSW, a specially designed rotating tool plunged into the abutting edges of the sheet to be welded and then transverse along the joint line. In this process, heat is generated by friction and severe plastic deformation allowing strong metallurgical bonds at the abutting interface. As compared to fusion welding technologies, there is no liquid weld pool formation, and hence excellent metallurgical

A. Chaudhary (✉) · M. Imam
Indian Institute of Technology Patna, Bihta, India
e-mail: ac3aug@gmail.com

C. Ramesh · V. Chinthapenta
Indian Institute of Technology Hyderabad, Sangareddy, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_13

properties such as fine microstructures, elimination of cracks in the weld zone, high depth of penetration is , comparatively less energy requirements, no requirement of shielding gas, and absence of consumable electrode [2]. As the capability of the FSW process expands, its demand in joining high-strength material also increases.

During welding of high-strength material, it is observed that the toolpin bend or deform due to high bending and torsional stresses. To overcome this problem, laser preheating of the sheet/plate is carried out ahead of the tool. Laser preheating softens the material so that bending and torsional stresses acting on the tool can be decreased. The present work aims to develop a finite element model for the laser-assisted FSW process. The proposed finite element has the two important industrial applications: (i) it helps to widen the process window by improving the weldability and (ii) improving tool durability by reducing the stresses, especially when FSW of high-strength material with larger sheet thickness needs to be conducted. A passively 3D FE coupled thermo-mechanical model has been developed by considering the thermal and mechanical aspects in ABAQUS/Explicit and to control the heavy distortion, arbitrary Lagrangian-Eulerian (ALE) formulation has been used. For incorporating the moving body heat source, DFLUX user-defined subroutine code was used.

13.2 Finite Element Method

A three-dimensional coupled Eulerian-Lagrangian thermo-mechanical FSW model is developed in ABAQUS/Explicit. The whole simulation is carried out in a dynamic, temp-displacement, explicit step. Stable time increment estimator is used for time marching. In the Lagrangian model, as the material deforms, mesh also deforms with it. In the Eulerian model, the mesh does not change as the material deforms. A large amount of heat is generated due to large plastic deformation and friction. The governing and material constitutive equations are given in Fig. 13.1. The workpiece material used in the present is 2024-T3 aluminum alloy. The dimensions and mesh density of the workpiece are shown in Fig. 13.1b, c. The mechanical and thermal properties are chosen in the simulation is based on the reported work by Mandal et al. [3]. The workpiece is modeled as three-dimensional deformable bodies, reference point and inertia are provided to the tool. Due to large deformation and mesh distortion, arbitrary Lagrangian-Eulerian (ALE) meshing has been considered in the simulation. A C3D8RT element has been chosen, which is a multi-material thermally coupled 8-noded element. Tool–workpiece interaction is defined by using the surface-to-surface contact formulation with tangential force on penalty basis.

Laser beam heating is solved in coupled temperature displacements, a pre-step to plunge. The Gaussian distribution is used to explain the shape of the laser beam heat source. The double-ellipsoidal heat source describes the shape of the body heat source. The laser beam intensity exponentially changes with a penetration depth of

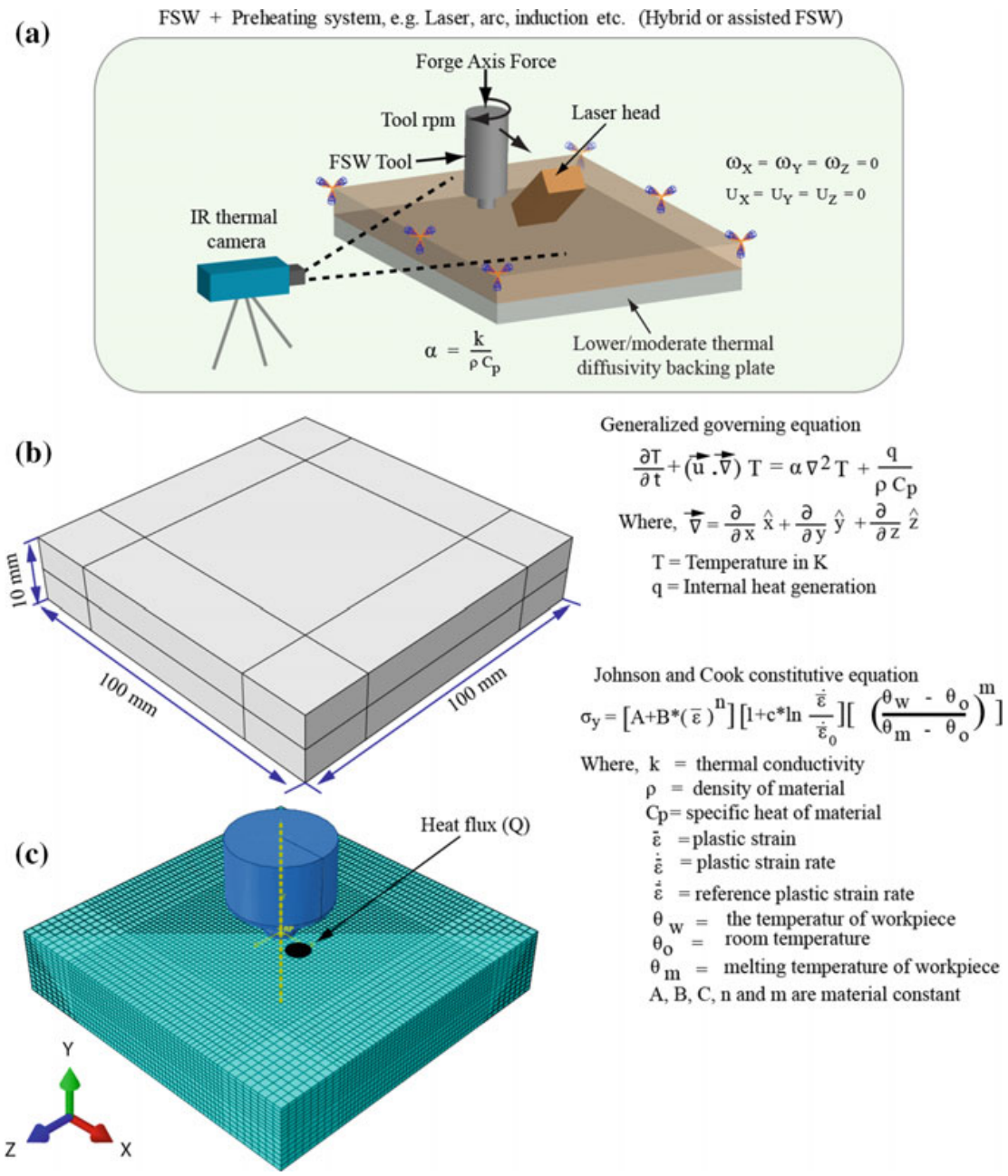


Fig. 13.1 **a** Schematic diagram of FSW with the preheating system, **b** dimensions of the deformable workpiece, and **c** FEA model with laser heat flux and governing equation of material model and heat transfer

welded element. The moving heat source parameters are controlled by changing DFLUX user-subroutine code which is written in FORTRAN programming language. The plunge stage of the tool is modeled in the next step based on the temperature-dependent properties from the laser pre-heating with dynamic, time–temperature explicit analysis. In the present work, FEM formulation is done by using ABAQUS commercial package. Goldak's model is used for simulating the laser heat source. Figure 13.2 shows the schematic of double ellipsoid Goldak heat source model. Double-ellipsoidal heat source Goldak's model was used to preheat

$$q_1(x,y,z) = \frac{6\sqrt{3}\, f_1 Q}{abc_1\pi\sqrt{\pi}} \exp\left(-3\frac{x^2}{a^2} - \frac{y^2}{b^2} - 3\frac{z^2}{C_1^2}\right)$$

$$q_2(x,y,z) = \frac{6\sqrt{3}\, f_2 Q}{abc_1\pi\sqrt{\pi}} \exp\left(-3\frac{x^2}{a^2} - \frac{y^2}{b^2} - 3\frac{z^2}{C_2^2}\right)$$

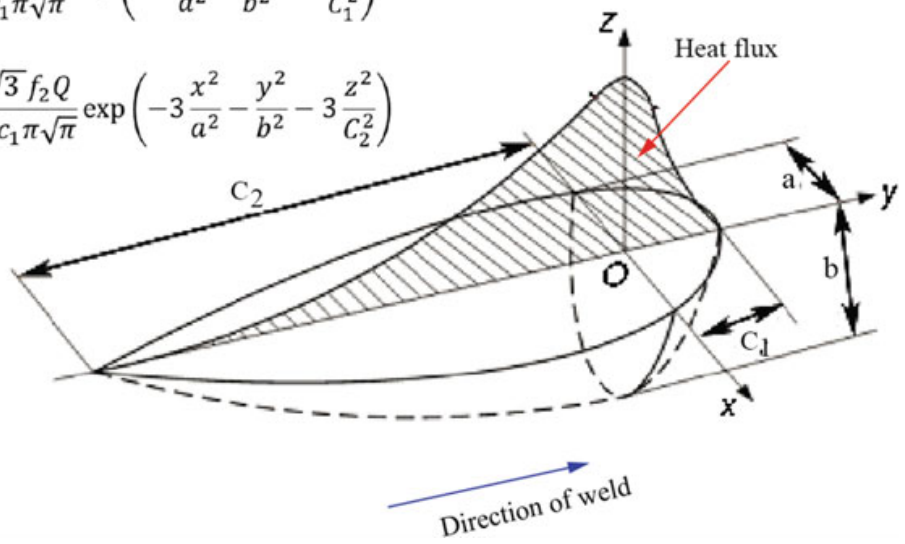


Fig. 13.2 Double ellipsoid Goldak heat source model a, b, c_1, and c_2 are sets of axes that define front and rear ellipsoid, and f_1 and f_2 represent distribution of heat source energy, respectively

the workpiece. Two-half ellipses are connected to single semi-axis. Power distribution of this double-ellipsoidal model heat source is given by the equations as shown in Fig. 13.2. Note that the resultant distribution of heat source is given by $q(x, y, z) = q_1(x, y, z) + q_2(x, y, z)$.

13.3 Results and Discussion

Figure 13.3a shows the relationship between welding parameters and a shift in weld productivity. Note that in this figure, the data for selection of control process parameters for plate thickness >19 mm are compared with the process window of 4-mm-thick sheet [4–8]. It can be cleary seen from this figure that there is a shift in weld productivity of FSW of thicker plate when compared with the smaller sheet thicknesses (<4 mm). The reason for the shift in the weld productivity can be understood as follows: (i) selection of process control parameters requires a more physical/scientific understanding of the process for welding thicker plates. To weld thicker plate, higher heat energy per unit length is needed. The control heat input parameters are tool rotational speed, axial force, and welding speed. In the case of thicker plate, large magnitude of axial force is required to ensure the optimum material mixing along the weld plate thickness. Therefore, the coupling of the tool rotational speed and axial force needs to be adjusted in such a way that it avoids the overheating of the material at the tool shoulder/workpiece interface. Otherwise, the higher degree of thermal softening will make the process unstable and the excess flash formation, (ii) the present industrial practice to avoid such a situation is the

weld-flip-weld approach. The first weld pass is made with the shorter tool probe length and then workpiece is flipped for welding from the other side. This approach involves cleaning of the bead obtained after the first weld pass, the additional time for the readjustment of the fixtures and requirement of large clamping forces for the adjustment of the workpiece with the fixtures because of the distortion of the workpiece after the first weld pass. In addition, the tool probe experiences the larger bending and shear stress in the case of thicker plates. The intricate tool probe profile design is detrimental for the tool life. To achieve better welding condition, the tool has been provided with probe surface which acts as a notch and the tool became weaker at that section and may lead to failure. Therefore, the tool material must be strong enough to withstand the large bending and torsion stresses acting on the tool when the welding speed is high. In this context, Laser pre-heating ahead of the tool has been found to be beneficial to solve such problems for welding of thicker plates and for welding of high-strength alloys. From Fig. 13.3b, it can be observed that for the high-strength material (mild- and high-strength steels) the yield strength does

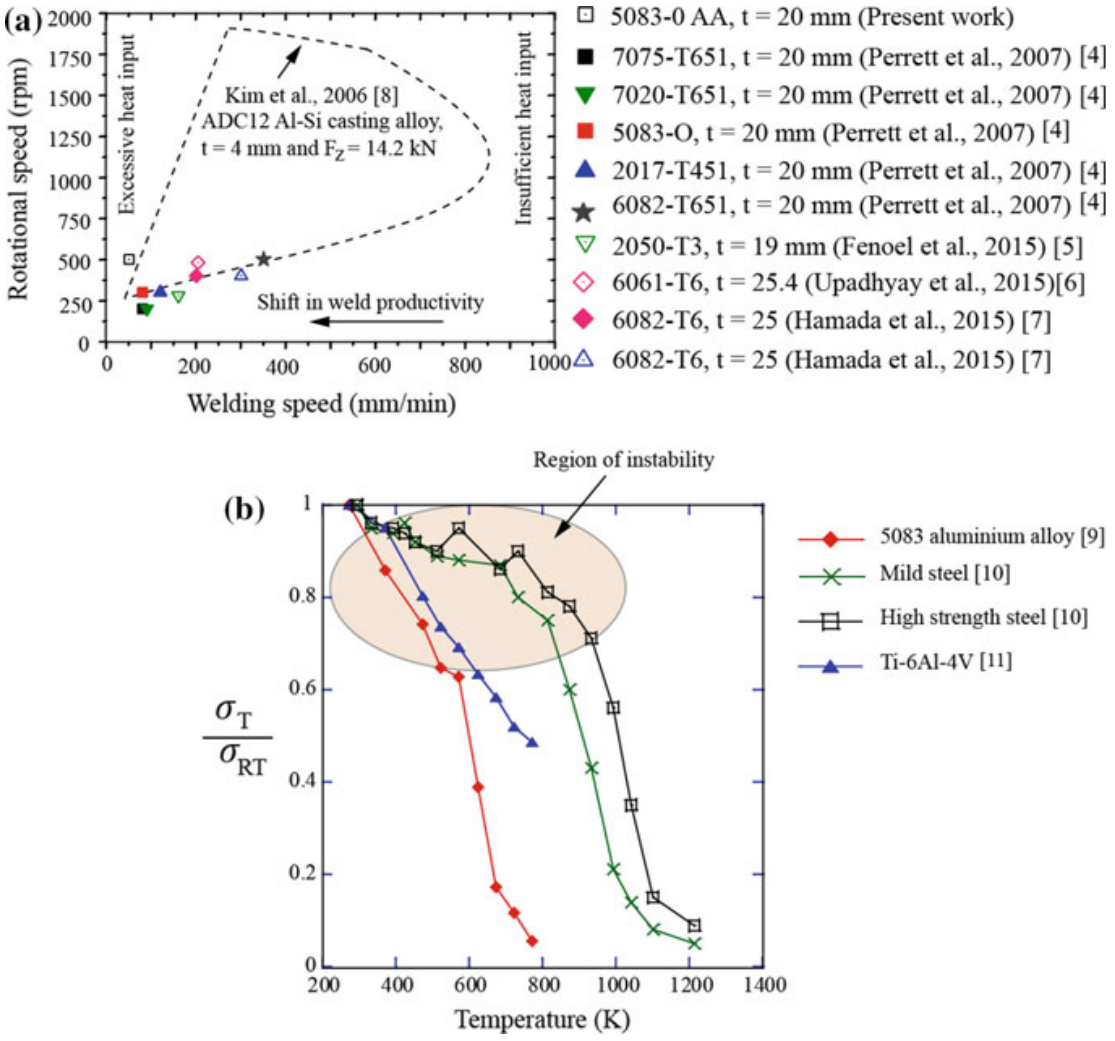


Fig. 13.3 **a** Process map for weld productivity in thick aluminum welds and **b** instability region for carrying out friction stir welding of high-strength materials

not vary significantly for the temperature range from 300 to 800 K (for details see the Fig. 13.3b). Note that σ_T/σ_{RT} is the ratio of the yield strength of the materials at elevated temperatures and room temperature, respectively. At the time of tool plunging, the temperature is not sufficiently high to cause the thermal softening of the work material and thereby leads to the tool failure. Adding the laser heating helps to solve this problem. The finite element simulation with laser pre heating has been carried out to compare the results and benefits of pre-heating over without pre-heating. From the predicted model, it can be observed that stress developed on the tool has been decreased significantly.

Figure 13.4a shows the comparisons of maximum principal stress values at three different locations during the plunging stage. Here, the maximum principal stress values are compared for both the cases (with and without preheating). The three locations (location 1, location 2, and location 3) shown here is for the tool plunging

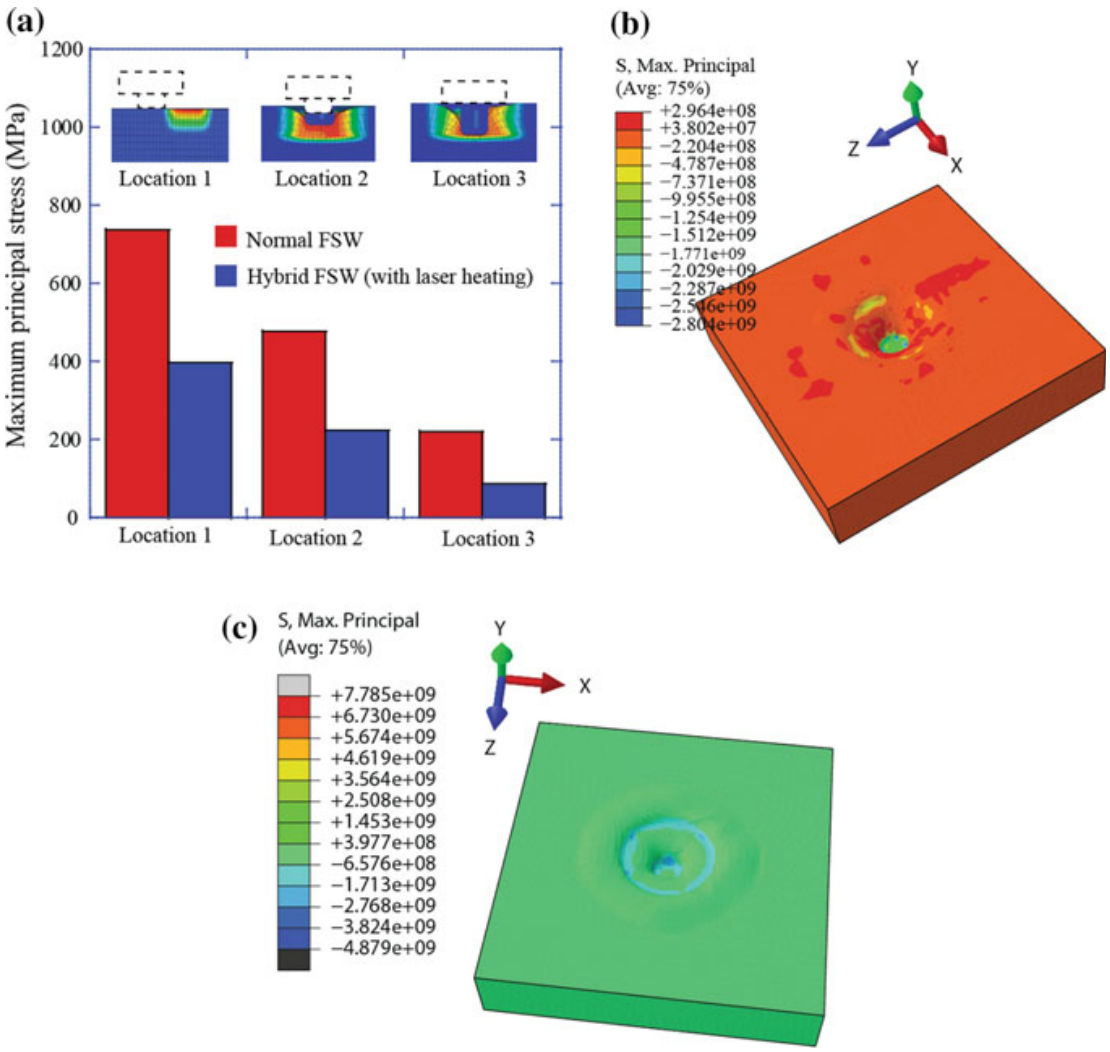


Fig. 13.4 **a** Comparisons of stress values at plunging locations 1, 2, and 3, respectively, **b** finite element simulation result showing the maximum principal stress value for normal FSW process at location 3, and **c** finite element simulation result showing the maximum principle stress value for Hybrid FSW process at location 2

distance of 0.6, 6, and 12 mm, respectively. Note that Fig. 13.4 b, c shows the predicted maximum principal stress on the workpiece at location 3 and location 2 for the case of normal FSW process and laser-assisted FSW. The stresses generated on the tool can be reduced in a very much amount in the case of laser heating as compared to the normal FSW process. Thus, it can be clearly understood that the laser heating is beneficial for causing the thermal softening of the material, particularly during the plunging. On similar lines, Yaduwanshi et al. [9] and Fei et al. [10] pointed out that the laser preheating is beneficial for the welding of dissimilar materials, particularly when aluminum and other high-strength materials (Steel, Titanium, etc.) are involved [11–13].

13.4 Conclusion

Laser-assisted friction stir welding has a potential to lower the stresses acting on the FSW tool by causing the thermal softening of the workpiece material. The finite element model suggests that the tool durability can be greatly improved. The addition of laser heating can also widen the process window, and hence better productivity can be achieved. The development of hybrid FSW system paves the way to establish this inexpensive and simple technique in heavy metal industries where metallurgical problems are still the main challenges to overcome.

References

1. Chena, H.-B., Yan, K., Lin, T., Chen, S.-B., Jiang, C.-Y., Zhao, Y.: The investigation of typical welding defects for 5456 aluminum alloy friction stir welds. Mater. Sci. Eng. A **433**, 64–69 (2006)
2. Imam, M., Sun, Y., Fujii, H., Aoki, Y., Ninshu, M.A., Tsutsumi, S., Murakawa, H.: Friction stir welding of thick aluminium welds—challenges and perspectives. In: Conference Paper, Friction Stir Welding and Processing IX, pp. 119–124. Springer (2017)
3. Mandal, S., Rice, J., Elmustafa, A.A.: Experimental and numerical investigation of the plunge stage in friction stir welding. J. Mater. Process. Technol. **203**(1), 411–419 (2008)
4. Perrett, J.G., Martin, J., Threadgil, P.L., Ahmed, M.M.Z.: Recent developments in friction stir welding of thick section aluminium alloys. In: Proceedings of 6th World Congress, Aluminium 2000, pp. 13–17, Florence, Italy (2007)
5. Avettand-Fenoel, M.-N., Taillard, R.: Heterogeneity of the nugget microstructure in a thick 2050 Al friction stirred weld. Metall. Mater. Trans. A **46**, 300–303 (2015)
6. Upadhyay, P., Reynolds, A.: Effect of backing plate thermal property on friction stir welding of 25 mm-thick AA6061. Metall. Mater. Trans. A **45**, 2091–2100 (2014)
7. Hamada, A.S., Jarvenpaa, A., Ahmed, M.M.Z., Jaskari, M., Wynne, B.P., Porter, D.A., Karjalainen, L.P.: The microstructural evolution of friction stir welded AA6082-T6 aluminium alloy during cyclic deformation. Mater. Sci. Eng. A **642**, 366–376 (2015)
8. Kim, Y.G., Fujii, H., Tsumura, T., Komazaki, T., Nakata, K.: Three defect types in friction stir welding of aluminum die casting alloy. Mater. Sci. Eng. A **415**, 250–254 (2006)

9. Yaduwanshi, D.K., Bag, S., Pal, S.: Numerical modeling and experimental investigation on plasma-assisted hybrid friction stir welding of dissimilar materials. Mater. Des. **92**, 166–183 (2016)
10. Fei, X., Ye, Y., Jin, L., Wang, H., Lv, S.: Special welding parameters study on Cu/Al joint in laser-heated friction stir welding. J. Mater. Process. Technol. **256**, 160–171 (2018)
11. Tello, K.E., Gerlich, A.P., Mendez, P.F.: Constants for hot deformation constitutive models for recent experimental data. Sci. Technol. Weld. Join. **15** (2010)
12. Chen, J., Young, B., Uy, B.: Behaviour of high strength structural steel at elevated temperatures. J. Struct. Eng. **132** (2006)
13. Kim, J., Kim, K.-H., Kwon, D.: Evaluation of high-temperature tensile properties of Ti-6Al-4V using instrumented indentation testing. Met. Mater. Int. **22**, 209–215 (2016)

Chapter 14
Swarm-Intelligence-Based Computation for Parametric Optimization of Electron Beam Fabrication

Sanjib Jaypuria, Amit Kumar Das and Dilip Kumar Pratihar

Abstract In the present paper, two recently developed intelligent swarm optimization algorithms were used for optimization of input variables in the electron beam welding process. The computational intelligence-based optimization algorithms used in this study were bonobo optimizer (BO) and cricket algorithm (CA), which are swarm-intelligence-based algorithms developed by mimicking self-organization, co-evolution, and learning of multiple agents in that particular swarm of the population. Bead-on-plate welding of CuCrZr alloy plate was carried out by electron beam. The welding runs were designed according to central composite design, and regression analysis was conducted to establish input–output dependency. Accelerating voltage, beam current, scanning speed, and focusing distance were considered as input variables of the process, whereas bead geometry parameters and microhardness of weld zone were regarded as the responses of the electron beam welding process. An optimization problem was formulated with an aim to minimize the weldment area without sacrificing the bead penetration and microhardness of the fusion zone. Penalty function approach was also used in this study to optimize the unconstrained problem. Three optimization algorithms were used in this study to solve the optimization constrained problem designed to optimize the input variables of the process yielding proper responses of the weld region. The performances of proposed approaches were evaluated based on percent deviation from that of the experimental result. Bonobo optimizer outperformed cricket algorithm and genetic algorithm for optimizing the process.

Keywords Bonobo optimizer · Cricket algorithm · Electron beam welding · Swarm intelligence · optimization

S. Jaypuria (✉) · A. K. Das · D. K. Pratihar
Soft Computing Lab, Department of Mechanical Engineering, Indian Institute of Technology Kharagpur, Kharagpur 721302, India
e-mail: sanjibjaypuria@iitkgp.ac.in

A. K. Das
e-mail: amit.besus@gmail.com

D. K. Pratihar
e-mail: dkpra@mech.iitkgp.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_14

14.1 Introduction

Manufacturing processes are always associated with trial and error method prior to the actual experimentation. It is always not beneficial to conduct these trial experiments due to the wastage of resources and time. Therefore, these days' industries are shifting toward the automation of manufacturing processes, which will reduce the lead time and consumption of resources [1]. In addition to input–output modeling, manufacturing industries also require optimization tools to enhance productivity by ensuring quality-based products. Most of the manufacturing processes deal with multi-input–output systems with complex interdependency of the variables; so, optimization is needed to get a suitable set of input parameters, which ensure a high-quality product. Electron beam welding is such a fusion welding process, where multiple process parameters are involved along with a certain degree of automation of the process. Electron beam welding has gained diverse applications in the fields like nuclear, energy, aerospace, and automotive in the recent era. The wide application of this process is because of low heat input, high aspect ratio, a minimal area of heat-affected zone and capability to weld reactive and reflective metals [2]. Here, electron beam acts as a heat source and fuses the joint by providing intense heat energy, which is being converted from the high-speed electron beam's kinetic energy. The quality of the joint is also decided by the vacuum environment of the welding [3].

Computational intelligence is a part of soft computing tools or metaheuristic algorithms, which are promising tools for modeling, optimization, and prediction of response in material and manufacturing engineering [4]. The most frequently used tools include artificial neural network (ANN), genetic algorithm (GA), fuzzy logic (FL), particle swarm optimization (PSO), and others. A few researchers studied the applicability of metaheuristic algorithms in the different manufacturing process and welding methods.

Benyounis and Olabi [5] reviewed linear regression, response surface method, artificial neural network (ANN), Taguchi method, and design of experiment (DoE) techniques used in welding process for optimization and modeling. It had been concluded from their work that GA and other metaheuristic algorithms are powerful optimization tools, especially in the irregular experimental region. The ability of these evolutionary approaches to operate in a vast set of search points to find optimal point was well appreciated in this work. Sathiya et al. [6] performed a laser welding of AISI 904L and investigated the relationship between laser input parameters and laser-welded joint characteristics. They first established the input–output relationships by artificial neural network and then, the developed model was used for optimization of responses using genetic algorithm. Jha et al. [7] modeled electron beam welding of reactive zircaloy by artificial neural network. The developed neural networks were optimized by GA, PSO, and BP algorithm. These developed bidirectional models had shown good prediction capability of the model. BPNN had proved the better performance as compared to the other two approaches. Park and Rhee [8] performed experiments on the laser welding of automotive grade

AA5182 aluminum alloy and studied the tensile strength of the welded joints. A neural network model tuned by genetic algorithm was proposed to predict the tensile strength using the experimental data. Dey et al. [9] conducted electron beam welding of stainless steel and experiments were designed according to central composite design (CCD). The objective of their study was to minimize the weld area without hampering the maximum bead penetration. Penalty function approach was used to obtain the suitable process parameters of the electron beam welding process. Kanigalpula et al. [10] studied the spiking behavior of partially penetrated ETP copper plates welded using electron beam. The constrained optimization problem for achieving the minimum standard deviation of penetration was solved using genetic algorithm, particle swarm optimization, and desirability function approach. The predicted results with various proposed approaches were verified with that of real experiments and a good agreement between them was achieved. In addition to this, there were a very few available literatures, where swarm-intelligence-based optimization algorithms were implemented on the manufacturing process and system for optimization and modeling.

It is observed from the literature that metaheuristic approaches are mostly used in the optimization of many manufacturing processes in addition to traditional tools because of their flexibility and accuracy. In addition to this, the modeling tool like neural network is also fine-tuned with these evolutionary approaches for the better efficiency of the developed neural network architecture. In most of the available literature, optimization tools like genetic algorithm and particle swarm optimization are used for optimization and modeling of the manufacturing process. It is also found from the research that there are only a few available works of literature, where swarm-intelligence-based approaches were used for optimization of electron beam fabrication.

Therefore, here, two recently developed, swarm-intelligence-based algorithms, e.g., Bonobo optimizer (BO) and cricket algorithm (CA), were used to optimize the process parameters of electron beam welding. The approaches were intended to minimize the weldment area and maximize the bead penetration and weld zone microhardness. The performance comparison and working principle were discussed in this work. In addition to this, genetic algorithm (GA), which is a bio-inspired tool, was also used to optimize the constrained problem. The performances of these developed models were evaluated in terms of absolute percentage deviation and convergence rate.

14.2 Experimentation and Data Collection

Precipitation hardened CuCrZr alloy plates of dimension 100 mm × 65 mm × 10 mm were used for electron beam welding in bead-on-plate configuration. The procured metal had 0.80 wt% Cr, 0.06 wt% Zr, and 0.13% impurities and remaining as copper. The primary input parameters of electron beam welding, namely accelerating voltage (V), beam current (I), welding speed (S), and focusing distance

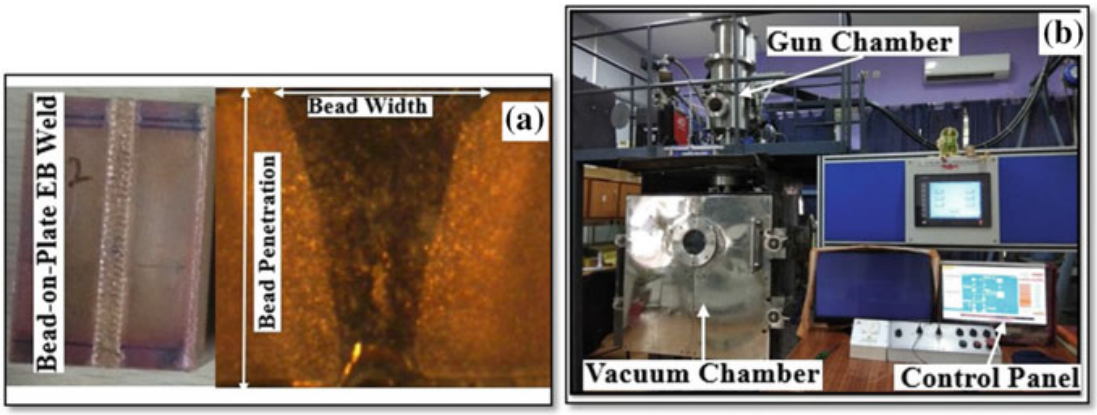


Fig. 14.1 Photograph of **a** weld bead geometry and **b** electron beam welding setup

(B) were taken as process variables, whereas microhardness of fusion zone and bead geometrical parameters, like bead penetration (BP), bead width (BW), microhardness of fusion zone (μ_{FZ}), was considered as the responses to be measured. Figure 14.1a indicates various geometrical features of the weld bead. The photograph of electron beam welding setup along with their principal components is given in Fig. 14.1b. A few trial runs were conducted to select the ranges of input variables. The actual experiments for this study were prepared by considering the method of central composite design (CCD) with three central points. A total of 27 combinations of input parameters were generated, and 81 experiments were conducted by considering three replicates of the original CCD [11]. The actual and test cases experimental data were taken from [11]. A section of the weld (15 × 10 10 mm) was cut from the obtained weld for measuring the weld responses. The polished and etched samples were examined using optical macroscope to measure bead penetration and width. Microhardness of fusion zone (FZ) was measured for each sample in Vickers microhardness tester at an applied load of 50-gf and dwell time of 15 s.

14.3 Method of Analysis: Proposed Approaches

The mathematical relationships between the inputs and output were established by nonlinear regression before solving the optimization problem. The proposed metaheuristic approaches, which were used as optimization algorithms in the developed problem, were briefly discussed below.

14.3.1 Regression Analysis

A nonlinear statistical regression analysis was conducted using Minitab15 software to establish input–output relationships of the process, as given below in Eq. (14.1).

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k} \sum_{i<j}^{k} \beta_{ij} x_i x_j + \varepsilon \tag{14.1}$$

where y is the response parameter; $\beta_0, \beta_i, \beta_{ii}$ are the coefficients to be computed using the least squares method. Here, ε denotes the error in fitting. The nonlinear regression equations for the responses were determined, as shown in Eqs. (14.2)–(14.4).

$$\begin{aligned} \text{BW} = & -172.103 + 2.16103\,V + 0.827001\,I - 0.0522837\,S \\ & + 0.479534\,B + 0.0104345\,V^2 - 0.00000530401\,I^2 \\ & - 0.00000157596\,S^2 + 0.000117136\,B^2 - 0.00057152\,VI \\ & - 0.0000567146\,VS - 0.00867625\,VB - 0.0000572391\,IS \\ & - 0.00193298\,IB + 0.000171340\,SB \end{aligned} \tag{14.2}$$

$$\begin{aligned} \text{BP} = & 5857.98 + 8.02798\,V + 0.0920434\,I + 0.0623438\,S \\ & - 33.0775\,B - 0.0313765\,V^2 - 0.00146853\,I^2 \\ & + 0.00000207134\,S^2 + 0.0456195\,B^2 \\ & + 0.0000113958\,VI - 0.0002530645\,VS \\ & - 0.0113766\,VB - 0.000145435\,IS \\ & + 0.00100352\,IB - 0.000114410\,SB \end{aligned} \tag{14.3}$$

$$\begin{aligned} \mu_{\text{FZ}} = & 435.876 + 41.1517\,V - 6.48847\,I - 0.289883\,S \\ & - 6.03015\,B - 0.285728\,V^2 + 0.00680864\,I^2 \\ & + 0.000101836\,S^2 + 0.00893827\,B^2 \\ & + 0.00541667\,VI + 0.000987500\,VS - 0.0291667\,VB \\ & + 0.000322917\,IS + 0.013\,IB + 0.000104167\,SB \end{aligned} \tag{14.4}$$

14.3.2 Problem Formulation

A constrained optimization problem was formulated to maximize the weld penetration and microhardness of the weld zone without increasing the weld area.

Minimize

$$A = 0.5 \times (\text{BW} \times \text{BP}) \tag{14.5}$$

subject to

$$\begin{aligned} &\mathrm{BP} \geq 0.7 \times t, \\ &|\mu_{\mathrm{PM}} - \mu_{\mathrm{FZ}}| \leq 59.0 \end{aligned}$$

and

$$\begin{aligned} &50 \leq V \leq 60 \\ &70 \leq I \leq 110 \\ &600 \leq S \leq 1000 \\ &365 \leq B \leq 375 \\ &\mu_{\mathrm{PM}} = 153 \\ &t = 10 \end{aligned}$$

Here, t and μ_{PM} indicate the thickness of the specimen in mm and microhardness of base or parent metal in VHN, respectively. A denotes the weldment area, which had to be minimized for the reduction of bead width and large heat-affected area. Therefore, the constrained optimization problem was designed to minimize weld area without scarifying the bead penetration and microhardness of the fusion zone. The above functional constraint was set by observing the microhardness of FZ for different combinations of input parameters lying within their respective ranges.

14.3.3 Genetic Algorithm (GA)

Genetic algorithm (GA) is a bio-inspired metaheuristic algorithm working based on Darwin's principle of natural selection and evolution. Therefore, concisely, a GA initiates the operation by selecting a random population of solutions and then improves the fitness level through iterative utilization of selection, crossover, and mutation operators. This iterative process stops when the stopping criterion is reached. Interested readers may refer to Pratihar [12] for a detailed discussion of the working principle of GA. The results of GA had been obtained from the literature for comparison purpose [11]. The working flowchart is given in Fig. 14.2a.

14.3.4 Bonobo Optimizer (BO)

Very recently, a new metaheuristic algorithm, named as bonobo optimizer, has been developed [13]. This newly proposed optimization algorithm imitates the social behavior and reproductive strategies of the bonobos. Generally, bonobos have been found to adopt a fusion–fission social practice, and moreover, they have four types of reproductive strategies, like promiscuous, restrictive, consortship, and extra-group mating. Furthermore, in the algorithm, two-phase conditions, namely

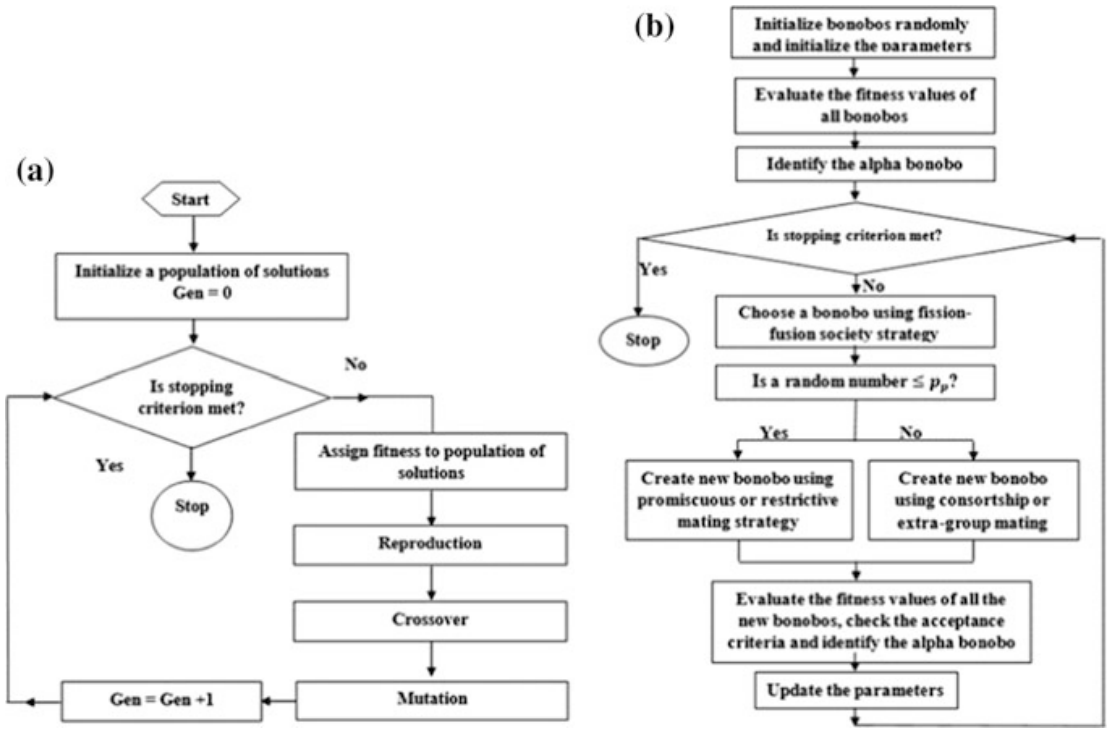


Fig. 14.2 Flowcharts of **a** genetic algorithm and **b** bonobo optimizer

positive phase and negative phase, have been considered and depending upon the phase, its associated phase probability (p_p), either exploitation or exploration phenomenon, is promoted. These strategies have been artificially modeled in the developed BO to make a more efficient optimization technique. Nevertheless, the parameters of the BO have been made adaptive to reach the globally optimum solution with the faster convergence rate. The flowchart of this BO has been shown in Fig. 14.2b.

14.3.5 Cricket Algorithm (CA)

Motivated from the nature and characteristics of an insect, namely cricket, a novel cricket algorithm had been proposed in recent time [14]. This algorithm was developed considering the functional aspects of some swarm-intelligence-based optimization methods, like bat algorithm, firefly algorithm, particle swarm optimization (PSO), etc. Also, a few peculiar characteristics of the crickets, such as forecasting the weather by courtesy of the number of fluttering of crickets and the sound frequency in the temperature, are artificially modeled to solve the optimization problems [15]. It had been applied to solve several standard benchmark functions and various real-world optimization problems. In those experiments, CA had shown its supremacy in performance compared to a few other optimization algorithms.

14.4 Results and Discussion

Sixteen test cases were used to validate this mathematical model. Figure 14.3 shows the absolute percentage deviation in the prediction of responses. The mean absolute percent error (MAPE) in the prediction of BW, BP, and μ_{FZ} were found to be 2.44, 4.84, and 1.14%, respectively. Similarly, the root mean squared error in prediction of BW, BP, and μ_{FZ} were found to be equal to 0.58, 1.03, and 4.47, respectively. In addition to this, *R*-squared value for each response in the regression model was also determined to check the closeness of data points to a fitted regression line. The values were found to be 97.95, 76.15, and 97.77% for BW, BP, and μ_{FZ}, respectively.

A real-coded genetic algorithm was utilized to solve the optimization problem in MATLAB 2015b environment. The minimum value of the fitness function (here the weldment area) was the objective of the study. A penalty function approach was adopted to handle constrained optimization problem, where infeasible solutions were penalized for being removed from the population of solutions. As the performance of the GA is dependent on its parameters, the parametric study was carried out to decide the set of optimal GA parameters by varying one parameter at a time and keeping the others unaltered [12]. The best fitness was obtained for the following GA parameters: $P_c = 0.9$; $P_m = 0.009$; $N = 100$, and $G_{max} = 150$. The GA could yield the optimal solution as follows: $V = 56.08$ kV; $I = 88.6$ mA; $S = 600$ mm/min, and $B = 365$ mm. The optimized value of the area of the weldment, as obtained by the GA, was 12.14 mm^2. A detailed comparison of various algorithms suggested results and is given in Table 14.1.

The common parameters, such as population size ($N = 100$) and maximum number of iterations ($iter_{max} = 150$), are taken the same for each of the optimization

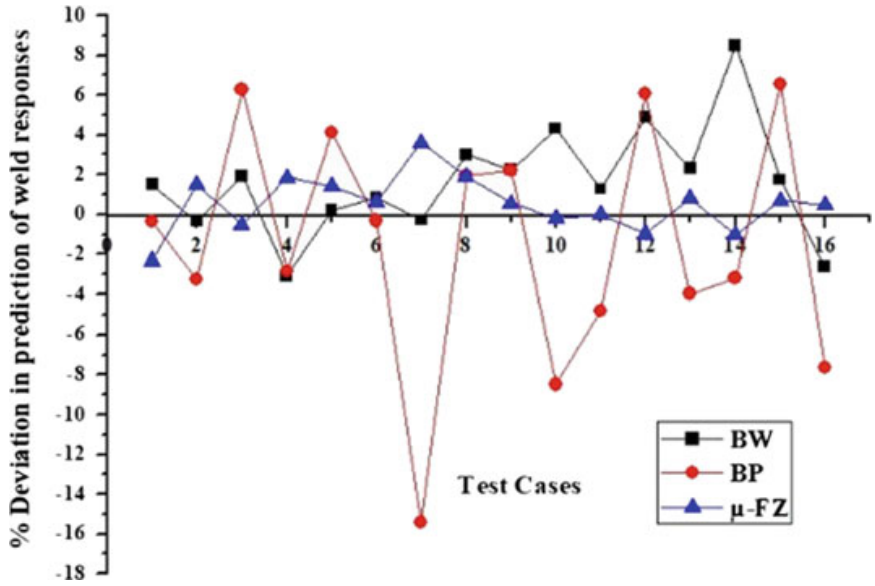


Fig. 14.3 Deviation (%) in the prediction

Table 14.1 Comparison of algorithm suggested results and experimental results

Proposed approaches	V (kW)	I (mA)	S (mm/min)	B (mm)	Area (mm^2)	BW (mm)	BP (mm)	μ_{FZ} (VHN)
CA	57	84	620	375	12.10	3.46	7.00	99.19
EXP	57	84	620	375	11.74	3.40	6.91	100.26
% error					3.093	1.70	1.36	1.066
BO	56	84	607	375	11.99	3.43	7.00	100.76
EXP	56	84	607	375	11.84	3.34	7.09	100.33
% error					1.544	2.54	0.97	0.428
GA [11]	56.08	88.6	600	365	12.14	3.45	7.04	99.66
EXP	56	88	600	365	12.65	3.61	7.01	96.87
% error					4.031	4.43	0.42	2.880

techniques. In addition, the algorithm-specific parameters used for BO are as follows: Initial probability for extra-group mating, $p_{\text{xgm_initial}} = 0.03$; sharing coefficient for alpha bonobo and the selected bonobo are found to be equal to 1.3 and 1.4, respectively; rate of change of phase probability, rcpp = 0.0036; maximum value for the temporary subgroup size factor, $\text{tsgs}_{\text{factor_max}} = 0.02$. The BO obtained the optimal solution as follows: V = 56 kV; I = 84 mA; S = 607 mm/min; and B = 375 mm. The optimized value of area of weldment as obtained by the BO was 11.99 mm^2. For CA algorithm, the following parameters were taken while it is running: minimum frequency, $Q_{\min} = 0$; $\beta_{\min} = 0.2$. The CA predicted the optimal solution as V = 57 kV; I = 84 mA; S = 620 mm/min; and B = 375 mm. The optimized value of area of weldment as obtained by the CA was 12.10 mm^2.

From Table 14.1, it is observed that BO could yield the best result compared to others for this optimization problem. Therefore, it can be concluded that BO could outperform GA and CA in this study.

The convergence plots of BO and CA algorithms are given in Fig. 14.4. The graph indicates that BO could yield a better solution with faster convergence rate

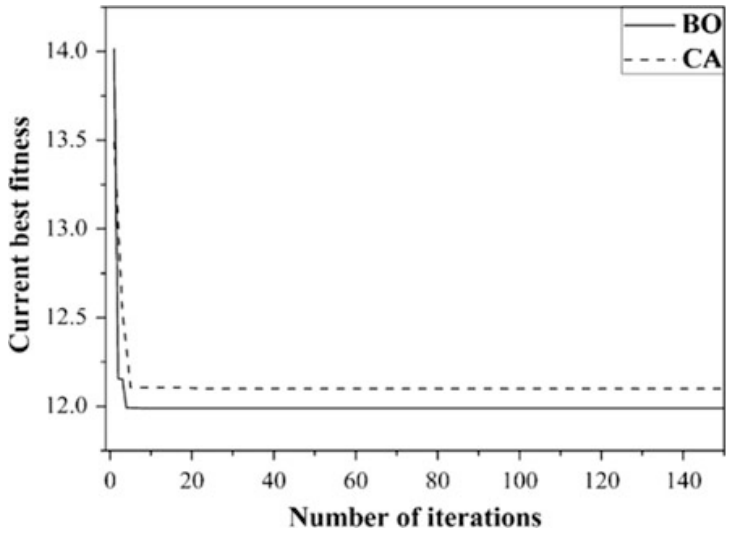


Fig. 14.4 Convergence plot for BO and CA

compared to CA. It can be inferred from the result of absolute percentage deviation and the result of the convergence rate that BO emerges as the winner for the optimization problem attempted in this study. In BO, the parameters of the algorithm are found to be adaptive in nature. According to the need of the situation, either exploitation or exploration phenomenon is promoted, or a proper balance between these two is maintained. In the case of a positive phase, the solutions are updated either using a promiscuous or restrictive type of reproductive strategy, and in this phase, the searching agents are observed to follow the current best solution. On the other hand, a more diversified search is found using either a consortship or extra-group mating strategy during a negative phase. Furthermore, the selection procedure of the mating partner based on the fission–fusion social behavior of the bonobos is also responsible for maintaining a right balance between the selection pressure and population diversity. Due to these reasons, BO is found to have an efficient searching mechanism to reach the globally optimum solution and hence, it has better performance compared to others.

14.5 Conclusion

A constrained optimization problem was formulated to yield minimum weld area without violating the requirement of high penetration and microhardness of the fusion zone. Mathematical relationship among input–output was established using regression, and the obtained equation was validated with experimental results. The following conclusions can be made from the present work.

- The formulated optimization problem was solved using BO and CA, which are swarm intelligence-based algorithm. In addition to this, the solution obtained using the evolutionary genetic algorithm was also taken from available literature for comparison of results among various approaches. The optimal results from different approaches were validated with that of real experiments. The maximum absolute percentage deviation in the prediction of optimal input parameters in all algorithms were found to be less than 4%, which indicates a closer match between algorithms suggested results and experimental results. The absolute percent deviation for all responses was within 2%, and the convergence rate was also faster in BO as compared to that of other approaches. Therefore, it could be concluded that BO performs better than GA and CA based on convergence and accuracy.
- It is advisable to maintain mid-level of voltage and current, a minimum level of welding speed and maximum level of focusing distance to achieve the minimum weld cross section along with a desirable level of penetration and microhardness in weld zone during the electron beam welding of copper alloy.

References

1. Zhong, R.Y., Xu, X., Klotz, E., Newman, S.T.: Intelligent manufacturing in the context of industry 4.0: a review. Engineering **3**(5), 616–630 (2017)
2. Jaypuria, S., Doshi, N., Pratihar, D.K.: Effects of welding parameters on mechanical properties in electron beam welded CuCrZr alloy plates. IOP Conf. Ser.: Mater. Sci. Eng. **338** (1), 012013 (2018)
3. Jaypuria, S., Meher, J., Kanigalpula, P.K.C., Pratihar, D.K.: Electron beam butt welding of Cu-Cr-Zr alloy plates: experimental investigations, studies on metallurgical and mechanical properties. Fusion Eng. Des. **137**, 209–220 (2018)
4. Casalino, G.: Computational intelligence for smart laser materials processing. Opt. Laser Technol. **100**, 165–175 (2018)
5. Benyounis, K.Y., Olabi, A.G.: Optimization of different welding processes using statistical and numerical approaches—a reference guide. Adv. Eng. Softw. **39**(6), 483–496 (2008)
6. Sathiya, P., Panneerselvam, K., Jaleel, M.A.: Optimization of laser welding process parameters for super austenitic stainless steel using artificial neural networks and genetic algorithm. Mater. Des. **1980–2015**(36), 490–498 (2012)
7. Jha, M.N., Pratihar, D.K., Bapat, A.V., Dey, V., Ali, M., Bagchi, A.C.: Knowledge-based systems using neural networks for electron beam welding process of reactive material. J. Intell. Manuf. **25**(6), 1315–1333 (2014)
8. Park, Y.W., Rhee, S.: Process modeling and parameter optimization using neural network and genetic algorithms for aluminum laser welding automation. Int. J. Adv. Manuf. Technol. **37** (9–10), 1014–1021 (2008)
9. Dey, V., Pratihar, D.K., Datta, G.L., Jha, M.N., Saha, T.K., Bapat, A.V.: Optimization of bead geometry in electron beam welding using a genetic algorithm. J. Mater. Process. Technol. **209** (3), 1151–1157 (2009)
10. Kanigalpula, P.K.C., Jaypuria, S., Pratihar, D.K., Jha, M.N.: Experimental investigations, input-output modeling, and optimization of spiking phenomenon in electron beam welding of ETP copper plates. Measurement **129**, 302–318 (2018)
11. Kanigalpula, P.K.C., Pratihar, D.K., Jha, M.N., Derose, J., Bapat, A.V., Pal, A.R.: Experimental investigations, input-output modeling and optimization for electron beam welding of Cu-Cr-Zr alloy plates. Int. J. Adv. Manuf. Technol. **85**(1–4), 711–726 (2016)
12. Pratihar, D.K.: Soft Computing: Fundamentals and Applications. Narosa Publishing House, New Delhi (2013)
13. Das, A.K., Pratihar, D.K.: A new bonobo optimizer (BO) for real-parameter optimization. In: 2019 IEEE Region 10 Symposium. IEEE, Kolkata, India (2019)
14. Canayaz, M., Karci, A.: Cricket behaviour-based evolutionary computation technique in solving engineering optimization problems. Appl. Intell. **44**(2), 362–376 (2016)
15. Canayaz, M., Karcı, A.: Investigation of cricket behaviours as evolutionary computation for system design optimization problems. Measurement **68**, 225–235 (2015)

Chapter 15
Effect of Pulsation on Temperature Distribution of Laser-Welded Dissimilar Joint of Thin Sheet

Bikash Kumar and Swarup Bag

Abstract Dissimilar joints are characterized by microstructural change and compositional gradient which evolve large variation in physical and mechanical properties across the welded joints. The combination of dissimilar material is challenging goal for development of parts with locally optimized parameter especially for Ti-SS joints. Due to high cost (like Ti-alloy), joining of dissimilar material (Ti-SS) helps to reduce the cost incurred by eliminating the otherwise unavoidable usage of costly and rare metal. The present study includes the pulsed laser welding of dissimilar joint of Ti-alloy and AISI 304. EDX analysis is performed on welded specimen for investigation of morphology attained by different zone. It is reported that fusion zone of dissimilar joint contains very fine-grained morphology because grains are not allowed to grow more at high heat input for shorter time span (pulsation). EDX analysis conveys the presence of intermetallic compounds which is majorly composed of Fe, Ti, and Cr. Cooling rate plays important role to determine microstructural variation. But it is difficult to measure it experimentally, so three-dimensional thermal model has been developed and validated with experimental result. Furthermore, average cooling rate is estimated for Ti and SS at different locations across the weld zone. Significant difference in thermo-physical properties like specific heat capacity and conductivity are major parameter due to which large difference has been found in the estimated cooling rate between Ti-alloy and stainless steel alloy at same location at same instant of time.

Keywords Dissimilar joint · Laser welding · EDX analysis · Pulsed thermal modelling · Cooling rate

B. Kumar (✉) · S. Bag
Department of Mechanical Engineering, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India
e-mail: bikash.kumar@iitg.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_15

15.1 Introduction

Owning to massive difference in chemical and physical characteristics, welding of dissimilar metal configuration is, however, a challenging task. In order to seize the privilege of certain attributes of each component to enhance the potential of product, dissimilar welding is accomplished. Particularly, dissimilar joints of Ti-alloy and stainless steel are used in petrochemical, chemical, electronics, nuclear and power generation industries [1]. The application of different alloy and metal for components provides greater flexibility to the designer and production engineer and often results in economic and technical advantage over products manufactured from single material. Costly metal (i.e. Ti-alloy) with eminent properties (i.e. low density, high strength, corrosion resistance and lightweight) can be used in critical location with relatively cheaper material (i.e. steel alloy) in order to improve productivity, production flexibility, weld quality and manufacturing opportunity.

In agreement with Fe–Ti diagram, the solubility of iron (Fe) in titanium is very low at room temperature beyond which intermetallic phases (i.e. Fe_2Ti, FeTi) are accounted [2]. These intermetallic morphologies are brittle in nature and leads to spontaneous crack formation during conventional fusion-welded joint. Mismatching of stress generated by thermal cycle (i.e. thermal stress) between parent metal resembles crack formation. To resolve and minimize the undesirable issue arises due to application of conventional process, low energy input process i.e. laser welding process can be efficiently utilized. Rapid heating and cooling cycle, accurate weld bead position, precise and localized beam, process flexibility, low residual stress and distortion generation are its principle features. Therefore, laser beam process is effective tool to control the brittle intermetallic formation at some magnitude.

Unlike continuous, pulse mode welding is characterized by a series of overlapping spot weld. The resulting pulsed weld is one where subtle variations in the pulsing parameters can lead to dramatic differences in quality of weld [3]. However, continuous waved welding process is accomplished of quasi-steady-state temperature distribution during fusion. Modern Nd: YAG laser process has capacity to shape the flux profile for each pulse at wide range of frequency and introduction of several additional parameters like pulse shaping, pulse energy, pulse duration, average power, etc. Owing to its good coupling efficiency, more flexible beam delivery, particularly pulsed mode Nd: YAG laser is a more convenient process for thin sheet welding as compared to continuous mode of CO_2 laser [4]. A precise amount of beam power can be applied for a short period of time causing small heat-affected zone (HAZ) and low distortion is possible. For the same average power, pulsed Nd: YAG laser can provide comparatively much higher penetration than continuous mode.

Matsunawa et al. [5] reported that Ti-alloy is susceptible towards pore formation due to impingement of hydrogen and other gases in melt pool. To overcome this issue, researcher implemented overlapping factor in pulse mode welding process.

Besides, it is shown that the number of porosity can be minimized with increasing overlapping factor, and even defect-free weld joint could be possible for more than 75% overlapping factor [6]. Few literatures have been found which demonstrate and characterize dissimilar joint of Ti-SS alloy. Hiraga et al. [7] investigated lap welding of Ti and AISI 304 of thickness 0.5 mm with a pulsed Nd: YAG laser to produce vacuum container used as thermo bottles. It revealed that when AISI 304 is used as upper sheet, then joint could be produced within a narrow range of welding condition; however, when Ti is used as upper sheet, joint had been found within wide range. Chen et al. [8] examined the dissimilar joint of Ti-alloy and stainless steel (SS304) of thickness 1 mm by using CO_2 laser and characterized the impact of laser beam offset on the microstructural morphology and mechanical properties of joints. Offsetting of laser beam towards steel results more durable joint having intermetallic compounds like FeTi + α-Ti, FeTi + Fe_2Ti + $Ti_5Fe_{17}Cr_5$ but when laser beam was offset towards Ti-alloy, spontaneous fracture happened. It is also reported that tensile strength of joint was higher when laser beam was offset towards SS304. However, characterization of bead geometry found for dissimilar joint based on EDX analysis is still hazy.

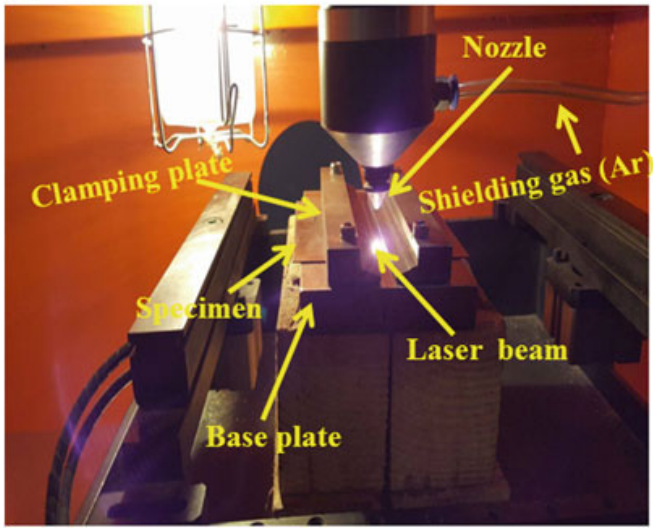
A number of FEM-based numerical models have been developed and investigated by different researcher for the simulation of temperature distribution [9, 10]. To determine the temperature distribution numerically by considering temperature-dependent domain properties and, latent heat of fusion and solidification, author has developed quasi-steady-state heat transfer model [11] for GTA welding. Moreover, the influence of welding velocity on double-ellipsoidal volumetric model has been suggested by another researcher [12]. However, effect of pulsation on the temperature distribution of dissimilar joint for thin sheet (thickness: 1 mm) by implementing volumetric heat source model is not focused so far.

The present study focuses on the primary investigation of experimentally observed macrostructural morphology and EDX analysis for dissimilar joint of Ti–6Al–4V alloy and AISI 304 alloy. Weld zone observed from experimental investigation has been compared with FEM-based 3D thermal (numerical) model. The influence of pulse-on temperature distribution during welding has been considered in the present analysis. Cooling rate plays important role for prediction of mechanical properties and evaluation of microstructure for laser welding process. So, based on numerically extracted temperature–time data, average cooling rate for different zones of each component (Ti-alloy and SS304) for post-welded samples have also been estimated. Furthermore, relation is established between microstructure evolved and cooling rate.

15.2 Experimental Investigation

The plate of 304 grade of stainless steel (60 × 50 × 1 mm^3) and Ti–6Al–4V alloy (60 × 50 × 1 mm^3) is chosen as specimen for pulsed Nd: YAG laser welding (capacity: 500 W) in butt joint configuration. Figure 15.1a depicts the welding

(a)



(b)

Process variable (optimized)	value
Peak power (kW)	1.8
Average power (W)	360
Pulse energy (J)	7.25
Heat input (J/mm)	80
Frequency (Hz)	50
Pulse width (ms)	4
Flow rate (L/min)	8
Weld speed (mm/s)	4.5
Offset distance (mm)	5

Fig. 15.1 **a** Schematic diagram of pulsed Nd: YAG laser welding set-up; **b** optimized parameter for dissimilar weld

set-up which is used to perform experiments. It is composed of fixed laser source, while the specimen along with fixture moves at a constant speed with auto-guided work table. The surfaces of specimens (i.e. Ti-alloy, SS304) are carefully pre-processed or polished to remove oxidized or coating layer and then clamped in fixture. As stable welding exists within a narrow range of parameter so heat input is considered as controlling parameter for present analysis to identify optimum process variables. Specimen is welded at an average power of 360 W with weld velocity of 4.5 mm/s (optimized parameter shown in Fig. 15.1b) for a constant pulse repetition rate of 50 Hz.

Welded specimens are cut across the welding line for metallographic analysis by using wire-electrodischarge machine (W-EDM) to avoid alteration of residual stress during cutting. Macro-graph of weld bead geometry is observed under upright optical microscope (*LEICA-S6D*). In order to investigate the morphology of joint and elemental composition, scanning electron microscope (SEM) outfitted with an energy dispersive X-ray spectroscopy (EDX) is used.

15.3 Theoretical Background

15.3.1 Thermal Analysis

A finite element (FE)-based three-dimensional heat transfer model for dissimilar laser welding has been developed by using ABAQUS commercial software. Temperature-dependent thermo-physical properties such as specific heat capacity, conductivity and constant density, latent heat of fusion are considered for the

simulation. A full-axisymmetric geometrical domain has been considered in which half geometry is assumed as Ti–6Al–4V alloy, whereas another half geometry is considered as stainless steel (304). Due to coherent and collimate nature of laser beam (spot dia. ~1 mm), the weld line is so narrow that suitably fine mesh is necessary to capture steep temperature. Thermal effects due to solidification of melt pool are simulated by considering the latent heat of fusion for both the material (Ti-alloy and SS). The temporal and spatial diversification of pulsed moving heat source is implemented through DFLUX user subroutine. The Gaussian form of heat flux is assumed over volume in such a way that flux density decreases in thickness (depth) direction and pulse attribute is also included over time domain. Double-ellipsoidal volumetric heat source model is considered for the real representation of laser beam [12]. For the inclusion of pulse characteristics, the temporal variation of flux intensity is assumed as 1 for pulse-on time and considered as 0 for pulse-off time. The average absorptivity of Ti-alloy and SS 304 are taken as 0.34 [13] and 0.17 [14], respectively. The nonlinear heat conduction equation based on energy conservation principle is given by

$$\rho_m C_p \left(\frac{\partial T}{\partial t} - V \frac{\partial T}{\partial x_1} \right) = \frac{\partial}{\partial x_i} \left(k_{ij} \frac{\partial T}{\partial x_j} \right) + h_g \tag{15.1}$$

where x_i for $i = 1, 2, 3$ represents x-, y- and z-axis, ρ_m is the density for Ti-alloy and stainless steel, T is the temperature, C_p is the specific heat, h_g is the internal heat generation or power generation under controlled volume, k_{ij} is the component of thermal conductivity and V is the welding velocity along x-axis. The boundary interaction associated with heat transfer between substrate and surrounding environment during fusion welding are implemented for the solution of conduction equation (shown in Fig. 15.2a).

a. Dirichlet boundary interaction, at time

$$t = 0, T(x, y, z, 0) = T_i \, 300 \text{ K} \tag{15.2}$$

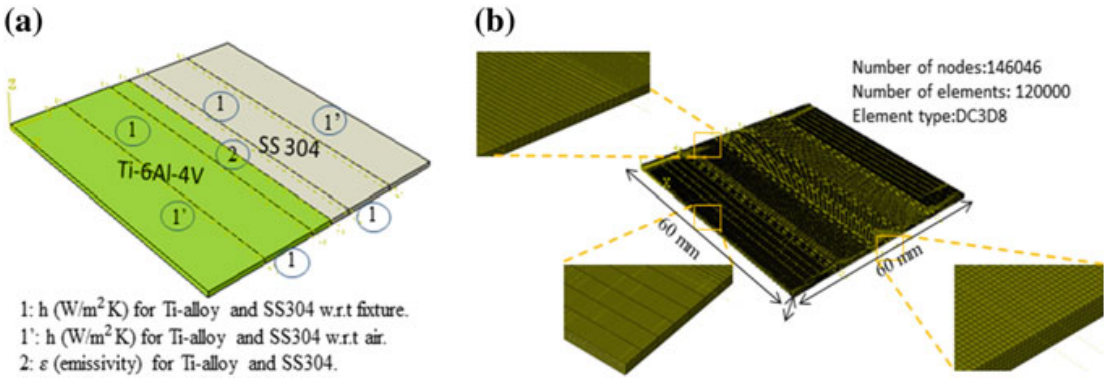


Fig. 15.2 **a** Thermal constraints for dissimilar modelling; **b** meshed geometrical configuration of thermal model

b. Neumann boundary condition; mathematically, heat interactions at all the surfaces can be expressed as

$$k_t \frac{\partial T}{\partial n} - q_f + \varepsilon_r \sigma_r \left(T^4 - T_i^4\right) + h_t (T - T_i) = 0 \tag{15.3}$$

where T_i is ambient temperature, k_t is isotropic thermal conductivity of the material, q_f is Gaussian distributed heat flux impinges on substrate, ε_r is the emissivity, σ_r denoted as Stefan–Boltzman constant (5.67×10^{-8} $\mathrm{Jm^{-2}\ s^{-1}\ K^{-4}}$), h_t is the convective heat transfer coefficient between workpiece and ambient surrounding.

The geometrical configuration of dissimilar welded joints with optimum uniform mesh discretization at fusion zone (FZ) and non-uniform meshing at far away distance for each component is shown in Fig. 15.2b.

15.4 Result and Discussions

15.4.1 EDS Analysis

Energy dispersive X-ray (EDX) spectroscopic analysis is performed to determine the elemental distribution and EDX-based scanning electron microscopic (SEM) investigation is carried to understand morphology attained by dissimilar joint of Ti-SS component. Figure 15.3 shows SEM images of different zones of welded specimen. Figure 15.3a depicts the fusion zone (FZ) which is composed of very fine-grained morphology and its grain boundary. Several small pores are also visible on the surface of grain is due to inclusion of hydrogen gas entrapment at high temperature in Ti-alloy. However, basic microstructure of austenitic stainless (304) is consists of equiaxed, austenitic twinned structure as shown in Fig. 15.3b. The austenitic structure completely transformed into delta ferrite upon heating. During welding, the liquid phase is nucleated at triple networks and grain boundary depending upon interface energy at solidus temperature. Upon further heating, the liquid phase is distributed along the grain boundary and generates a continuous junction along the grain edge as clearly visible in Fig. 15.3a.

Due to pulse phenomena, components experience higher temperature for very small interval of time due to which weld pool lasts for short duration at fusion zone. As time spent by the substrate at high temperature domain decreases, grains are not allowed to grow more during cooling which resembles very fine-grained microstructure in FZ. EDS analysis as spectrum 12 (as shown in Fig. 15.4) confirms that fusion zone is majorly composed of elements like Fe (55%), Cr (15%), C (12%), Ti (6%) and Ni (6%). There is possibility that these elements form intermetallic compound like FeTi + Fe_2Ti + $Ti_5Fe_{17}Cr_5$ + Ni_3Ti and some form of

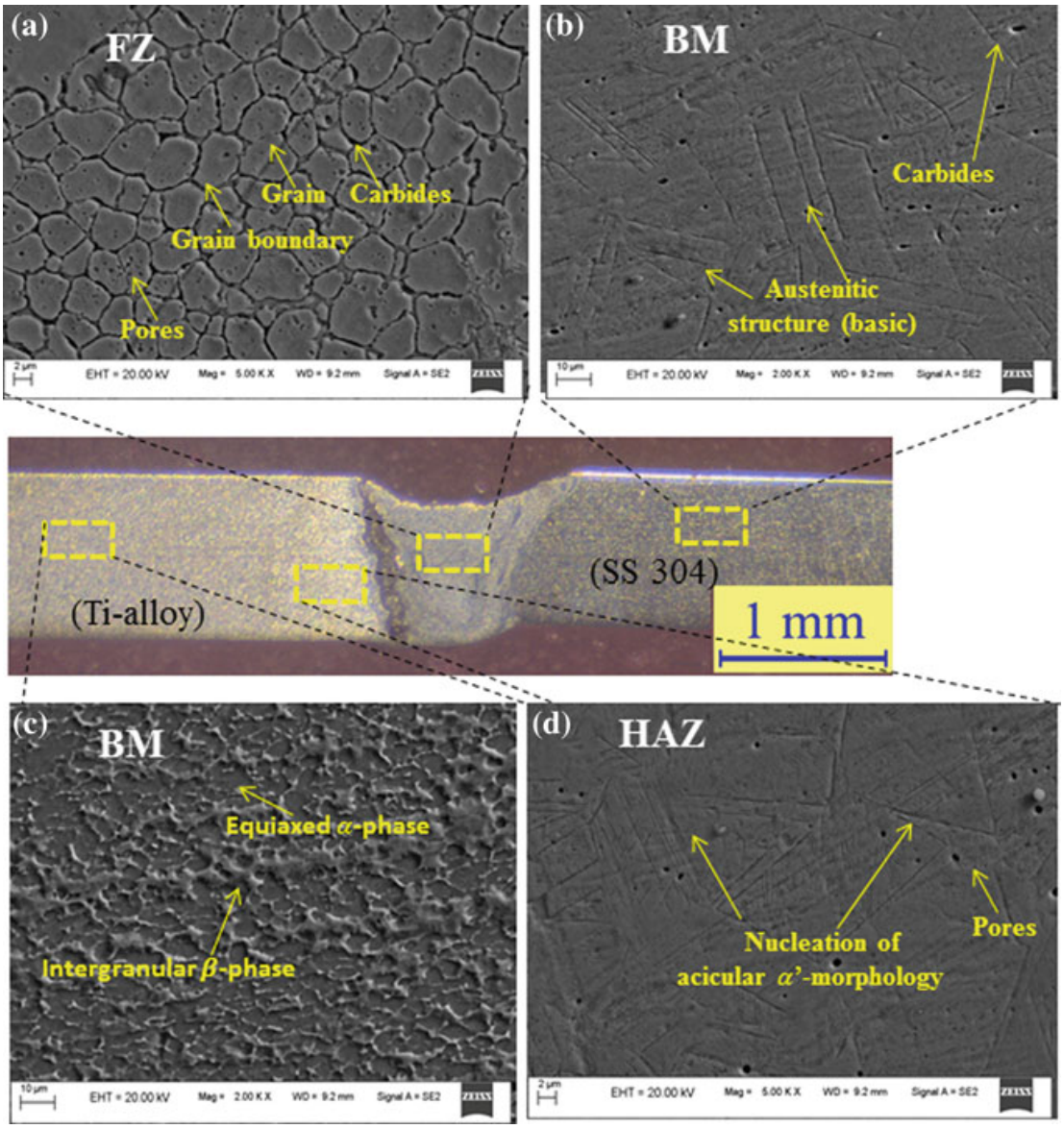


Fig. 15.3 SEM image of **a** FZ; **b** base material (stainless steel); **c** base material (Ti-alloy); **d** HAZ (Ti-alloy)

carbides due to diffusion of Ti-alloy with SS 304 at elevated temperature. Figure 15.3c depicts the basic microstructure of Ti–6Al–4V, consists of equiaxed α-phase (dark region) and line-like white region (light region) as intergranular β-phase. Figure 15.3d signifies the heat-affected zone at the verge of molten zone at Ti-side which clarifies the nucleation of acicular type α' brittle martensitic microstructure.

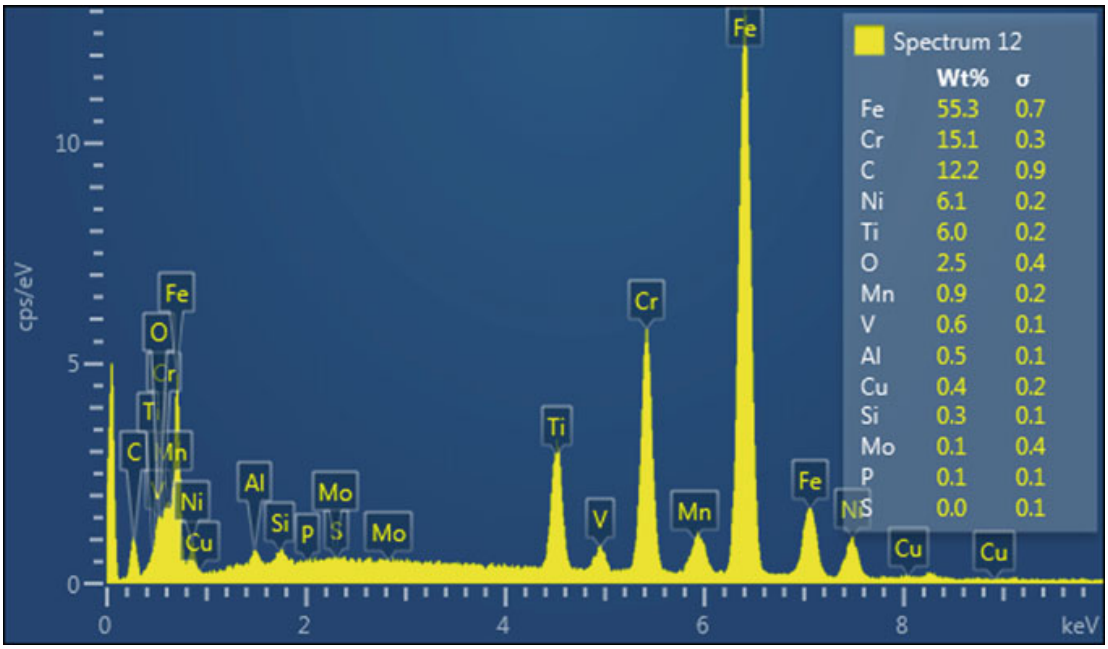


Fig. 15.4 EDS-based spectrum for fusion zone (FZ) of dissimilar weld joint

15.4.2 Numerical Result

FEM-based three-dimensional thermal modelling is undertaken to predict weld bead geometry and to study variation in nodal temperature in both components of dissimilar weld during pulsed laser welding. The numerical simulation of temperature distribution is carried out corresponding to welding speed 4.5 mm/s and average power of 360 W for the heat input of 80 J/mm. Figure 15.5a manifests the comparison between experimental and numerical weld bead profile for dissimilar joint of Ti-alloy and AISI 304. It reveals that macrostructural characteristics (weld bead profile, HAZ) of simulated domain are of similar pattern as anticipated by experimental result. Heat affected a zone which manifests solid-state phase transformation represented by region between blue and yellow colour band, mushy zones are characterized by liquidus (Ti-alloy $\sim$ 1933 K; SS $\sim$ 1723 K) and solidus temperature (Ti-alloy $\sim$ 1877 K; SS $\sim$ 1673 K) is designated as yellow band, whereas FZ is represented by red colour layer for both components. The contour plot of spatial and temporal movement of heat flux in pulsed off and on condition is expressed in Fig. 15.5b. It shows that the red colour band is only visible in pulse-on condition due to attainment of higher temperature for short period which is attributed to entrapment of direct laser beam on the substrate. Experimentally obtained bead profile reveals that the evaporation phenomena also involved for the present optimized process parameter. However, numerical result did not exhibit evaporation occurrence since the maximum temperature attained here is well below of boiling temperature (3500 K). Therefore, it can conclude that conduction mode of welding is prevailing over keyhole model for developed thermal model.

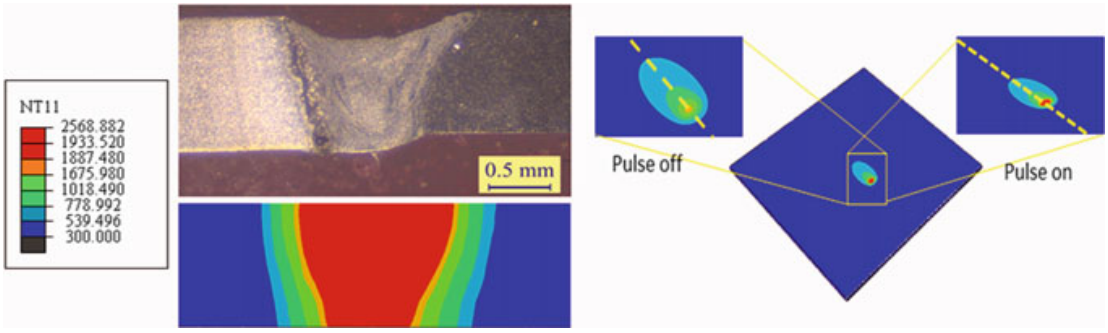


Fig. 15.5 **a** Comparison of experimental and numerical weld bead profile; **b** contour plot of spatial movement of laser bear at pulse-on condition and pulse-off condition

Figure 15.6 depicts the temperature–time variation in Ti-alloy and stainless steel part during pulsed welding. It shows the needle-type structure at peak temperature for both materials. Pulse-on and pulse-off appearances lead to instantaneous temperature gradient at weld line or at closer distance from weld line which is attributed to formation of peak and valley for very short time duration. It is obvious from the figure that peak and valley existence gradually decreases away from the weld line for both the material. Figure 15.6a reveals the attained temperature across the weld line for Ti-alloy side while on stainless steel side is shown in Fig. 15.6b. For both cases, achieved temperature is gradually decreasing as the distance increases from the weld line, and this is due to the fact that at weld line, substrate is in direct contact with laser beam; however, as the distance increases, less amount of heat conduction occurs across the molten zone.

The temperature attained by Ti-alloy is about 2350 K at 0.5 mm distance of weld line, whereas SS 304 achieved 2000 K at same distance across the weld line. The higher diffusive heat transfer coefficient and absorptivity of Ti-alloy as compared with SS304 resemble higher amount of heat diffusion throughout the surface unlike stainless steel. Figure 15.7a conveys that temperature distribution in the

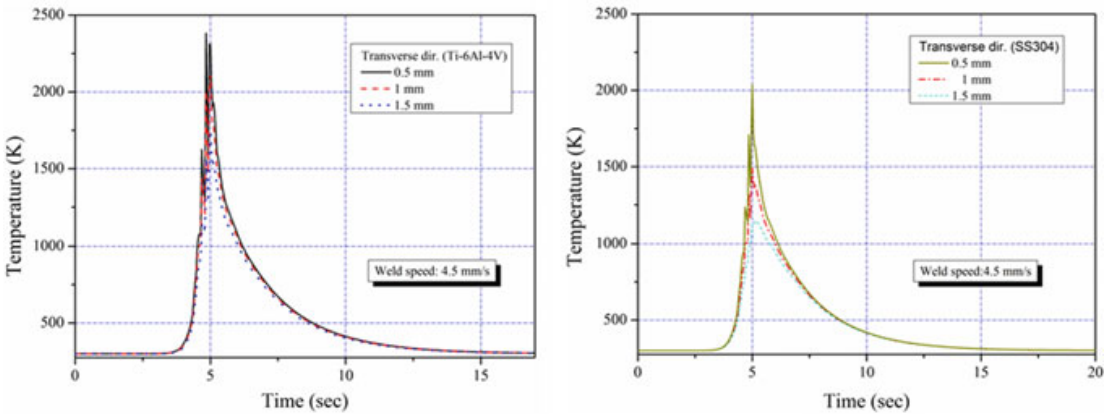


Fig. 15.6 Temperature distribution at different location across the weld zone, **a** Ti–6Al–4V alloy; **b** SS 304

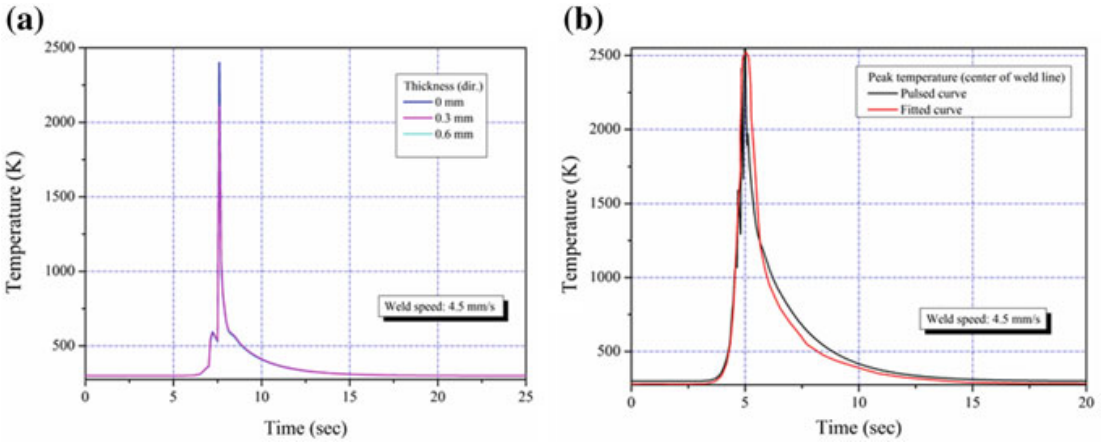


Fig. 15.7 **a** Temperature profile in thickness direction at weld line; **b** temperature profile obtained at mid of weld line in pulsed condition and fitted curve by Gaussian function

component in thickness direction at centre point of weld pool. At the top of the surface, attained temperature is the same as that of peak temperature ($\sim$2500 K). Drop of the temperature throughout the thickness is very low up to 0.5 mm thereafter it increases but not so much because of the thinness of the weldments.

15.4.3 Cooling Rate Estimation

Researchers have already described that cooling rate corresponding to laser welding of thin sheet plays significant role for the prediction of different microstructures attained by weldments and its associated mechanical properties. Senkov et al. [15] investigated the effect of controlled cooling rate on the morphology and mechanical properties of Ti-alloy. The weld thermal cycle mainly consists of continuous heating and followed by cooling it up to ambient environment. It is unsuitable or crucial to determine the attained temperature–time history during welding at fusion zone due to involvement of melting phenomena. Hence, simulation-based time–temperature data are implemented for the estimation of cooling rate at different location of both components of weldments.

In the present work, pulsed mode of welding is considered for the investigation so the temperature profile exhibits instantaneous temperature variation with respect to time at higher temperature. For the sake of simplicity, the achieved temperature profile at weld zone is curve fitted by using Gaussian basic function which provided satisfactory continuous curve. Fitted continuous curve is implemented for the calculation of cooling rate in the present study as depicted in Fig. 15.7b. The temperature at which remarkable phase transformation occurs during cooling cycle is considered for the estimation of cooling rate here. Kumar et al. [16] reported the alpha-dissolution and beta-transus temperature for the dual-phase Ti-alloy as 940 and 1275 K. Mathematical formulation for the estimation of cooling rate is also

revealed in the literature. The phase transformation temperature range for stainless steel is assumed as 773 and 1073 K.

Figure 15.8 shows the temperature distribution at different location across the weld line within assumed phase transformation temperature range for Ti–6Al–4V and steel alloy. These curves reveal that temperature profile within this range is approximately linear for both components. Therefore, it is obvious that slope of these curve can predict fairly accurate average cooling rate. Figure 15.8a depicts that the time spent for the attainment of phase transformation temperature at a distance of 0.5 mm (closer to weld line) is comparatively more than (i.e. 1.4 s) that of 1 mm or 1.5 mm across the weld line for steel alloy. Due to direct contact of weld line with heat flux, nearby region of fusion zone experiences high heat input which sustain weld pool for comparatively longer period of time. Therefore, it takes higher solidification time. Hence, average cooling rate is lower (i.e. 220 K/s) at this location, and at 1 and 1.5 mm location, cooling rate are 258 and 277 K/s, respectively.

Figure 15.8b conveys the temperature distribution within alpha-dissolution temperature and beta-transus temperature range for Ti-alloy. It shows that 0.84 s of total time is elapsed for the attainment of phase transformation temperature at the location of 0.5 mm across the weld centre while somewhat less time is taken at the location of 1 and 1.5 mm across the weld line. So, the estimated cooling rate at closer (weld line) distance is comparatively less (i.e. 398 K/s), whereas cooling rate attained at 1 and 1.5 mm distance are 420, 424 K/s, respectively.

It is observed that the estimated cooling rate at same location of two different materials is having remarkable difference, i.e. 220 K/s for stainless steel and 398 K/s for Ti–6Al–4V at 0.5 mm distance from weld centre. Because, cooling rate eminently associated with the thermal behaviour of material. Accomplishment of temperature depends upon the thermo-physical properties like specific heat, diffusion coefficient, thermal conductivity, etc. So, it can deduce that significant difference in these properties of Ti-alloy and SS 304 is responsible for great difference in cooling rate.

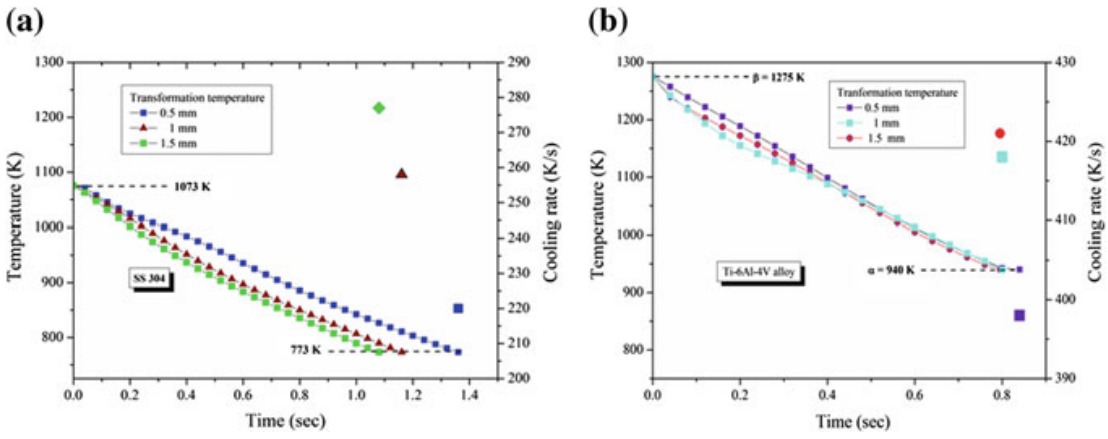


Fig. 15.8 Temperature profile within the range of phase transformation and cooling rate, **a** stainless Steel 304; **b** Ti–6Al–4V alloy

It is already reported in the literature that microstructural evolution can be determined with the help of appropriate cooling rate calculation. It is well established fact that 400 K/s is critical cooling rate for dual-phase Ti-alloy, up to which massive diffusional nucleated secondary lamellar structure can be obtained and beyond which hexagonal acicular α' martensitic structure is possible. Widmanstatten weave-basket-like morphology can be produced at very low cooling rate in dual-phase Ti-alloy. EDS analysis reveals that the nucleation of acicular α' martensitic structure (Fig. 15.3d) has been found at heat-affected zone (nearby region of fusion zone) and it is also observed that the estimated cooling rate at HAZ location is approximately 420 K/s which is well higher that critical cooling rate as reported in the literature. Therefore, it can be concluded that the estimated average cooling rate based on numerical model is fair and accurate and showing good agreement with experimentally observed morphology in the present study.

15.5 Conclusion

Pulsed laser butt welding of titanium alloy to stainless steel has been performed at optimized process variable. EDS-based morphological study and FEM-based temperature distribution have been studied in the present work. Based on this study, conclusions are summarized as following:

- After solidification of weld zone, fine-grained morphology is revealed. Because, due to pulsation effect, high temperature attained at weld zone is only for small period of time that lasts weld pool for shorter period. So, nucleated grain during re-solidification from liquid to solid phase is not allowed to grow more, whereas nucleated acicular α' martensitic structure is observed in HAZ of Ti-alloy side.
- Energy dispersive X-ray (EDX) spectroscopic analysis confirms that FZ majorly contains Fe, Cr, C and Ti elements that resemble formation of intermetallic compound like FeTi + Fe_2Ti + $Ti_5Fe_{17}Cr_5$ + Ni_3Ti during fusion of both components (Ti-alloy and SS 304) at high temperature.
- Good agreement has been observed between experimentally and numerically welds bead profile. However, evaporation of material at top surface is clearly shown in experimental result but attained maximum temperature of weld zone is well below the boiling temperature so an evaporation phenomenon is neglected in 3D modelling.
- The maximum temperature attained at Ti-side is higher than that of stainless steel side. Significant difference in diffusion coefficient and absorptivity is the possible reason for the same.
- As it is difficult to measure temperature–time data at weld zone experimentally, so cooling rate is estimated at different location (i.e. at a distance of 0.5, 1 and 1.5 mm across the weld line) by using numerical simulated result. Maximum cooling rate has been found at 1.5 mm distance for both of the components of weldments, i.e. 424 K/s for Ti-alloy and 277 K/s for stainless steel.

- It is revealed that remarkable difference in thermo-physical properties such as specific heat capacity, conductivity, diffusivity, etc. are the feasible reason for the large difference in cooling rate of both of the material at the same location.
- Estimated average cooling rate for HAZ of Ti-alloy has predicted similar type of morphology as observed in experimental investigation and also reported by the other researchers.

Acknowledgements The authors gratefully acknowledge Central Instrument Facility (CIF) and Central workshop of IITG for providing experimental facility.

References

1. Sun, Z., Ion, J.C.: Laser welding of dissimilar metal combinations. J. Mater. Sci. **30**(17), 4205–4214 (1995)
2. Mousavi, S.A.A.A., Sartangi, P.F.: Effect of post-weld heat treatment on the interface microstructure of explosively welded titanium–stainless steel composite. Mater. Sci. Eng. A **494**(1-2), 329–336 (2008)
3. Lapsanska, H., Chmelickova, H., Hrabovsky, M.: Effect of beam energy on weld geometric characteristics in Nd: YAG laser overlapping spot welding of thin AISI 304 stainless steel sheets. Metall. Mater. Trans. B **41**(5), 1108–1115 (2010)
4. Tzeng, Y.-F.: Parametric analysis of the pulsed Nd: YAG laser seam-welding process. J. Mater. Process. Technol. **102**(1-3), 40–47 (2000)
5. Matsunawa, A., et al.: Dynamics of keyhole and molten pool in laser welding. J. Laser Appl. **10**(6), 247–254 (1998)
6. Gao, X.-L., et al.: Porosity and microstructure in pulsed Nd: YAG laser welded Ti6Al4V sheet. J. Mater. Process. Technol. **214**(7), 1316–1325 (2014)
7. Hiraga, H., et al.: Nd: YAG laser welding of pure titanium to stainless steel. Weld. Int. **16**(8), 623–631 (2002)
8. Chen, S., et al.: Microstructures and mechanical property of laser butt welding of titanium alloy to stainless steel. Mater. Des. **53**, 504–511 (2014)
9. Bag, S., et al.: Efficient estimation of volumetric heat source in fusion welding process simulation. Weld. World **56**(11-12), 88–97 (2012)
10. Baruah, M., Bag, S.: Influence of heat input in microwelding of titanium alloy by micro plasma arc. J. Mater. Process. Technol. **231**, 100–112 (2016)
11. Bag, S., De, A.: Development of a three-dimensional heat-transfer model for the gas tungsten arc welding process using the finite element method coupled with a genetic algorithm–based identification of uncertain input parameters. Metall. Mater. Trans. A **39**(11), 2698–2710 (2008)
12. Yadaiah, N., Bag, S.: Effect of heat source parameters in thermal and mechanical analysis of linear GTA welding process. ISIJ Int. **52**(11), 2069–2075 (2012)
13. Akbari, M., et al.: Experimental and numerical investigation of temperature distribution and melt pool geometry during pulsed laser welding of Ti6Al4V alloy. Opt. Laser Technol. **59**, 52–59 (2014)
14. Zhu, X.K., Chao, Y.J.: Numerical simulation of transient temperature and residual stresses in friction stir welding of 304L stainless steel. J. Mater. Process. Technol. **146**(2), 263–272 (2004)

15. Senkov, O.N., et al.: Effect of cooling rate on microstructure of Ti-6Al-4V forging. Mater. Sci. Technol. **18**(12), 1471–1478 (2002)
16. Kumar, B., Kebede, D., Bag, S.: Microstructure evolution in thin sheet laser welding of titanium alloy. Int. J. Mechatron. Manuf. Syst. **11**(2-3), 203–229 (2018)

Chapter 16
Activated TIG Welding of AISI 321 Austenitic Stainless Steel for Predicting Parametric Influences on Weld Strength of Tensile Test—Experimental and Finite Element Method Approach

S. Mohan Kumar, N. Siva Shanmugam and K. Sankaranarayanasamy

Abstract Tungsten inert gas (TIG) welding is the most common manufacturing process used to join materials like stainless steel, titanium and aluminium alloys due to their high-quality and inexpensive welds. Furthermore, the difficulty in welding plates having a thickness of 4 mm or more is overcome by employing activated TIG (A-TIG) welding process which uses an activating flux for high depth of penetration in a single pass. Uni-axial tensile test is conducted to assess the tensile strength of the material by experimentation. In this study, a uni-axial tensile test of base metal (BM) and weld metal (WM) samples are simulated using ABAQUS to evaluate the accuracy of finite element (FE) simulation results with the experimental results for predicting the tensile strength. The results show that the stress–strain values predicted by the FE analysis agree with experimental results. Also, the fracture behaviour of experimentation and FE simulation is identical with ductile mode of fracture. The fracture location of the sample in FE analysis is found very similar to experimental fractured samples. The results of ferrite measurement indicate that concentration of delta-ferrite in the WM (5.9 FN) is higher than BM (1.2 FN) content and show better mechanical behaviour in the A-TIG weldments. Also, scanning electron microscope (SEM) shows that the failure of BM and WM resembled to ductile mode-type fracture.

Keywords A-TIG welding · AISI 321 · Tensile test · Finite element analysis · Fractography

S. Mohan Kumar · N. Siva Shanmugam (✉) · K. Sankaranarayanasamy
Department of Mechanical Engineering, National Institute of Technology Tiruchirappalli, Tiruchirappalli 620015, Tamil Nadu, India
e-mail: nsiva@nitt.edu

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_16

16.1 Introduction

AISI 321 material has excellent resistance to corrosion [1]. Widely used in automotive industries, jet aircraft components, welded pressure vessels in the defence sector and high-speed impact on firewalls, the encounter of static and dynamics loads on AISI 321 is a common problem. The demand for the controlled weld shape and deeper penetration in TIG welding led to the novel modification of the well-known TIG process, called as A-TIG welding. A-TIG welding was introduced during the 1960s by the E.O. Paton Institute of Electric Welding [2]. The A-TIG joining process uses a thin layer of flux which is spread over the welding region before welding [3]. Tensile testing provides engineers with a wealth of data about the mechanical behaviour of the metals, elastic modulus, tensile strength, yield strength, the percentage of elongation and the stress–strain relationship of the materials. Hence, it is essential to understand the mechanical behaviour and metallurgical aspects that affect the results of the tensile test [4]. Reisgen et al. [5] conducted an FE analysis of uni-axial tensile test to evaluate the tensile strength of laser-welded TRIP700 steel sheet. Garcia-Garino et al. [6] carried out an FE simulation of tensile test for a metal specimen and found good agreement with the analytical and experimental results. Patel et al. [7] studied the effect of activated fluxes like SiO_2 and TiO_2 on the mechanical behaviour of AISI 321 using the TIG welding. However, research works related to FE simulation of tensile test of A-TIG-welded samples are few. Hence, in the present study, tensile test was carried out on base- and A-TIG-welded butt joint of AISI 321 and compared with the experimental and FE analysis results for validation.

16.2 Experimentation Procedure

Bead on plate trials is performed to choose optimum welding parameter. Selection of perfect welding current, arc length and welding speed defines the requirement of the good weld bead. The material selection in the present study is commercially available type AISI 321 austenitic stainless steel. Tables 16.1 and 16.2 show the chemical composition and the physical properties of AISI 321, respectively. Welding equipment model Fronius MagicWave 4000 is used for joining AISI 321 with experimental set-up shown in Fig. 16.1. Particularly during A-TIG joining process, a very thin film of activating flux is painted to the material top surface

Table 16.1 Elemental composition of AISI 321 (wt%)

Ni	Cr	Mn	Si	C	P	S
9.356	17.560	1.007	0.349	0.032	0.031	0.016
Fe	W	Ti	Cu	Mo	Al	Nb
70.775	0.036	0.356	0.356	0.113	0.006	0.003

Table 16.2 Physical properties of AISI 321

Density (kg/m^3)	Young's modulus (GPa)
8090	193

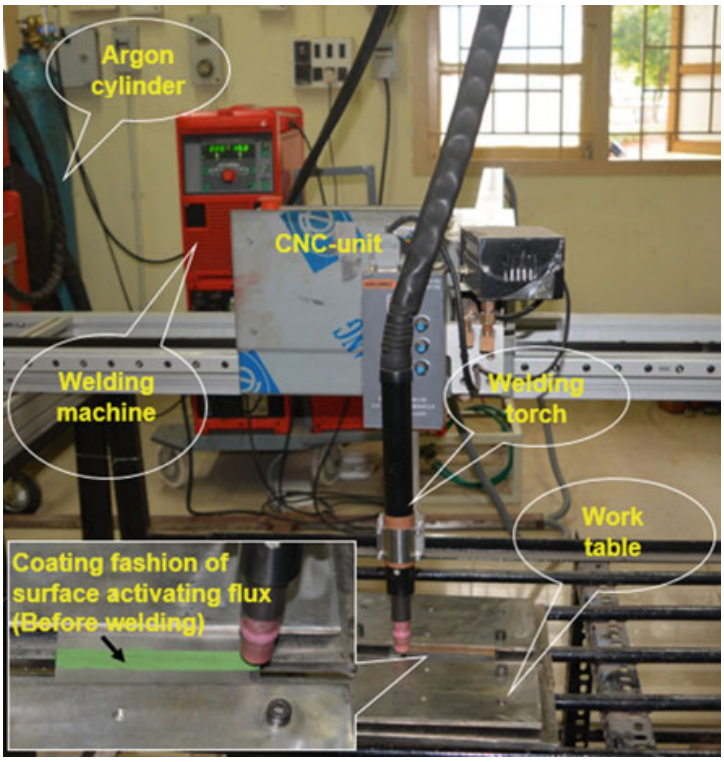


Fig. 16.1 Experimental set-up for A-TIG welding

before joining so as to increase the depth of penetration (DOP). Based on our previous study [8], the commercial Edison Welding Institute (EWI) flux is used in this work, and corresponding ingredients of activated flux are shown in Table 16.3. This increase in DOP is attributed to the mechanisms of arc constriction and Marangoni reversal flow. Before joining the AISI 321 plate of 6 mm thickness was cut into strips of 100 × 100 mm and roughly polished with sheets of silicon carbide to remove the impurities in surface and further scrubbed with acetone. The input weld parameters selected after examinations from the trials of the bead on the plate are shown in Table 16.4. Figure 16.2 represents the welded plate with butt joint configuration. Further, Fig. 16.2b shows the macrostructure of the butt joint with achieved full depth of penetration. The sample for the tensile test is cut from the plate specimen following ASTM E8/E8M-16a standard as shown in Fig. 16.2c.

Table 16.3 Composition of EWI flux

Flux ingredients	Weight (%)
Silicon oxide (SiO_2)	0–20
Chromium III oxide (Cr_2O_3)	0–50
Titanium oxide (TiO_2)	20–60

Table 16.4 A-TIG welding input process parameters

Process parameters	Values	Units
Welding current	220	Amps.
Welding speed	120	mm/min
Arc length	3	mm
Welding voltage	14.9	V
Welding heat input	1.4751	kJ/mm
Electrode diameter	2.4	mm
Shielding gas flow rate (constant—argon purity 99.99%)	20	l/min

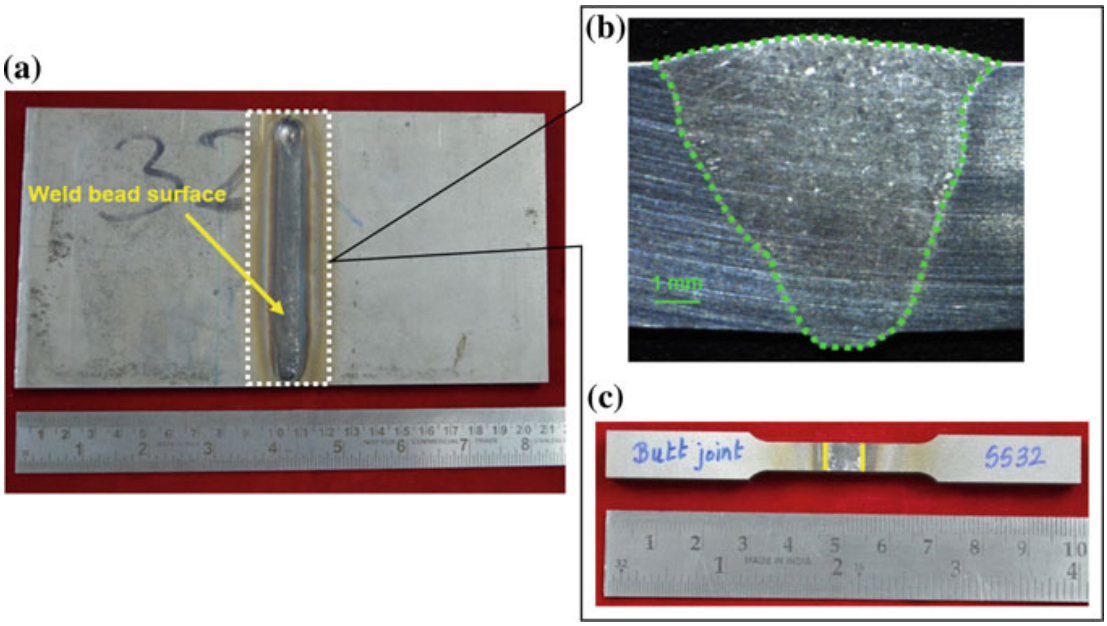


Fig. 16.2 **a** Butt weld geometry made of A-TIG welding, **b** macrostructure of butt joint and **c** sample for the tensile test

16.3 Results and Discussion

16.3.1 Experimental and FE Simulation of the Tensile Test

The AISI 321 tensile specimens of BM and WM were prepared as per standard. The wire-cut EDM process was used to section the specimens. Tensile tests were carried out using Tinius Olsen 50 KN load cell at room temperature with a loading rate of 1 mm/min until failure.

FE analysis of tensile test was performed using ABAQUS—a commercial available FE analysis software. Brief illustration of FEM procedure is shown in Fig. 16.3. The part was modelled in ABAQUS, and physical material properties such as density, Young's modulus and Poisson's ratio assigned are 8090 kg/m^3,

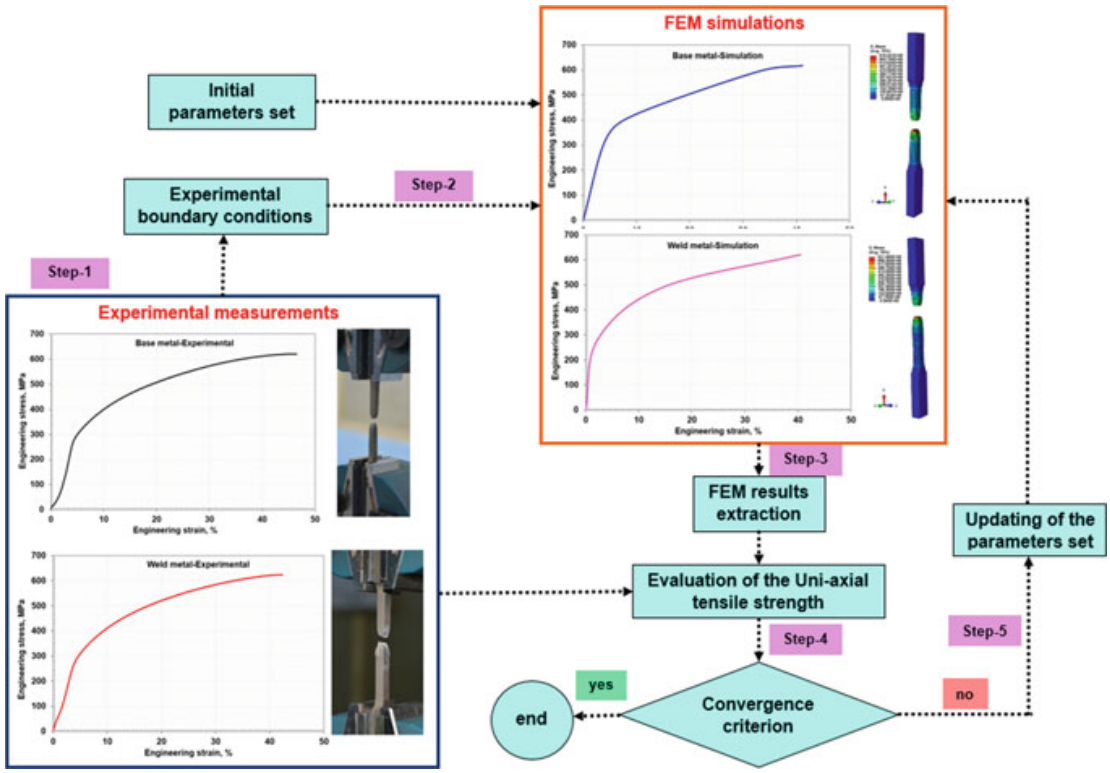


Fig. 16.3 Flow chart of FE method for uni-axial tensile test

Table 16.5 Tensile properties of AISI 321

Part	Yield stress (MPa)	Yield strain (%)	Tensile stress (MPa)	Tensile strain (%)
BM	213	3.536	621	46.56
WM	215	3.088	624	42.4

193 GPa and 0.3, respectively. From uni-axial tensile test, mechanical properties obtained for BM and WM specimens are assigned for plasticity as shown in Table 16.5.

The dynamic explicit analysis was conducted in consideration of nonlinear geometry. Suitable damage material model is necessary to initiate fracture and subsequent failure in the FE model. It is assumed that the material behaviour is isotropic. Hence, the Johnson–Cook model is utilized for analysing the failure. Johnson–Cook (JC) material model is one such model used in the present analysis for material failure, considering the effects of pressure, temperature, equivalent stress and strain rate. JC model's damage parameters assigned for ductile damage in AISI 321 are shown in Table 16.8 [9]. Based on the critical literature review, it is understood that the behaviour of both AISI 316L and AISI 321 austenitic grade stainless steels is almost same. For confirming the similarity in mechanical properties, a uni-axial tensile test has been conducted for both AISI 316L and AISI 321. From the uni-axial tensile test results (Fig. 16.4), the following data are extracted,

Table 16.6 Uni-axial tensile test results of AISI 321 and AISI 316L

Material name	Yield stress (MPa)	Tensile stress (MPa)	Tensile strain (%)
AISI 321	213	621	46.56
AISI 316L	208	594	40.96

Table 16.7 Physical properties

Material name	Density (kg/m^3)	Young's modulus (GPa)
AISI 321	8090	193
AISI 316L	8000	193

Table 16.8 JC damage parameters assigned for AISI 321

D1	D2	D3	D4	D5
0.05	3.44	2.12	0.002	0.61

and the same have been presented in Table 16.6. The density and Young's modulus are also similar in nature (refer Table 16.7).

Owing to the similarities between the two grades of ASS, the JC damage parameters of AISI 316L material are utilized in the present study for performing the tensile test simulation of AISI 321.

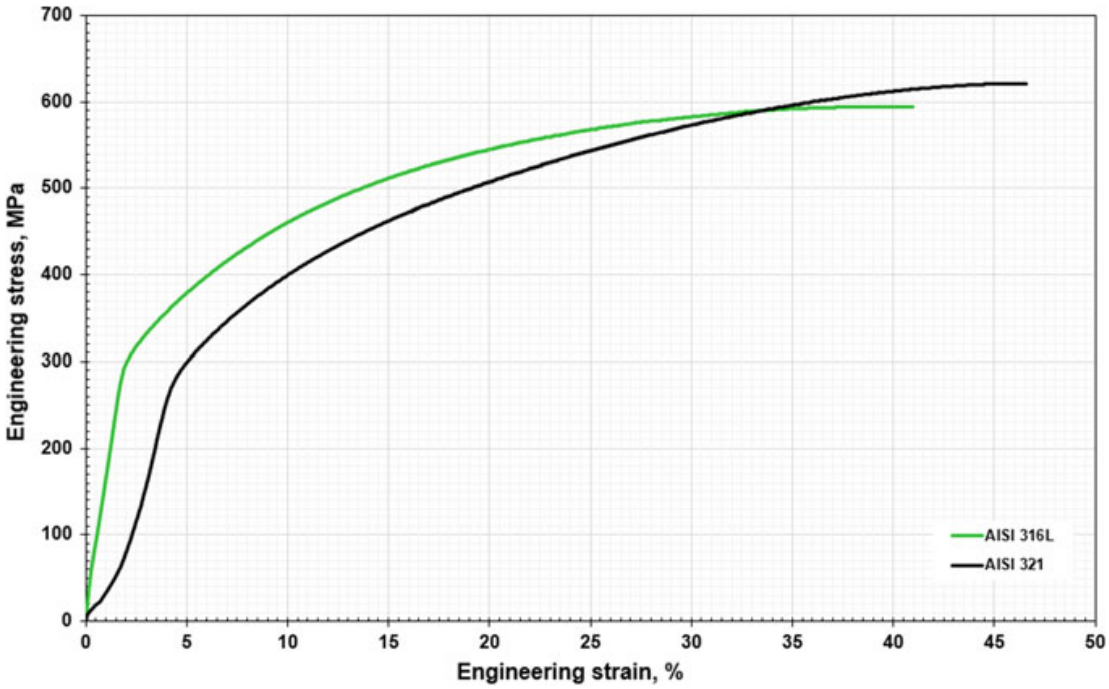


Fig. 16.4 Stress-versus-strain plot for AISI 321 and AISI 316L

The predicted results from the simulation of the uni-axial tensile test also confirm with an error % of 0.51 and 0.48 for BM and WM of AISI 321 steel in comparison with experimental tensile test. JC material model is mathematically expressed by Eqs. 16.1 and 16.2.

The failure is expected when parameter $D = 1$

$$D = \sum \frac{\Delta\varepsilon}{\varepsilon_f} \tag{16.1}$$

$$\bar{\varepsilon}_f = [D_1 + D_2 \exp D_3\sigma^*][1 + D_4 \ln \dot{\varepsilon}^*][1 + D_5 T^*] \tag{16.2}$$

where

$\Delta\varepsilon$ is the increment of effective plastic strain during an increment in loading.
$\bar{\varepsilon}_f$ is the equivalent strain to fracture.
$\dot{\varepsilon}^*$ is the dimensionless strain rate.
T^* is the homologous temperature.
σ^* is the average of the three normal stresses to the von Mises equivalent stress assumed $\sigma^* \leq 1.5$.
D_1, D_2, D_3, D_4 and D_5 are material-dependent constants.

The experimental setting provides boundary conditions for the tensile specimen. One end of the specimen is ENCASTERED (U1 = U2 = U3 = UR1 = UR2 = UR3 = 0), where movement is restricted in all degrees of freedom, while another end is allowed to move in the X-direction for BM and WM (refer Fig. 16.5a, b).

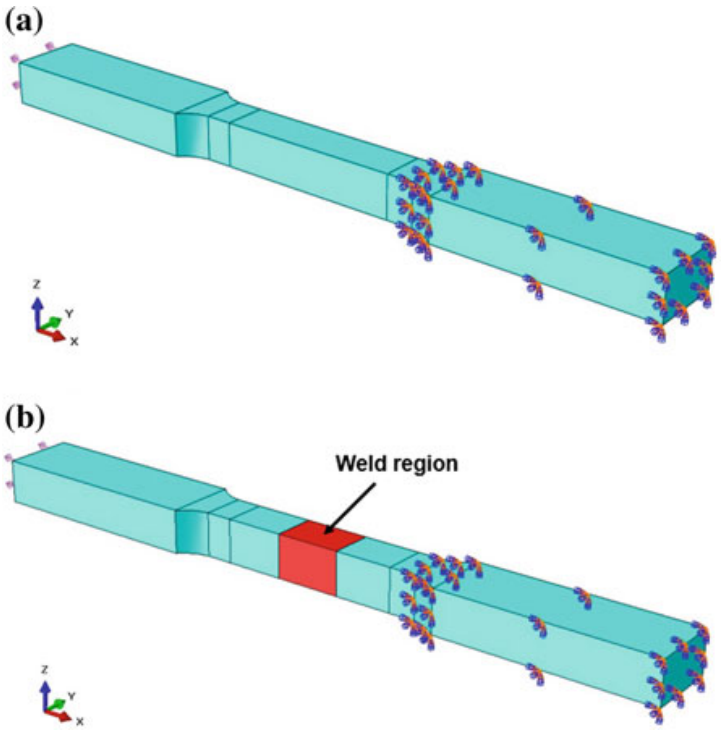


Fig. 16.5 Boundary conditions for the modelled tensile test specimens, **a** BM and **b** WM

Table 16.9 Meshing parameters used for uni-axial tensile test simulation

Part	Shape of element	Meshing technique	Elements generated	Element type
BM and WM	Hexagonal	Structured	6672	C3D8R (an eight-node linear brick, reduced integration with hourglass control)

The meshing of the model is conducted by employing fine mesh at the gauge length section and coarse mesh at the gripping section (to reduce the computational time) as results are expected only at gauge length section. The parameters used for meshing in FE analysis are shown in Table 16.9. Further, the meshed tensile test model of BM and WM is described in Fig. 16.6a, b separately.

Figure 16.7a shows the experimental tensile test depicting the fracture behaviour of the BM. Similarly, a tensile test conducted using the FE model is shown in Fig. 16.7b with a subsequent fracture. A good match can be observed in fracture at the centre of the specimen in both the analyses. Also, the fracture behaviour of the BM and WM samples was observed with the ductile mode of fracture and found similar with FE analysis as shown in Figs. 16.8 and 16.10. Likewise, for WM, the fracture region obtained after performing experimental and FE analysis is shown in Fig. 16.9a, b, respectively. From our earlier investigations [8], the mechanical properties of HAZ were better than the BM using A-TIG process. From the results

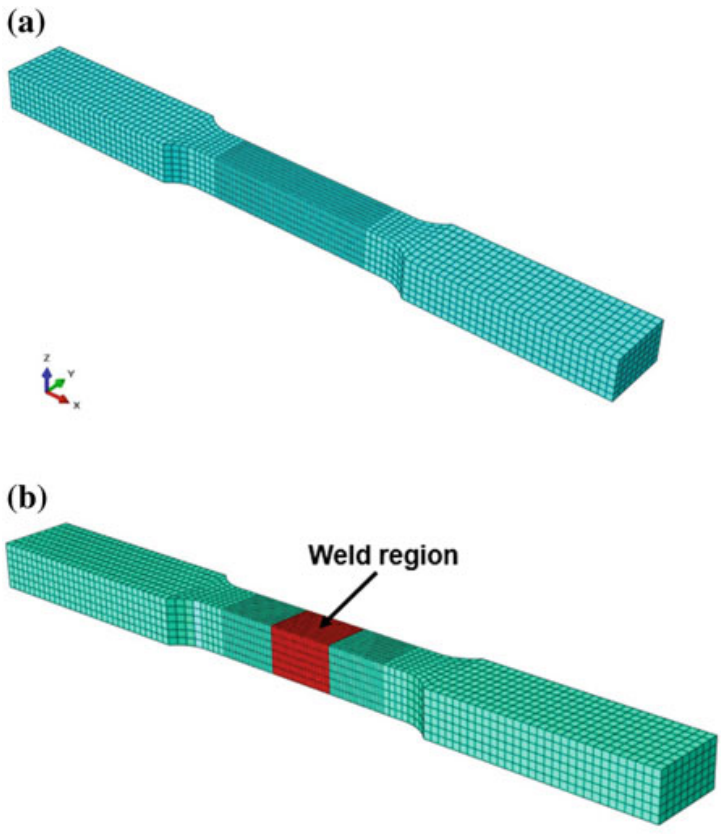


Fig. 16.6 Modelled tensile test specimens with mesh, **a** BM and **b** WM

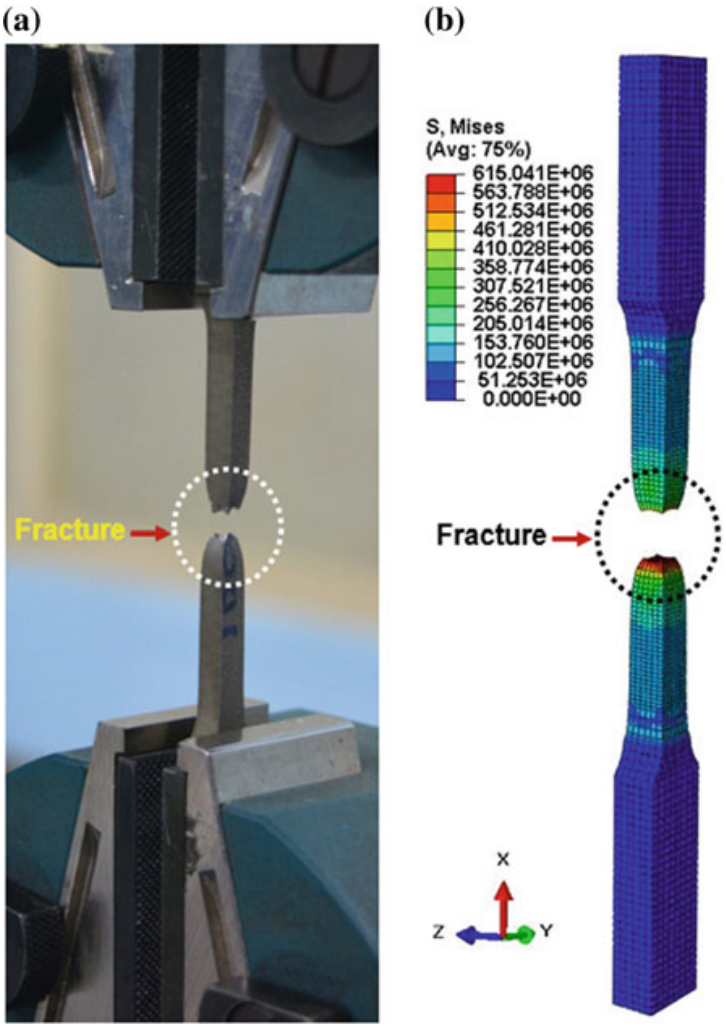


Fig. 16.7 Comparison of the BM tensile test between experimental (**a**) and FE analysis (**b**)

of uni-axial tensile test on both BM and WM, it is evident that the fracture occurred in the BM regions. Hence, the effect of heat-affected zone in the FE simulation is not considered. Excellent agreement is revealed through visual comparison as fracture occurred always in the BM, i.e. ductile mode of failure as shown in Fig. 16.10 between experimental and FE analysis.

The stress–strain curves determined from the tensile test are shown in Fig. 16.11. The tensile strength for the BM and WM tested experimentally at 621 and 624 MPa, respectively, while the FE simulation tensile strength for the BM and WM is 617.8 and 621 MPa, respectively. Table 16.10 presents the tensile strength and the crack position for BM and WM specimens. It was observed that welded specimen has slightly higher tensile strength compared to base, and percentage of elongation of weld specimen decreased considerably when compared to the BM specimen in experimental and FE analysis. From Table 16.10, it is evident that excellent result is obtained in terms of tensile strength for base and weld metal with the percentage of error 0.51 and 0.48%, respectively. Results of ferrite measurement (Fischer FERITSCOPE FMP30) in BM and WM zones of AISI 321 are 1.2 and 5.9 FN, respectively. It is confirmed that the increase in tensile strength and decrease in

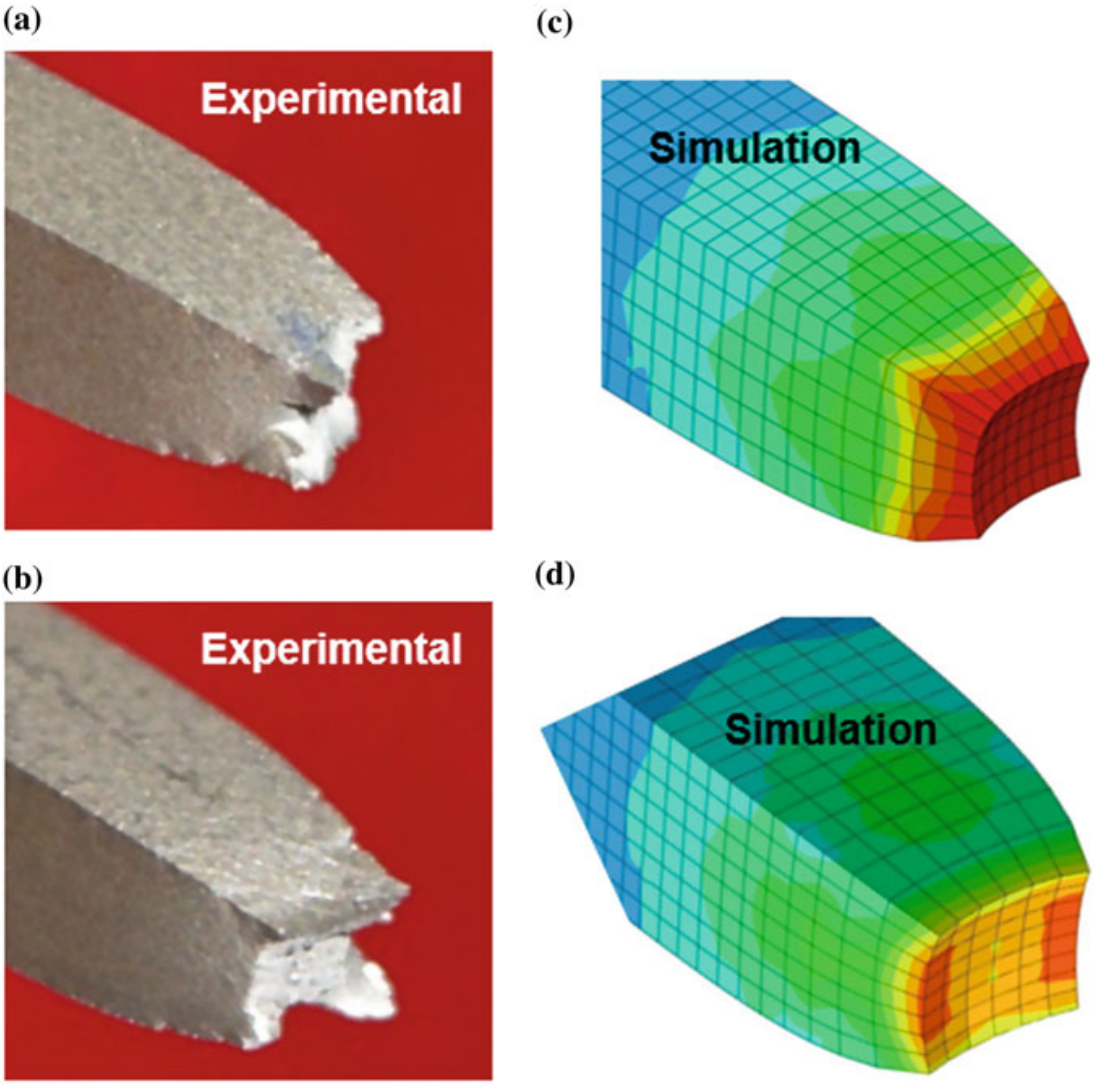


Fig. 16.8 Comparison of BM tensile test fracture surface mode between experimental (**a** and **b**) and FE analysis (**c** and **d**)

ductility are attributed to the hike in the delta-ferrite content. The tensile test results can corroborate that the A-TIG WM tensile strength is better than BM [10]. The experimental and FE simulation joint efficiency of the WM is 101.1 and 100.51%, respectively. The fractured specimens also showed a ductile fracture.

16.3.2 Fractography Examination

The fractured tensile specimen is cleaned ultrasonically and examined using SEM [11]. The tensile strength of the WM is higher than that of the BM only due to the influence of the welding heat input leading to rearrangement of the molecular orientation. Further, as is clear from Figs. 16.12 and 16.13, it has characteristics

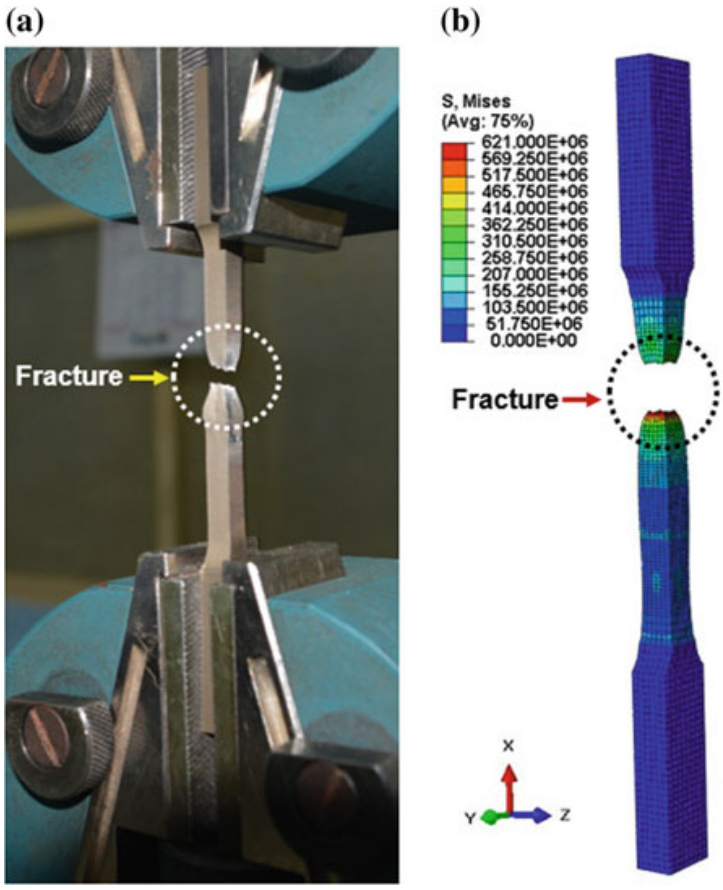


Fig. 16.9 Comparison of the WM tensile test between experimental (**a**) and FE analysis (**b**)

similar to those of the BM and the WM. It is understood from the fracture surface of the WM sample that there are microvoids and fine dimples as compared with the BM and such fine dimples contribute to the improvement of the load-carrying capacity of the WM. Also, on the fractured surface of the WM and the BM, it has similar characteristics, that is, the evolution of a honeycomb structure with voids and dimples confirming the ductile failure.

16.4 Conclusions

The following most important inferences are summarized from the current study.

- A-TIG welding was performed on AISI 321 with optimal input parameters achieving good depth of penetration and bead width for butt joint configuration.
- Experimental tests performed on BM and WM were evaluated for mechanical strength. The uni-axial tensile test showed higher strength for WM—624 MPa compared to BM—621 MPa. The increase in weld strength and reduction in ductility are due to the increase in the delta-ferrite content.
- The prediction by FE analysis was observed with a tensile strength of 621 and 617.8 MPa for WM and BM, respectively.

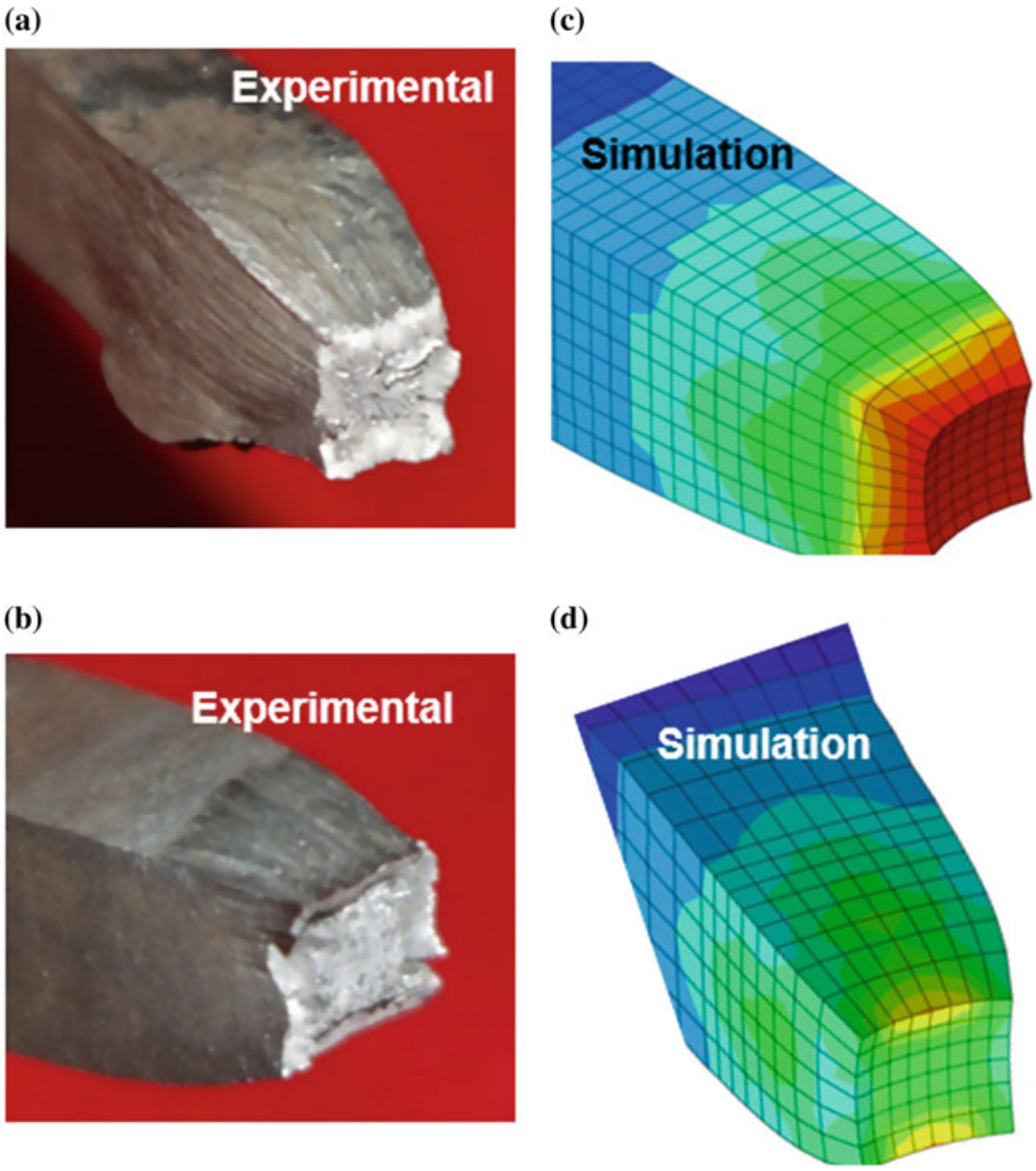


Fig. 16.10 Comparison of WM tensile test fracture surface mode between experimental (**a** and **b**) and FE analysis (**c** and **d**)

- The fracture position of the FE analysis tensile test is found very similar to fractured samples.
- The proper validation by calculating the percentage of error for the tensile test with an error is less than 1%.
- The observation of fractography reveals that voids and dimples are found in the fractured area of the tensile test sample and confirms ductile mode of failure.

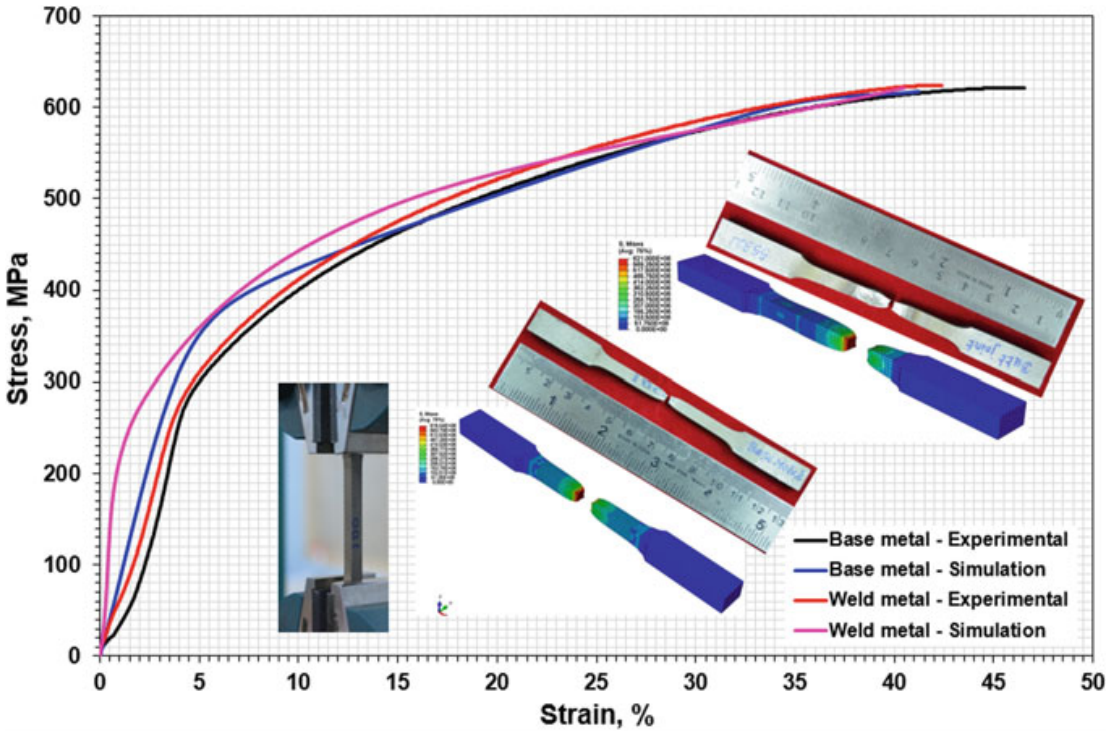


Fig. 16.11 Stress-versus-strain plot for BM and WM

Table 16.10 Uni-axial tensile test results

Sample type	Experimental tensile strength (MPa)	Simulation tensile strength (MPa)	% of error	Position of fracture
BM	621	617.8	0.51	Centre
WM	624	621.0	0.48	BM

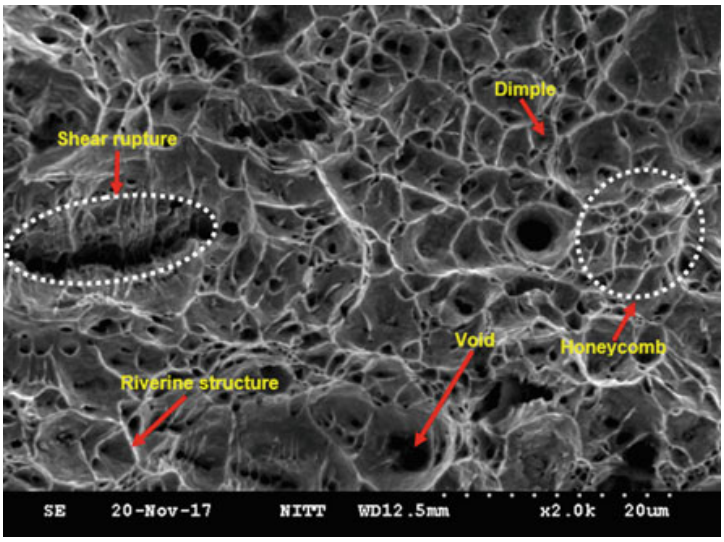


Fig. 16.12 SEM photomicrograph of fractured zone in BM tensile test sample

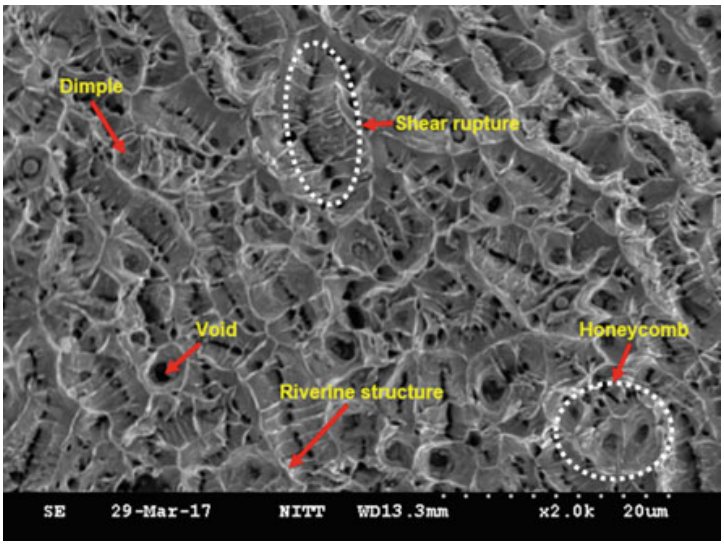


Fig. 16.13 SEM photomicrograph of fractured zone in WM tensile test sample

References

1. Nakhodchi, S., Shokuhfar, A., Iraj, S.A., Thomas, B.G.: Evolution of temperature distribution and microstructure in multipass welded AISI 321 stainless steel plates with different thicknesses. J. Press. Vessel Technol. **137**(6), 061405-1-15 (2015)
2. Fujii, H., Sato, T., Lu, S., Nogi, K.: Development of an advanced A-TIG (AA-TIG) welding method by control of Marangoni convection. Mater. Sci. Eng. A **495**, 296–303 (2008)
3. Vidyarthy, R.S., Dwivedi, D.K.: Activating flux tungsten inert gas welding for enhanced weld penetration: J. Manuf. Process. **22**, 211–228 (2016)
4. Joun, M., Choi, I., Eom, J., Lee, M.: Finite element analysis of tensile testing with emphasis on necking: Comput. Mater. Sci. **41**, 63–69 (2017)
5. Reisgen, U., Schleser, M., Mokrov, O., Ahmed, E.: Numerical and experimental investigation of tensile behavior of laser beam welded TRIP700 steel. ISIJ Int. **51**(3), 429–434 (2011)
6. Garcia-Garino, C., Gabaldon, F., Goicoleab, J.M.: Finite element simulation of the simple tension test in metals. Finite Elem. Anal. Des. **42**, 1187–1197 (2006)
7. Patel, A.B., Patel, S.P.: The effect of activating fluxes in TIG welding by using Anova for SS 321. Int. J. Eng. Res. Appl. **4**, 41–48 (2014)
8. Mohan Kumar, S., Siva Shanmugam, N.: Studies on the weldability, mechanical properties and microstructural characterization of activated flux TIG welding of AISI 321 austenitic stainless steel. Mater. Res. Express **5**, 106524, 1–48 (2018)
9. Sawakar, N., Boob, G.: Finite element based simulation of orthogonal cutting process to determine residual stress induced. Int. J. Comput. Appl. 0975–8887 (2014)
10. Liu, G.-H., Liu, M.-H., Yi, Y.-Y.: Activated flux tungsten inert gas welding of 8 mm-thick AISI 304 austenitic stainless steel. J. Cent. South Univ. **22**, 800–805 (2015)
11. Pouranvari, M., Asgari, H.R., Mosavizadch, S.M., Mariachi, P.H., Goodarzi, M.: Effect of weld nugget size on overload failure mode of resistance spot welds. Sci. Technol. Weld. Join. **2**, 217–225 (2007)

Chapter 17
Design Optimization of Welding Parameters for Multi-response Optimization in TIG Welding Using RSM-Based Grey Relational Analysis

S. C. Moi, R. Rudrapati, A. Bandyopadhyay and P. K. Pal

Abstract This paper presents a research work on optimization of input process parameters using response surface methodology (RSM) along with grey relational analysis to enhance the weldment quality responses in tungsten inert gas welding of 316L austenitic stainless steels. The process control parameters selected for the investigation are welding current, welding speed, and gas flow rate. The experiments have been performed using Box–Behnken design of RSM with three factors, three levels each. Experiments have been performed and output responses: ultimate tensile strength (UTS) and weld width (WW) have also been measured. Grey relational analysis technique has been applied on experimental data to convert multi-output problem into single-response problem. Mathematical model has been developed correlating the process parameters with grey relational grade. The effects of welding input parameters on multi-performance characteristics have been studied graphical contour plots which made from the developed mathematical model. Optimal process parametric has been obtained at higher grey relational grade value. The results of ANOVA and contour plots indicating that output responses, UTS and WW, are significantly influenced by welding input parameters. The concluding remarks have been drawn from the study.

Keywords TIG welding · 316L austenitic stainless steel · Response surface methodology · Grey relational analysis · Welding performance

S. C. Moi · A. Bandyopadhyay · P. K. Pal
Mechanical Engineering Department, Jadavpur University, Kolkata, India
e-mail: sc_moi1@rediffmail.com

A. Bandyopadhyay
e-mail: asishbanerjee@yahoo.com

P. K. Pal
e-mail: pradippal54@yahoo.com

R. Rudrapati (✉)
Mechanical Engineering Department, Hawassa University, Awasa, Ethiopia
e-mail: rameshrudrapati@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_17

17.1 Introduction

The selection of the material for any industrial application is based on its mechanical properties like strength, ductility resistance to oxidation and/corrosion, sensitization, etc. Variety of austenitic stainless steel materials which are used for many industrial purposes, type 316L is usually considered over other austenitic steel materials as a structural material, because, of its higher corrosion resistance and excellent mechanical properties both at low and high temperature [1]. Austenitic stainless steel (type 316L) is one of the most consumable materials in orthopedic implants [2], and it is considered as a major structural part for sodium cooled fast reactor applications [3], etc. TIG welding is very useful for similar or dissimilar welding of metals compared to laser welding, because, of high welding quality, high stability in welding of variety of industrial applications. Tungsten inert gas (TIG) welding with and without use of filler material is generally applied welding methods for joining of stainless steels for the industrial applications like pipe, automotive exhaust gas system, chemical industry-based equipment, power plants [3], etc. Welding of 316L austenitic stainless steel (ASS) is important area of research [4] due to corrosion and crack formation resistance [5].

In TIG welding, correct selection of welding process parametric setting is very important to decide the quality of the weld specimen. The quality levels of weldments is highly depending on many interactive input variables such as, the gap maintained between the joined parts, evaporation of parts, torch angle, quantity of the shielding gases present, pulse frequencies and the electrode tip angle selected, etc. when conduction of welding operation. It is very important to select the optimal welding process parametric combination to obtain required weld performances [6]. The systematic analysis and optimization approaches are required to produce desired quality welding and to enhance the quality responses of joining parts [7] but it required to study and understand the welding process extensively because many factors control the welding process [8].

The experimental analysis and optimization of any welding process are difficult job due to many interlinked parameters involved while solving it. The experimental analysis of welding is often a very costly as well as time-consuming process [9]. With a view to achieving the desired quality characteristics in TIG welding of 316L austenitic stainless steels, statistically designed experiments based on the Box–Behnken design of RSM have been used to conduct experimental study. The Box–Behnken experimental design-based RSM is very useful to study system or process effectively to obtain the required information about the direct and interaction effects of input process variables on performance measures with only fifteen number of experimental runs. The input parameters considered in the present work are current, speed, and gas flow rate. The welding performance characteristics measured after the experimentation are ultimate tensile strength (UTS) and weld width. Solving the multi-performance characteristics is not an easy task. Ghosh et al. [10], Wahule and Wasankar [11], and Bahar et al. [12] stated that grey relational analysis (GRA) is efficient for solving multi-objective problems. In the present research investigation,

GRA is utilized to solve the multi-responses by converting multi-output parameters into single response. RSM has been used to make mathematical equation between the input factors and grey relational grade. Two-dimensional contour plots have been made RSM to analyze the factor effects on output response(s).

17.1.1 Literature Survey

Recently, some papers reported in literature to establish a new experimental optimization approach using Taguchi method to analyze TIG welding with minimum number experimental runs.

Durgutlu [13] had investigated the microstructure, penetration, and mechanical properties of 316L ASS in TIG welding operation. They tested the effects of shielding gas (i.e., hydrogen in argon) while welding of 316L ASS. They found highest tensile strength of sample at 1.5% H_2-Ar from their study. Soltani and Tayebi [5] had conducted a comparative analysis of welding of 304L and 316L stainless steels using laser and TIG welding operations. They considered yield strength, elongation, hardness, and weld width as output responses. Authors stated from their study that TIG welding has produced better weld results compared to laser welding. Thomas et al. [4] had been investigated the fatigue crack growth rate in stainless steel (type 316LN) sheets joined by using traditional multi-pass TIG and activated TIG welding processes. Researchers had been found that better results while welding of multi-pass welding of 316LN stainless steel compared to activated TIG welding. Cheng et al. [14] studied the significances of TIG flux on the quality responses of dissimilar welds between mild steel G3131 and austenitic stainless steel 316L. The experimental runs were performed to investigate the effects of CaO, Fe_2O_3, Cr_2O_3, and SiO_2. They found residual slag on TIG weld surface due to oxide flux. Improvements in weld joint penetration and increased weld depth to width ratio were obtained by using SiO_2 powder. Reddy [15] conducted experiments to weld AISI 4140 and AISI 316 with and without using filler metal in TIG welding. The quality responses were evaluated from the study.

Ghosh et al. [10] made an investigation to visualize the welding factor effects on multi-performance characteristics of weld joint in metal inert gas (MIG) welding of ferritic stainless steel to austenitic stainless steel with the use of grey relational analysis and Taguchi method. Authors mentioned that welding responses were highly depend on the proper selection input parametric combination. Researchers had also been stated that grey relational analysis is efficient to deal with multi-responses problems. Wahule and Wasankar [11] had been carried out experimental research work to optimize the multi-outputs such as tensile strength (TS) and % elongation (PE) in TIG welding of SS304 and Fe-410 materials. They considered grey relational analysis technique to solve multi-objective problem (TS and PE) into single objective problem. Researchers used analysis of variance tool to study the significances of input parameters on both the output responses. They stated from their work that input welding parameters are significantly affecting the mechanical properties of welded

joint. They optimized multi-objectives: TS and PE simultaneously by using grey relational analysis. Bahar et al. [12] had conducted analysis on TIG welding of SS316 and MS1020 materials to enhance the welding performance. Authors selected current, welding speed, and gas flow rate as input parameters and hardness, percentage elongation, and UTS as output responses. GRA had been used to optimize the multi-output variables. Factor effects on all the responses were analyzed by analysis of variance and graphical main effect plots. Optimum input combination had been obtained by grey based Taguchi method. Authors had stated that GRA was very useful to optimize multi-objectives simultaneously.

From the literature review, one can identify that various aspects of TIG welding had been investigated by the researchers. These aspects include parametric studies related to enhance the welded joint quality characteristics as well. However, research is still being continued. This indicates the need for further research. The effects of weld process parameters on weld joint quality characteristics like UTS and weld width though studied by some investigators, conclusive relationship between process parameters, and quality characteristics—explaining the effects will only emerge through more extensive research.

17.2 Response Surface Methodology (RSM) and Grey Relational Analysis (GRA)

RSM is used to determine the relationships between output response(s) and a set of quantitative input parameters. RSM is very useful to predict the improved product quality over the quality achieved. In the present work, RSM's Box–Behnken experimental table with three factors, three levels each, and fifteen runs (Table 17.1) are chosen. The mathematical equation is then postulated that illustrates the relationship between the input process variables and response characteristic. The quantitative form of a mathematical model to postulate the relationships between input and output variables is given as follows:

$$Y = f(\text{current, speed, gas flow rate}) \tag{17.1}$$

where Y is the output response. Here, it is assumed that the independent control variables are continuous and controllable by experiments with negligible errors. The complete explanation of the process behavior can be obtained through second-order polynomial equation is given below.

$$\begin{aligned} Y = {} & \beta_0 + \beta_1(\text{current}) + \beta_2(\text{speed}) + \beta_3(\text{gas flow rate}) + \beta_{11}(\text{current} * \text{current}) \\ & + \beta_{22}(\text{speed} * \text{speed}) + \beta_{33}(\text{gas flow rate} * \text{gas flow rate}) + \beta_{12}(\text{current} * \text{speed}) \\ & + \beta_{13}(\text{current} * \text{gas flow rate}) + \beta_{23}(\text{speed} * \text{gas flow rate}) \end{aligned} \tag{17.2}$$

Table 17.1 Box–Behnken design, output responses, and grey relational grade and ordering

Run	Input parameters			Output responses		Grey relational analysis	
	Current (A)	Speed (cm/min)	GFR (l/min)	UTS (MPa)	Weld width (mm)	Grey relational grade	Grey relational order
1	100	12	9	612.13	7.61	0.418719	13
2	150	12	9	616.42	6.78	0.542807	11
3	100	18	9	608.48	6.22	0.711446	7
4	150	18	9	630.37	6.35	0.76222	5
5	100	15	6	621.26	6.81	0.561772	9
6	150	15	6	612.91	6.64	0.559544	10
7	100	15	12	580.61	7.37	0.354331	15
8	150	15	12	607.69	6.82	0.498924	12
9	125	12	6	624.15	6.83	0.575156	8
10	125	18	6	616.24	6.18	0.770095	4
11	125	12	12	592.42	7.44	0.371909	14
12	125	18	12	615.88	6.24	0.729702	6
13	125	15	9	639.42	6.55	0.782465	3
14	125	15	9	642.64	6.54	0.832558	1
15	125	15	9	639.67	6.56	0.782788	2

where Y = output response and the betas are coefficients of linear, quadratic, and interaction of welding process parameters. The term β_0 is the intercept term; β_1, β_2, and β_3 are the linear terms; β_{11}, β_{22}, and β_{33} are the squared terms; and β_{12}, β_{13}, and β_{23} are the interactions between the input factors. This empirical mathematical model is helpful to identify the optimum welding parametric setting to obtain desired weld performance.

GRA is mainly utilized to solve the multi-attribute decision-making problems. It considers multi-objectives simultaneously to identify the best alternative. Data preprocessing steps of GRA can be expressed as:

Step 1: In the GRA, first step is to preprocess the experimental data because the factor goals and directions may be different in each other. The preprocessing of the data is called grey relational generation. The equations for calculating higher-the-better and lower-the-better are given in Eqs. 17.3 and 17.4, respectively. Higher-the-better:

$$X_i(k) = \frac{Y_i(k) - \min Y_i(k)}{\max Y_i(k) - \min Y_i(k)} \tag{17.3}$$

where $X_i(k)$ is the value after grey relational generation while min $Y_i(k)$ and max $Y_i(k)$ are, respectively, the smallest and largest values of $Y_i(k)$ for the kth response. Lower-the-better:

$$x_i(k) = \frac{\max y_i(k) - y_i(k)}{\max y_i(k) - \min y_i(k)} \tag{17.4}$$

where $x_i(k)$ is grey relational generation; min $y_i(k)$ and max $y_i(k)$ are the smallest and largest values of $y_i(k)$ for the kth response.
Step 2: Grey relation coefficient
By normalizing, grey relational coefficient (GRC) is calculated by using Eq. 17.5 which is given below

$$\xi_i(k) = \frac{\Delta \min + r\Delta \max}{\Delta o_i k + r\Delta \max} \tag{17.5}$$

where $\Delta o_i = \| x_0(k) _ x_i(k) \|$ is variance of the absolute value between $x_0(k)$ and $x_i(k)$ and here "r" is distinguishing coefficient used to adjust the difference of the relational coefficients; Δ min and Δ max are minimum and maximum values of Δo_i for the kth response; "r" is within the set {0, 1}. $\xi_i(k)$ is grey relation coefficient.
Step 3: Determination of the grey relational grade (GRG) for each experiment combination

$$\alpha_i = \sum_{k=1}^{n} \xi_i(k)/n \tag{17.6}$$

where "n" is the number of performance characteristics; $\xi_i(k)$ = grey relation coefficient of each response; α_i is grey relational grade.
Step 4: Grey relational ordering: In GRA, final step is to provide the grades of GRG an ascending order to all factor combinations. Optimal parametric setting can be obtained at which parametric combination possess highest grade value.

17.3 Experimental Procedure

As already mentioned, that present work is planned to obtain the better weld qualities on 316L austenitic stainless steel material in TIG welding operation by using RSM and grey relational analysis. The selected dimensions of base material, i.e., 316L austenitic stainless steel are: 50 mm × 80 mm × 3 mm, and 1.6 mm diameter of ER316L austenitic stainless steel filler rod has been selected for the work. Juang and Tarng [16] and Ugur et al. [17] had mentioned that correct selection of process parameters had highest influence on the weld pool geometry as well as mechanical properties of welded joint in TIG welding operation. The important process parameters: welding current, gas flow rate, and welding speed have been selected for the present experimentation investigation. The selected input parametric settings and their levels are current = 100, 125, and 150 A; speed = 12, 15, and 18 cm/min; and gas flow rate = 6, 9, and 12 l/min. The experiments have been performed as per

Fig. 17.1 Experimental setup

Box–Behnken design of response surface methodology. Figure 17.1 shows the photographic view of a semi-automatic TIG welding machine (Model-SUPERGEN320, M/C SR. No-12111097670 and Make-Ador Welding Ltd.). Box–Behnken design matrix has been shown in Table 17.1. After TIG welding operation, visual and X-ray radiographic tests have been conducted on welded specimen to check the defects in it. However, visual and X-ray radiographic test results have not been discussed in the present paper. Output responses, ultimate tensile strength (UTS) and weld width (WW), have been measured and given in Table 17.1.

17.4 Converting Multi-response into Single-Response Problem

Grey relational analysis has been conducted on experimental data to translate multi-responses into single objective problem by following the steps mentioned above. Present work, UTS and WW, must be maximized to enhance the mechanical properties and WW needs to be minimized to reduce the material requirement. As mentioned above, in GRA, first step is to preprocess the data. Here, Eqs. 17.3 and 17.4 have been used to preprocess the data of UTS and WW, respectively. Equation 17.5 has been used to calculate the grey relational coefficient. Finally, Eq. 17.6 has been used to determine the grey relational grade (GRG) by averaging the grey relational coefficients (GRC) of both the responses, UTS and WW. Grey relational ordering has been made to grey relational grade values. The GRG values and ordering are given in Table 17.1.

17.5 Results and Analysis

The experiments have been made and output responses are measured and shown in Table 17.1. GRA is used to translate multi-objectives, i.e., UTS and WW into single response, grey relational grade (Table 17.1). RSM is applied on grey relational grade to postulate the relationships between welding factors: current, speed and gas flow rate, and grey relational grade. Contour plots have been made from developed mathematical equation by using RSM application. Optimum welding condition obtained highest grey relational grade.

17.5.1 Mathematical Modeling and Contour Plots

Developing mathematical model from the experimental data can create a more scope to study and analyze the experimental results. In the present investigation, RSM has been utilized to develop the mathematical relationships between the welding process parameters and grey relational grade (GRG) as given in Table 17.1. Here, GRG is developed from the output responses: UTS and PE as already mentioned above. RSM is one of the empirical mathematical modeling tools used to carry out to develop mathematical relationships between input and output parameters. General polynomial quadratic model of is already given in Eq. 17.2. The constant values of β_0, β_1, β_2, β_3, β_{11}, β_{22}, β_{33}, β_{12}, $\beta_{13,}$ and β_{23} as given in Eq. 17.3 are calculated from experimental data (i.e., GRG in the present case) as shown in Table 17.1 using least square. The developed models are shown in Eq. 17.7.

$$\begin{aligned} Y = & -6.33747 + 0.0486566 * \text{current} + 0.404600 * \text{speed} \\ & + 0.212709 * \text{gas flow rate} - 1.98029\text{E} - 04 * \text{current} * \text{current} \\ & - 0.0127831 * \text{speed} * \text{speed} - 0.0103304 * \text{gas flow rate} * \text{gas flow rate} \\ & + 2.32751\text{E} - 05 * \text{current} * \text{speed} + 5.16623\text{E} - 05 * \text{current} * \text{gas flow rate} \\ & - 0.00207718 * \text{speed} * \text{gas flow rate} \end{aligned} \tag{17.7}$$

where Y is output response.

The contour plots are two-dimensional plots which are very useful to study the main and interaction effects of two factors on outcome in a single plot [18]. In the present work, contour plots have been drawn from the developed mathematical model (Eq. 17.7) of grey relational grade (GRG) by RSM application from the MINITAB 16.1 software. The relationship between GRG with the welding input parameters has been demonstrated graphically by these plots. Each line in the contour plots is a constant—response line. The contour lines in a plot, comparatively straighter lines, indicate lesser or no interaction effect on the output variable.

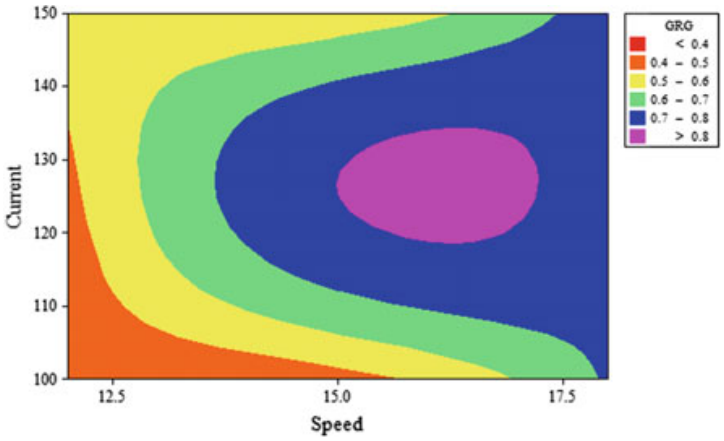


Fig. 17.2 Contour plot showing combined effects of current and speed on GRG

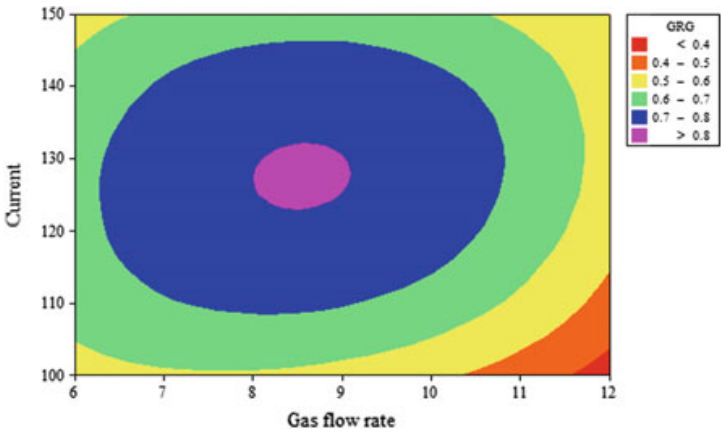


Fig. 17.3 Contour plot showing combined effects of current and gas flow rate GRG

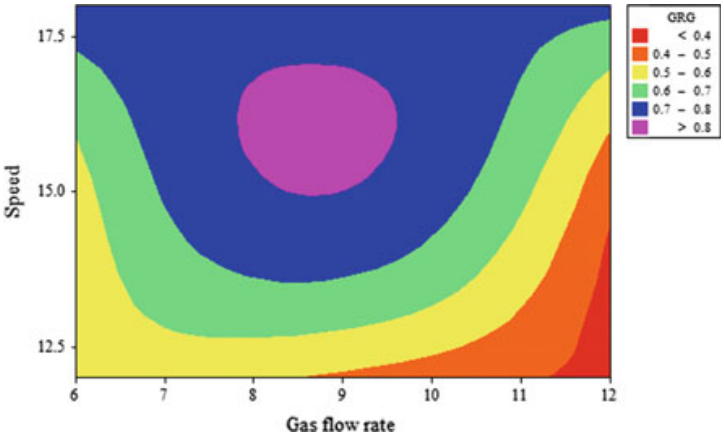


Fig. 17.4 Contour plot showing combined effects of speed and gas flow rate on GRG

The elliptical nature or bent/circular contours in a plot shows that input parameters are significant on the response. The contour plots for GRG is shown in Figs. 17.2, 17.3, and 17.4.

From contour plots of grey relational grade (GRG) as given in Figs. 17.2, 17.3, and 17.4, it is noted that all the welding input variables are most significant for UTS and weld width. From the contour plots, one can identify the optimum parametric setting where better outcome obtained. In a contour plot, each shade has set of two

input parameters which projecting certain output values. From the all contour plots shown in Figs. 17.2, 17.3 and 17.4, it is observed that input parameters current = 125 A, speed = 15 cm/min and gas flow rate = 9 l/min, is one of optimum parametric setting which produces higher grey relational grade. Same parametric setting has also been suggested by grey relational analysis (Table 17.1) for predicting multi-response characteristics combinedly.

17.6 Conclusions

The experiments have been conducted to study, model, and optimize the weld responses: ultimate tensile strength (UTS) and weld width (WW) combinedly in TIG welding on 316L austenitic stainless steel plates. Following are the conclusions drawn from the present research investigation:

1. Multi-objective problem (UTS and WW) is converted into single-response problem (i.e., grey relational grade) by GRA approach
2. Mathematical relationships between welding parameters and grey relational grade by response surface methodology (RSM)
3. Contour plots are drawn using RSM application
4. The effects of welding parameters on both the responses are studied from contour plots
5. Input welding parameters are most significant for both the responses
6. Optimum welding condition: current = 125 A, speed = 15 cm/min, and gas flow rate = 9 l/min is obtained by GRA
7. RSM and GRA are very useful to optimize the welding conditions of TIG welding 316L austenitic stainless steel.

References

1. Samanta, S.K., Mitra, S.K., Pal, T.K.: Effect of rare earth elements on microstructure and oxidation behavior in TIG weldments of AISI 316L stainless steel. Mater. Sci. Eng., A **430**, 242–247 (2006)
2. Dadfar, M., Fathi, M.H., Karimzadeh, F., Dadfar, M.R., Saatchi, A.: Effect of TIG welding on corrosion behavior of 316L stainless steel. Mater. Lett. **61**, 2343–2346 (2007)
3. Sakthivel, T., Vasudevan, M., Laha, K., Parameswaran, P., Chandravathi, K.S., Mathew, M. D., Bhaduri, A.K.: Comparison of creep rupture behaviour of type 316L(N) austenitic stainless-steel joints welded by TIG and activated TIG welding processes. Mater. Sci. Eng., A **528**, 6971–6980 (2011)
4. Thomas, M., Raghu, V., Prakash, S., Ganesh, S.R., Vasudevan, M.: High temperature fatigue crack growth rate studies in stainless steel 316L(N) welds processed by A-TIG and MP-TIG welding. In: 12th International Fatigue Congress (2018)

5. Soltani, H.M., Tayebi, M.: Comparative study of AISI 304L to AISI 316L stainless steels joints by TIG and Nd:YAG laser welding. J. Alloys Compd. (2018). https://doi.org/10.1016/j.jallcom.2018.06.302
6. Konkol, P.J., Koons, G.F.: Optimization of parameters for two wire AC - ACSAW. Weld. J. **27**(12), 367s–374s (1978)
7. Tarng, Y.S., Tsai, H.L., Yeh, S.S.: Modeling, optimization and classification of weld quality in tungsten inert gas welding. Int. J. Mach. Tools Manuf. **39**, 1427–1438 (1999)
8. Sudhakaran, R., VelMurugan, V., Sivasakthivel, P.S.: Effect of process parameters on depth of penetration in gas tungsten arc welded (GTAW) 202 grade stainless steel plates using response surface methodology. TJER **9**(1), 64–79 (2012)
9. Correia, D.S., Gonçalves, C.V., da Cunha Jr., S.S., Ferraresi, V.A.: Comparison between genetic algorithms and response surface methodology in GMAW welding optimization. J. Mater. Process. Technol. **160**, 70–76 (2005)
10. Ghosh, N., Pal, P.K., Goutam, N.: Investigation on dissimilar welding of AISI 409 ferritic stainless steel to AISI 316L austenitic stainless steel by using grey based Taguchi method. Adv. Mater. Process. Technol. (2018). 2374068X.2018.1451182
11. Wahule, A., Wasankar, K.: Multi-response optimization of process parameters of TIG welding for dissimilar metals (SS-304 and Fe-410) using grey relational analysis. Int. Res. J. Eng. Technol. **5**(6), 986–993 (2018)
12. Bahar, D., Venkatesh, A., Vinay, B.K., Siva, T.R.K.P.: Multi objective optimization of TIG welding parameters for joining dissimilar materials. Int. J. Eng. Sci. Res. **8**, 26–33 (2018)
13. Durgutlu, A.: Experimental investigation of the effect of hydrogen in argon as a shielding gas on TIG welding of austenitic stainless steel. Mater. Des. **25**, 19–23 (2004)
14. Cheng, H.K., Kuang, H.T., Chang, P.C.: Effect of activated TIG flux on performance of dissimilar welds between mild steel and stainless steel. Key Eng. Mater. **479**, 74–80 (2011)
15. Reddy, M.P.: Assessment of mechanical properties of AISI 4140 and AISI 316 dissimilar weldments. In: 7th International Conference on Materials for Advance Technology. Elsevier, Amsterdam (2014)
16. Juang, S.C., Tarng, Y.S.: Process parameter selection for optimizing the weld pool geometry in the tungsten inert gas welding of stainless steel. J. Mater. Process. Technol. **122**, 33–37 (2002)
17. Ugur, E., Melih, B, Yugut, K., Sueda, O.: Optimization of weld bead geometry in TIG welding process using grey relation analysis and Taguchi method. MTAEC9 **43**(3), 143–149 (2009)
18. Moi, S.C., Rudrapati, R., Bandyopadhyay, A., Pal. P.K.: Determination of tungsten inert gas welding input parameters to attain maximum tensile strength of 316 l austenitic stainless steel. J. Mech. Eng. Strojnicky Casopis **68**, 231–248 (2018)

Chapter 18
Subjective Factors Consideration in the Selection of Welding Technique for Welded Tube Manufacturing

Ravindra S. Saluja and Varinder Singh

Abstract In choosing the right welding process for an application like welded tube manufacturing, the manufacturing engineers have to take into account a large number of selection criteria. Choosing the right welding process for any application is difficult not only due to multiple factors affecting the decision but also due to a number of emerging new welding techniques becoming available as alternatives. Some of the alternative welding processes for manufacturing tubes have been Fretz-Moon process (FM), low-frequency (LF) as well as high-frequency (HF) alternating current resistance welding, direct current (DC)-based resistance welding in addition to gas metal arc welding (GMAW), and submerged arc welding (SAW) in some specific applications. A systematic examination of the decision situation may lead us to an appropriate choice rather than basing the decision on one factor or the other in stand-alone manner. This paper presents an analytic hierarchy process (AHP)- based decision analysis for the selection of welding process in a case of manufacturing small diameter steel tubes.

Keywords Steel tubes · Welding process · Decision analysis · AHP

18.1 Introduction

Tube manufacturing is one of the major industries supporting almost all other sectors dealing with engineering products. Most of the tubes are manufactured by a technique involving roll bending a sheet or strip and welding the ends in a continuous operation to give it a tube shape. This type of manufacturing technique is preferred one in comparison to seamless tube manufacturing as it offers significant

R. S. Saluja (✉) · V. Singh
Department of Mechanical Engineering, K. K. Birla Goa Campus, BITS Pilani, Sancoale 403726, Goa, India
e-mail: rsaluja@goa.bits-pilani.ac.in

V. Singh
e-mail: vsingh@goa.bits-pilani.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_18

advantages on economics and production front. Tubes manufactured by this technique are commonly called welded tubes and are able to satisfy general quality requirements. By subjecting such products to post-processing operations such as cold drawing and some other operations, even stringent quality requirements are met easily. The choice of welding technique plays a major role in defining the cost and quality of such tubes. In general, the techniques of welding deployed for manufacturing different tubes are decided based on stand-alone technical considerations. In the past, many researchers [1–4] have pointed out the inadequacies of such approach while deciding welding techniques for different applications. The application of multiple criteria decision tools for systematic consideration of varied factors has been a topic of active research in the recent past wherein such investigations have resulted in well thought out selection processes for applications as trivial as butt joints for aluminum parts to hard facing of carbon steels by welding harder material.

For commercially important tube manufacturing industry, no systematic treatment is yet reported in literature for consideration of multiple alternatives available and the multiple attributes affecting the decision. An appropriate choice of welding for a mass scale industry of tube manufacturing may make a significant difference in the competitiveness of the industry in terms of cost and quality in addition to the some benefits even for general society in terms of resource sustainability. In the modern times, a vast amount of technological research is resulting in the advent of newer techniques in welding technology. The present research may also pave way for consideration of newer alternatives which may result in fruitful and commercially exploitable innovations. Thus, a systematic decision-making process is very much needed that can take the multiple attributes as well as multiple alternatives into consideration for selection of the most appropriate welding technique for welded tube manufacturing. Different multi-attribute decision-making (MADM) tools may be used for analyzing this decision. Analytic hierarchy process (AHP) is one of the popular ones used by a number of researchers for decision making in different applications [5]. Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) is another multi-attribute method that has been used to deal with a number of multi-attribute decision problems [1, 6–8]. A hybrid AHP-TOPSIS approach has also been in use to provide decision support in many applications. Similarly, there are many more MADM tools, all with the aim of resolving the multi-attribute decision situation in a systematic manner while following different algorithms. On the other hand, the fuzzy scales have also been adopted by a number of researchers within AHP, TOPSIS, and different other MADM methodologies [9–11]. Researchers [11] have also discussed that the use of different MADM tools do not create significant difference in results and the choice of the tool may be based on minimizing the computational load. AHP in its basic form also holds significant promise and is widely used to address decision situations involving subjective judgments in pairwise comparison of attributes as well as alternatives in the domain of the chosen attributes of decision making.

The present paper investigates the choice of welding technique for small diameter tube manufacturing considering multiple subjective attributes. AHP

method is employed as it offers a simple way to systematically arrive at the user preference to meet the goal of selection process.

18.2 Methodology and Implementation

The overall methodology used in this paper for investigating decision making in welding process selection for small diameter tube manufacturing while considering subjective factor is graphically presented in Fig. 18.1.

18.2.1 AHP Model

For calculating the weights and user preference scores based on the subjective factors, AHP method is used. The general steps in AHP are described below which are based on established literature sources [3–5].

(a) Index the welding processes that engineers can take into consideration for potential pick.
(b) Find the attributes with measurable values that could be used in decision making.

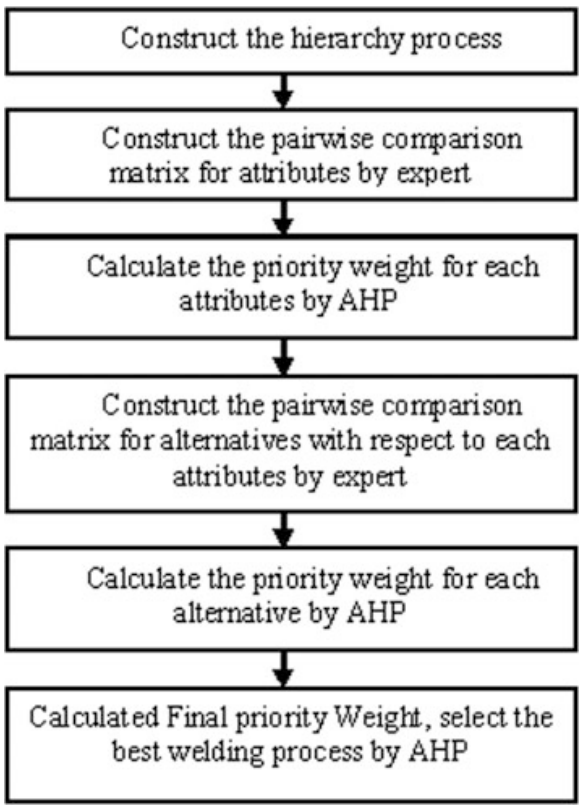


Fig. 18.1 Welding process selection based on the subjective factor by using an AHP methodology for tube manufacturing

(c) Next, the pairwise comparison matrix of set of attributes is completed by capturing relative importance as per expert views.
(d) Once the pairwise comparison has been captured in a matrix for attributes, the normalized values are estimated. The values of each column are normalized by dividing each value in column by the sum of the values in the same column. Then, the average of each row is calculated to provide priorities weights (PW) of the criteria.
(e) Similar to steps (c and d), the pairwise comparisons of all the alternatives is used to obtain priority weights (PW) of each alternative with respect to different attributes.
(f) Also, before weights are calculated in steps (c, d, and e), it is necessary to check the pairwise comparisons are consistent. For that purpose, a measure known as consistency ratio (CR) is obtained as below.

$$\mathrm{CR} = \mathrm{CI}/\mathrm{RI}, \tag{18.1}$$

where consistency index,

$$(\mathrm{CI}) = (\lambda_{\max} - n)/(n-1) \tag{18.2}$$

$\lambda_{\max}$ = Maximum eigen value
n = Dimension of matrix.
And RI is a random index based on a table in reference [5].
So as long as CR < 0.1, the evaluations arrived at are considered consistent.
(g) For AHP-based selection, the weighted normalized decision matrix is calculated by summations of the products of attribute weights and alternative weights to obtain the final rankings.
(h) Compute the weighted priority scores for each alternative.
(i) In the end, the ranks are assigned to alternatives having the highest weighted priority score.

18.2.2 Implementation of Decision-Making Model

The case of manufacturing welded tube with the following design specifications has been considered for this study.

Material: Low carbon steel.
Diameter: 15 mm
Thickness: 2 mm

Based on the design specifications and the guidelines of the Society of Manufacturing Engineering [12, 13], four welding processes have been shortlisted for this case for further study namely Fretz-Moon method (FM), direct current

electric resistance welding (DC), low-frequency electric resistance welding (LF), and high-frequency electric resistance welding (HF). Many research papers [1–4] dealing with selection of welding processes for different applications use many different attributes.

The hierarchical structure of the decision problem is presented in Fig. 18.2 where the top level indicates the overall goal, which is to bring the best welding process available for tube manufacturing on mass scale. As the attributes important for the current decision making, namely setup time requirement (SR), worker skill requirement (WSR), occupational factor (OF), ease of automation (EA), and quality (QY), are explained in Table 18.1, at the second level, such attributes and their relative importance toward achievement of the goal of selecting the best alternative are analyzed through pairwise comparison. At the next level, the four potential alternatives, namely Fretz-Moon process (FM), direct current electric resistance welding (DC), low-frequency electric resistance welding (LF), and high-frequency electric resistance welding (HF), are considered for pairwise comparison in the direct of each attribute.

For pairwise comparison of different attributes as well as alternatives, the scale of comparison is presented in Table 18.2. It is used to fill the pairwise comparison elements for each attribute in Table 18.3. It is also used to fill pairwise comparison for each alternative with respect to all considered attribute in Table 18.4.

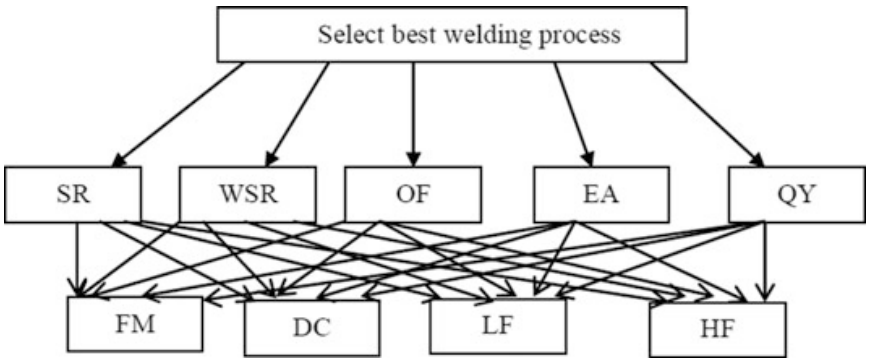


Fig. 18.2 Schematic diagram of analytic hierarchy process

Table 18.1 Description of attributes considered [3, 4]

S. No.	Subjective attributes	Description
1	Setup time requirement (SR)	Welding parameters setting, cleaning the base metal, and clamping joint in fixtures
2	Worker skill requirement (WSR)	Refer to the skill level of worker required
3	Occupational factor (OF)	Arc glare, smoke, and fumes
4	Ease of automation (EA)	Manual, semi-automatic, and fully automatic
5	Quality (QY)	Weld bead appearance and surface finish

Table 18.2 Strength of preference for pairwise comparison [3–5]

Strength of importance	Verbal judgement
1	Equally preferable
3	Moderately preferable
5	Strongly preferable
7	Very strongly preferable
9	Extremely strongly preferable
2, 4, 6, 8	Intermediate values
1/2, 1/3, 1/4 etc.	Reciprocal of above values

18.3 Results and Discussion

The pairwise comparison of attributes for this case of tube manufacturing is presented in Table 18.3. As evident from the priority weight scores in Table 18.3, the quality of the welding process is considered as the most important attributes followed by occupational factor and so on.

Similar way, the pairwise comparison of welding processes with regard to each of the considered attributes and values of priority weights (PW) is presented in Table 18.4. It may be noted that the calculated consistency ratio by using Eq. (18.1) in Tables 18.3 and 18.4 is found to be less than 0.1 as per the requirement of AHP procedures.

The results of Table 18.4 are summarized in Table 18.5 and overall composite priority weights of alternatives are calculated according to step (f) and step (g) as mentioned in Sect. 18.2.1 and are reported in Table 18.6. From the overall priority weights for different alternatives as available in Table 18.6, the highest value for direct current (DC) resistance welding is observed.

18.3.1 Discussion

It may be observed that DC resistance welding is standing high in the order of preference. The reason for this outcome is that the quality of weld in terms of clean

Table 18.3 Comparison of subjective attributes

Attributes	SR	WSR	OF	EA	QY	Priority weight
SR	1.000	0.111	0.143	0.200	0.111	0.030
WSR	9.000	1.000	0.333	0.500	0.333	0.138
OF	7.000	3.000	1.000	3.000	0.333	0.254
EA	5.000	2.000	0.333	1.000	0.250	0.138
QY	9.000	3.000	3.000	4.000	1.000	0.439
CR = 0.094						

Table 18.4 Pairwise comparison matrix of processes with respect to the attributes

Attributes	Process	FM	DC	LF	HF	Priority weights	Consistency ratio
Setup time requirement (SR)	FM	1.00	0.20	0.20	0.14	0.05	0.02
	DC	5.00	1.00	1.00	0.33	0.21	
	LF	5.00	1.00	1.00	0.33	0.21	
	HF	7.00	3.00	3.00	1.00	0.52	
Worker skill requirement (WSR)	FM	1.00	3.00	3.00	3.00	0.50	0
	DC	0.33	1.00	1.00	1.00	0.17	
	LF	0.33	1.00	1.00	1.00	0.17	
	HF	0.33	1.00	1.00	1.00	0.17	
Occupational factor (OF)	FM	1.00	0.20	0.20	0.14	0.05	0.024
	DC	5.00	1.00	1.00	0.33	0.21	
	LF	5.00	1.00	1.00	0.33	0.21	
	HF	7.00	3.00	3.00	1.00	0.52	
Ease of automation (EA)	FM	1.00	0.20	0.20	0.20	0.06	0
	DC	5.00	1.00	1.00	1.00	0.31	
	LF	5.00	1.00	1.00	1.00	0.31	
	HF	5.00	1.00	1.00	1.00	0.31	
Quality (QY)	FM	1.00	0.14	0.50	0.25	0.06	0.08
	DC	7.00	1.00	5.00	3.00	0.59	
	LF	2.00	0.20	1.00	0.33	0.11	
	HF	4.00	0.33	3.00	1.00	0.23	

Table 18.5 Summary of priority weights of alternative welding processes

Process	SR	WSR	OF	EA	QY
FM	0.05	0.50	0.05	0.06	0.06
DC	0.21	0.17	0.21	0.31	0.59
LF	0.21	0.17	0.21	0.31	0.11
HF	0.52	0.17	0.52	0.31	0.23

Table 18.6 Ranking of welding processes by AHP model

Process	SR	WSR	OF	EA	QY	Composite priority weight	Ranking
FM	0.002	0.069	0.013	0.008	0.028	0.12	4
DC	0.006	0.023	0.054	0.043	0.260	0.39	1
LF	0.006	0.023	0.054	0.043	0.048	0.17	3
HF	0.016	0.023	0.133	0.043	0.103	0.32	2

weld is of highest importance in the present case of small diameter tube manufacturing on mass scale, as the post weld flash removal is extremely difficult. The DC resistance welding is otherwise a poorer choice in comparison to a more popular alternative of high-frequency alternating current resistance welding as is evident from Table 18.4. For medium diameter tubes, the cleaner quality weld may acquire lesser importance and the choice may tilt toward high-frequency alternating current resistance welding as flash removal is easier owing to larger diameter. The utility of such decision analysis is that it considers different practical aspects and uses the provided information systematically, which help the design engineer and manufacturing engineer to choose the most appropriate process considering several attributes simultaneously.

18.4 Conclusions and Future Scope

In this work, the AHP model has been used for investigating the issue of selecting appropriate welding process for a case of small diameter tube manufacturing on mass scale while considering multiple attributes, namely setup time requirement, worker skill requirement, occupational factor, ease of automation, and quality of weld. The conclusions drawn from the study and the future scope are presented below:

i. For the case of manufacturing a welded tube with the following design specifications: material—carbon steel, diameter—15 mm and thickness—2 mm, the direct current (DC)-based resistance welding is found to be the most appropriate process while considering multiple attributes.
ii. AHP-based analysis is found to be simple for adaptation by experts while systematically collecting expert opinion on different factors in the otherwise complex problem of selection involving multiple attribute and choices as well.
iii. In future, there is scope to consider emerging welding processes for multiple attribute-based analysis before actual commercial deployment. Also, there is scope to consider multiple MADM methods as well as fuzzy approach due to subjective judgments involved.

References

1. Mostafa, J., Vahdat, S.E.: A fuzzy multi-attribute approach to select the welding process at high pressure vessel manufacturing. J. Manuf. Process. **14**, 250–256 (2012)
2. Darwish, S.M., Tamimi, A.A., Habdan, S.A.: A knowledge base for metal welding process selection. Int. J. Mach. Tools Manuf. **3**, 1007–1023 (1997)
3. Ravisankar, V., Balasubramanian, V., Muralidharan, C.: Selection of welding process to fabricate butt joints of high strength aluminum alloys using analytic hierarchic process. Mater. Des. **27**, 373–380 (2006)

4. Balasubramanian, V., Varahamoorthy, R., Ramachandran, C.S., Muralidharan, C.: Selection of welding process for hard facing on carbon steels based on quantitative and qualitative factors. Int. J. Adv. Manuf. Technol. **40**, 887–897 (2009)
5. Saaty, T.L.: Analytic Hierarchy Process. McGraw Hill Publications, New York (1980)
6. Singh, V., Agrawal, V.P., Deb, P.: A decision making method for selection of finish process. In: Proceedings of IEEE International Conference on Industrial Engineering and Engineering Management, pp. 38–42, Macao, China, 7th–10th December (IEEM-2010). https://doi.org/10.1109/ieem.2010.5674417
7. Rao, R.V.: Machinability evaluation of work materials using a combined multiple attributes decision-making method. Int. J. Adv. Manuf. Technol. **28**, 221–227 (2006)
8. Rao, R.V., Davim, J.P.: A decision-making framework model for material selection using a combined multiple attributes decision-making method. Int. J. Adv. Manuf. Technol. **35**, 751–760 (2008)
9. Chang, D.Y.: Applications of the extent analysis method on fuzzy AHP. Eur. J. Oper. Res. **95** (3), 649–655 (1996)
10. Deng, H.: Multi-criteria analysis with fuzzy pair-wise comparison. Int. J. Approximate Reasoning **21**(3), 215–231 (1999)
11. Sabzi, H.Z., King, J.P., Gard, C.C., Abudu, S.: Statistical and analytical comparison of multi-criteria decision-making techniques under fuzzy environment. Oper. Res. Perspect. **3**, 92–117 (2016)
12. Baheti, M.: Everything About Welded Tube Technology. DeeTee Industries Ltd., Indore, India (2016)
13. Veilleux, R.F., Petro L.W.: Tool and manufacturing engineer's handbook. In: Manufacturing Management, vol. 5, 4th edn. Society of Manufacturing Engineers (1998). ISBN 978-0-872633-06-3

Chapter 19
GMAW Investigation of AISI 201 Stainless Steel and Industry Need Optimization Using Genetic Algorithm

Krity Lata, M. Chandrasekaran, S. K. Tamang, R. Ramesh and N. K. Rana

Abstract In this work, the welding investigation of AISI 201-grade stainless steel (120 mm × 60 mm with 4 mm thickness) using gas metal arc welding (GMAW) has been carried out. Four weld parameters, viz. wire feed rate, welding voltage, nozzle-to-plate distance and welding speed, are used to investigate different weld bead characteristics [i.e. penetration (*P*), bead width (*W*) and bead height (*H*)]. The welding experiments are performed using Taguchi L9 experimental design, and each run is completed in a single pass. The predictive models are developed to predict weld bead geometry, and the performance of the model is validated. The increase in wire feed rate and voltage increases penetration and bead width, while an increase in nozzle-to-plate distance decreases the value of penetration and bead width. The welding parameters are optimized using genetic algorithm to determine the optimal combinations of welding parameters for better quality of component having maximum *P* with minimum *W* and *H* at which the maximum weld quality be achieved.

Keywords AISI 201-grade stainless steel · GMAW · Weld bead · Genetic algorithm

19.1 Introduction

Nitrogen-strengthened austenitic stainless steel (200 series) has found more applications in modern industries compared to conventional AISI 300 series of stainless steel components [1]. The price of nickel is relatively higher compared to other austenite formation elements such as manganese, and changing so, it becomes

K. Lata · M. Chandrasekaran (✉) · S. K. Tamang · N. K. Rana
Department of Mechanical Engineering, NERIST (DU), Nirjuli 791109, Arunachal Pradesh, India
e-mail: mchse1@yahoo.com

R. Ramesh
PSG College of Technology, Coimbatore, Tamil Nadu, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_19

necessary to increase such a material like AISI 201, as an option, with the same weight of chromium 18%, and instead of nickel, 7 wt% of manganese is added to stabilize the austenitic phase, and weight of nickel is decreased from 8 to 3 wt% in stainless steel of AISI 201 grade [2]. Hence, AISI 201 SS is a high Cr–Mn–N stainless steel which is free of nickel. This material is useful in manufacturing of battle tanks in defence applications possessing low cost, better corrosion resistance and excellent mechanical properties. This material can be used for conventional construction applications such as nuclear reactors, thermal power plants, civil construction, vessels and heat exchangers and for multiple industrial applications such as food and dairy processing equipment and kitchen equipment for consumer use.

Gas metal arc welding (GMAW) is a renowned soldering process used for various ferrous and non-ferrous materials. This is a cheaper and cost-effective process, without flaws, without faults, with proper appearance and with easy mechanizing [3]. In this process of welding, an arc is struck between a continuous consumable wire electrode and the workpiece metal. To provide a good-quality weld having economic production of desired weld bead geometry, optimum selection of welding parameters is important. A number of researches were conducted on welding investigation and parametric optimization in different welding processes on joining various materials. Choudhury and Chandrasekaran [4] reviewed previous research works published in the areas of welding investigation and optimization on aerospace materials. In gas tungsten arc welding (GTAW), Juang and Tarng [5] studied the selection of optimal process parameters to obtain optimum welding bead geometry. The effects of hardness and penetration depth were investigated by Ibrahim et al. [6] on various MIG welding process parameters. The optimum parameter combination was assessed by Datta et al. [7], using the Taguchi approach and the Grey relational analysis, in order to reach an optimum penetration depth and minimum bead height, bead height and depth of the welds. They concluded that Taguchi–Grey method adopted achieves an improved weld quality. The effect of weld speed on microstructure, mechanical characteristics and corrosion studies of AISI 201 GTA welded stainless steel sheets was studied by Chuaiphan and Srijaroenpramong [2]. Kim et al. [8] studied gas metal arc welding process parameters modelling and optimizing with GA and RSM. Sathiya et al. [9] have examined laser welding processes of the AISI 904L super austenitic stainless steel materials' assemblies and process parameters for optimization with the help of soft computing methods such as ANN and GA. The author has attempted to investigate the effect of various weld parameters on weld bead characteristics during the AISI 201 SS GMAW process on beads on plates in the present investigation. The experimental results are analysed and weld parameters are optimized using GA to get better penetration with minimum bead width and bead height. The result provides important guidelines for the selection of weld parameters for particular application in fabrication industries.

19.2 GMAW Experimentation Procedure

The experiments designed were based on Taguchi L9 OA having 4 factors and 3 levels. The welding parameters were chosen as wire feed rate (*W*), welding voltage (*V*), nozzle-to-plate distance (*N*) and welding speed (WS), and response parameters were chosen as penetration (*P*), bead width (*W*) and bead height (*H*). Experiments were done on bead on plate (AISI 201 stainless steel of size 120 mm × 60 mm, 4 mm thick) as per design matrix using TransPuls Synergic 3200 MIG welding machine. The experimental set-up was prepared as shown in Fig. 19.1. GMAW machine was connected to shielding gas (98% Ar + 2% O_2) cylinder, and gas flow rate was controlled electronically using this machine. AISI 316L wire rod having diameter 1.2 mm was selected as welding electrode and was attached to the welding torch which was further connected to the set-up table and fixed at a torch angle of 90°. DCEP electrical polarity was used as it concentrated more heat on melting the filler material in the welding area. The variations in welding voltage automatically changed the current in this experimental set-up. Before welding, the workpiece was cleaned, degreased and clamped in the experimental set-up. After the complete configuration, the tests were performed randomly using the design matrix so that any systemic error was avoided. By conducting test runs, the working range of all the selected factors was determined by changing one factor while the remaining factor was kept as constant values. Thus, when inspecting the weld bead for a smooth appearance, no visible defects such as undercut and surface porosity were obtained in various test experiments and given in Table 19.1, and the range of parameters of process variables was defined.

Figure 19.2 shows welded specimens. After welding, the specimen was cut and grinded in a grinder disc and polished using nylon polishing cloth with levigated polishing alumina solution. The specimen was etched with etchant solution (15% HCL + 10% HNO_3 + 2–3 drop glycerine) and then washed with alcohol solution.

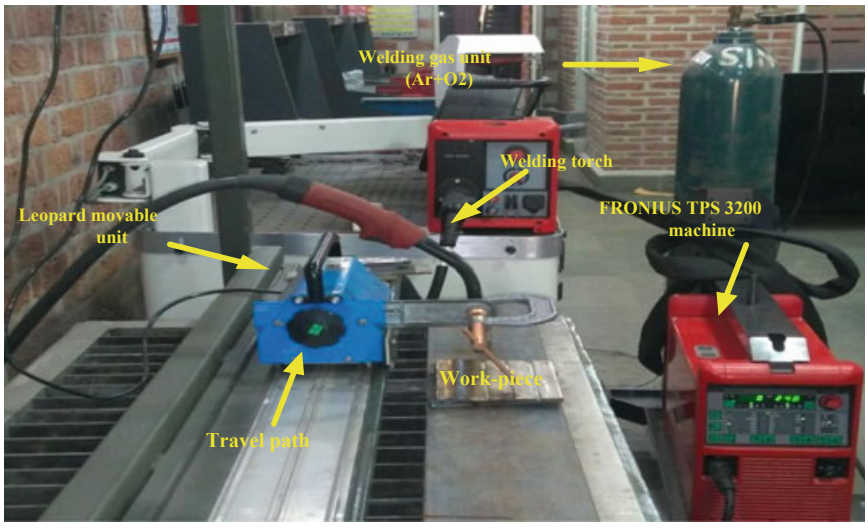
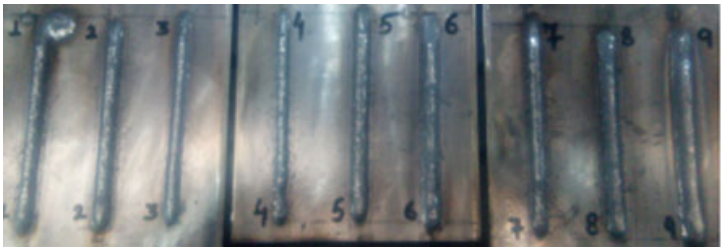


Fig. 19.1 GMAW experimental set-up

Table 19.1 GMAW parameters and their levels

Parameter	Unit	Symbol	Level 1	Level 2	Level 3
Wire feed rate	m/min	WFR	6	7	8
Welding voltage	V	*V*	20	22	24
Nozzle-to-plate distance	mm	*N*	16	20	24
Welding speed	mm/min	WS	220	260	300

Fig. 19.2 Weld runs

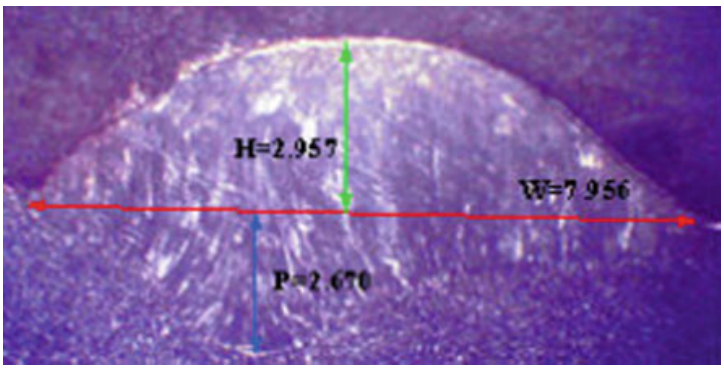


The bead profile parameters such as penetration, bead width and bead height were measured using 'stereo microscope' to get bead profile of experiments. The measured weld profiles are shown in Fig. 19.3.

19.2.1 Effect of Welding Parameters On the Performance Measures

The experimental results obtained from GMAW experiments were analysed through graphical figure. It shows and studies the effects of each parameter on responses. The effect of wire feed rate (WFR) on different geometries of weld beads, such as penetration, bead width and bead height, is shown in Fig. 19.4. The increase in wire feed rate, penetration and bead width is observed. For all wire feed values, bead height is constant.

Fig. 19.3 Microstructure measurement of a weld bead profile



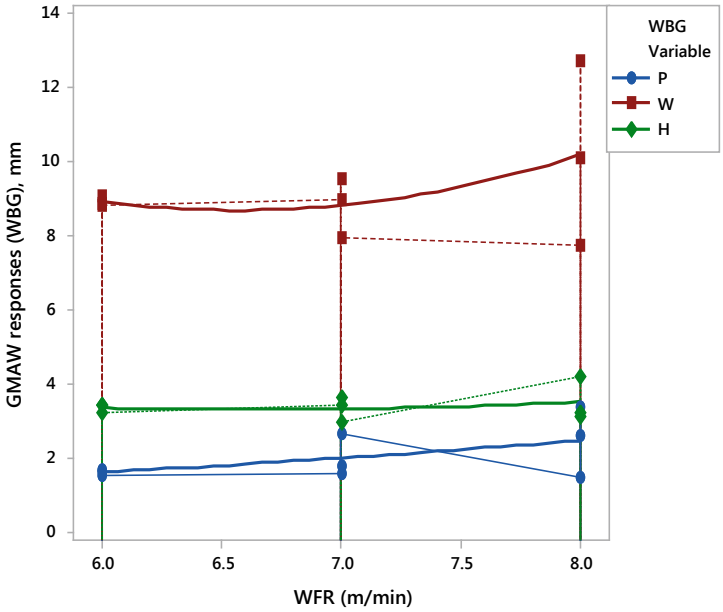


Fig. 19.4 Effect of wire feed rate on P, W and H

Similarly, the effect of other weld parameters is studied. It is observed that increase in welding voltage increases penetration and bead width, whereas bead height decreases very slightly. The increase in the value of the nozzle-to-plate distance lowers the penetration value and the width of the bead.

19.3 Development of RSM Modelling

The experimental results are used to develop predictive modelling of different weld characteristics using response surface methodology (RSM). RSM is statistical and empirical method being widely used for the development of predictive models. Three different models are developed for predicting P, W and H characteristics of weld bead geometry.

In this work, functional correlation between different weld bead characteristics and process parameters is found using regression analysis. The linear models with interaction terms show better correlation. In general, linear model with interaction terms is given in Eq. 19.1.

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{j=i+1}^{k} \sum_{i=0}^{k} \beta_{ij} x_i x_j + \varepsilon, \tag{19.1}$$

where y is the response and x_i (1, 2, …, k) is the level of k variable. β_0 is constant, β's are model coefficients, and ε is modelling error, respectively. The regression model is developed using statistical software Minitab 17®. The predictive model equations obtained for P, W and H are given in Eqs. 19.2, 19.3 and 19.4,

respectively. The correlation coefficient (R^2) values of each model equation are above 95%.

$$P = 34.11 - 5.093\,\text{WFR} - 1.358\,V - 0.427\,N + 0.00254\,\text{WS} + 0.2176\,\text{WFR} \times V + 0.0379\,\text{WFR} \times N + 0.0057\,V \times N, \quad (19.2)$$

(R^2 = 99.96%, R^2 = 99.67%)

$$W = 187.6 - 23.6\,\text{WFR} - 9.24\,V - 0.58\,N + 0.0124\,\text{WS} + 1.180\,\text{WFR} \times V - 0.074\,\text{WFR} \times N + 0.062\,V \times N, \quad (19.3)$$

(R^2 = 95.03%, R^2 = 60.23%)

$$H = -14.6 + 3.46\,\text{WFR} + 0.896\,V - 0.185\,N - 0.00489\,\text{WS} - 0.163\,\text{WFR} \times V + 0.112\,\text{WFR} \times N + 0.0057 \times N, \quad (19.4)$$

(R^2 = 97.15%, R = 77.23)

Table 19.2 presents predictive performance of the model comparing with the experimental results, and it shows an average error of 0.621, 4.780 and 1.89% for P, W and H, respectively. Also, the maximum error percentage is well within 5.5%, except one data set for W. The established models can be used with considerable accuracy to determine penetration, bead width and bead height.

Table 19.2 Model predictions

S. No.	Experimental results			Model prediction			% Error		
	P	W	H	P	W	H	P	W	H
1	1.686	8.863	3.427	1.693	8.984	3.383	0.42	1.360	1.280
2	1.600	9.105	3.415	1.575	8.504	3.236	1.58	1.600	5.240
3	1.542	8.851	3.224	1.549	9.016	3.181	0.46	1.860	1.330
4	1.554	8.976	3.403	1.571	9.360	3.401	1.09	4.270	0.060
5	1.764	9.546	3.607	1.781	9.952	3.600	0.96	4.250	0.200
6	2.670	7.956	2.957	2.685	8.344	2.950	0.56	4.870	0.230
7	1.452	7.751	4.190	1.447	7.656	4.095	0.35	1.220	2.260
8	2.581	10.103	3.216	2.575	9.992	3.117	0.24	1.090	3.070
9	3.378	12.736	3.139	3.372	14.968	3.04	0.18	17.520	3.150
Average percentage error (APE)							0.621	4.780	1.890

19.4 Parametric Analysis

The effect of weld parameters on various weld measures is analysed through contour figures. Three weld bead characteristics, viz. penetration, bead width and bead height, are studied over the range of parameters under consideration. The effect of responses is investigated by two different parameters, while other parameters are kept constant at middle level. As the wire feed rate and voltage increase, the penetration and bead width increase, while the value of penetration and bead width decreases with increase in the value of nozzle-to-plate distance. In this contour figures, you can better understand how weld parameters interact with various weld bead responses.

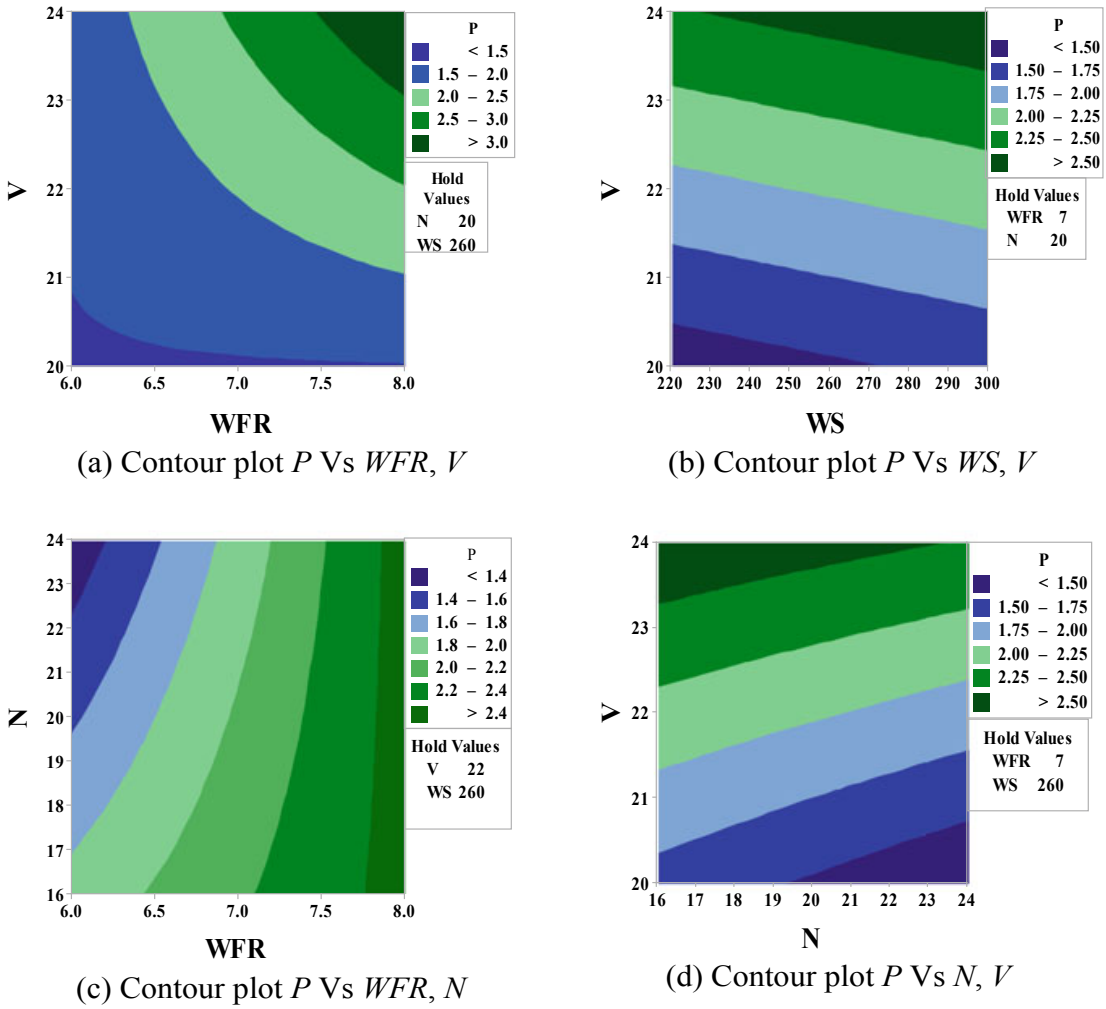
19.4.1 Contour Figures for Penetration (P)

From the results, it is clear that the welding voltage (V), interaction terms WFR $\times$ V and wire feed rate (WFR) are the most significant parameters that affect the penetration. Four figures, viz. P versus WFR and V, WS and V, WFR and N, and V and N, are figured, and it is shown in Fig. 19.5. The dark green colour indicates maximum penetration, and dark blue colour indicates minimum penetration.

The variation in V and WFR on P is shown in Fig. 19.5a. It shows that V and WFR bears a direct relationship with penetration. Increased value of V and WFR increases P in a range of 1.5–3 mm. The Fig. 19.5b also shows that V and N direct variation in P, and Fig. 19.5c shows that as the value of WFR increases, the value of P increases. A minimal effect of N is observed in case of penetration. Figure 19.5d shows the variation of P with N and V that shows direct relationship. The deeper penetration needs high values of input factors.

19.4.2 Contour Figures for Bead Width (W)

Figure 19.6 shows the contour figures for bead width (W) versus WFR and V, WFR and N, V and WS, *and* V and N. From the Fig. 19.6a, it is observed that WFR and N holds a direct relationship with bead width, and as WFR and N increases, the value W also increases. Figure 19.6b shows that with an increase in V, the value of W decreases. From the Fig. 19.6c, W increases with increase in the value of N and WS, and in Fig. 19.6d, low value of N shows minimum W, and W slightly increases with increase in N. W is almost unaffected with an increase in V.



(a) Contour plot *P* Vs *WFR*, *V*

(b) Contour plot *P* Vs *WS*, *V*

(c) Contour plot *P* Vs *WFR*, *N*

(d) Contour plot *P* Vs *N*, *V*

Fig. 19.5 Contour figures for penetration

19.4.3 Contour Figures for Bead Height (H)

Contour figures of *H* versus WFR and *V*, WFR and WS, *and V* and *N* shown in Fig. 19.7 provide better understanding of weld parameter on *H*. From the Fig. 19.7a, it is observed that the value of height (*H*) increases with increase in wire feed rate (*W*) and remains unaffected by *V* value. From Fig. 19.7b, it is also observed that with an increase in WFR, bead height increases and *H* decreases with an increase in welding speed (WS). From the Fig. 19.7c, *H* is maximum at low value of *V* and minimum when the welding voltage (*V*) increases and *H* increases with increase in nozzle-to-plate distance (*N*).

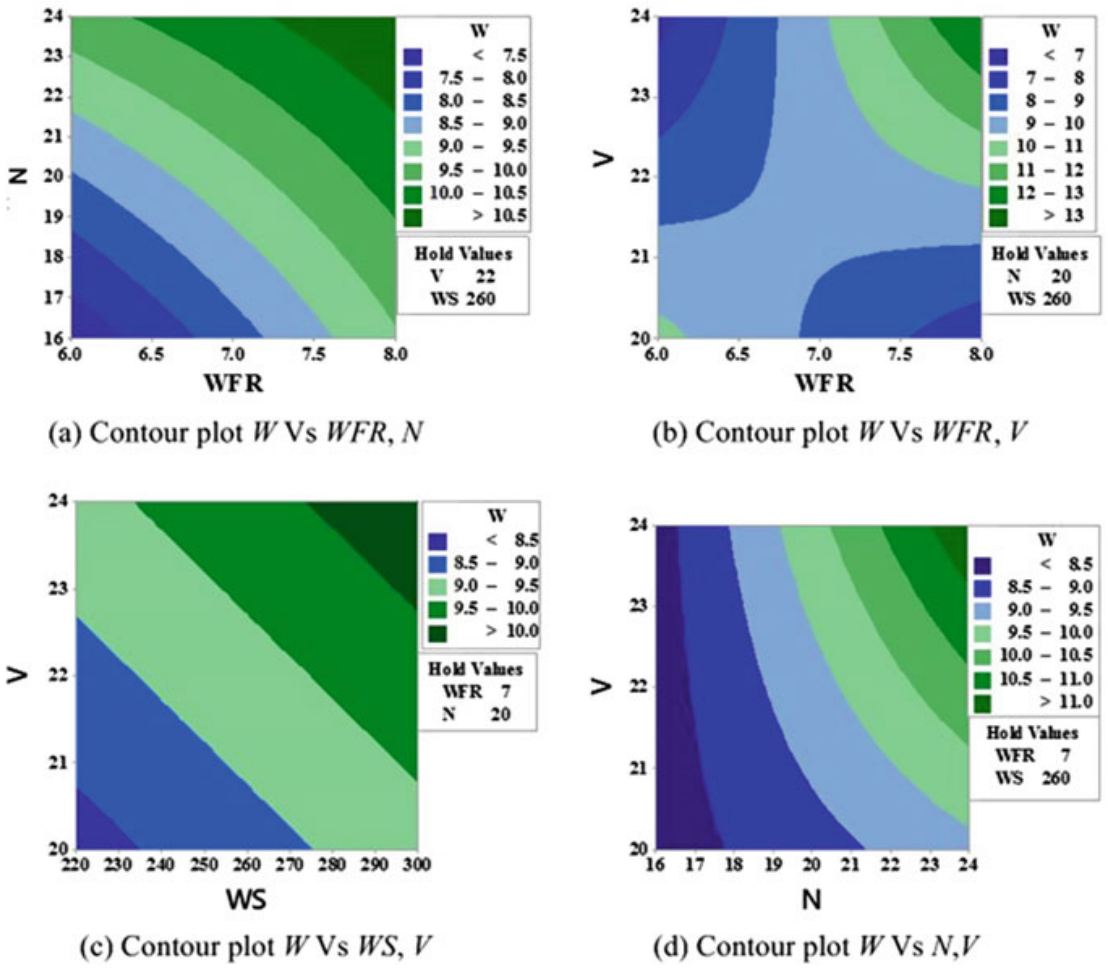
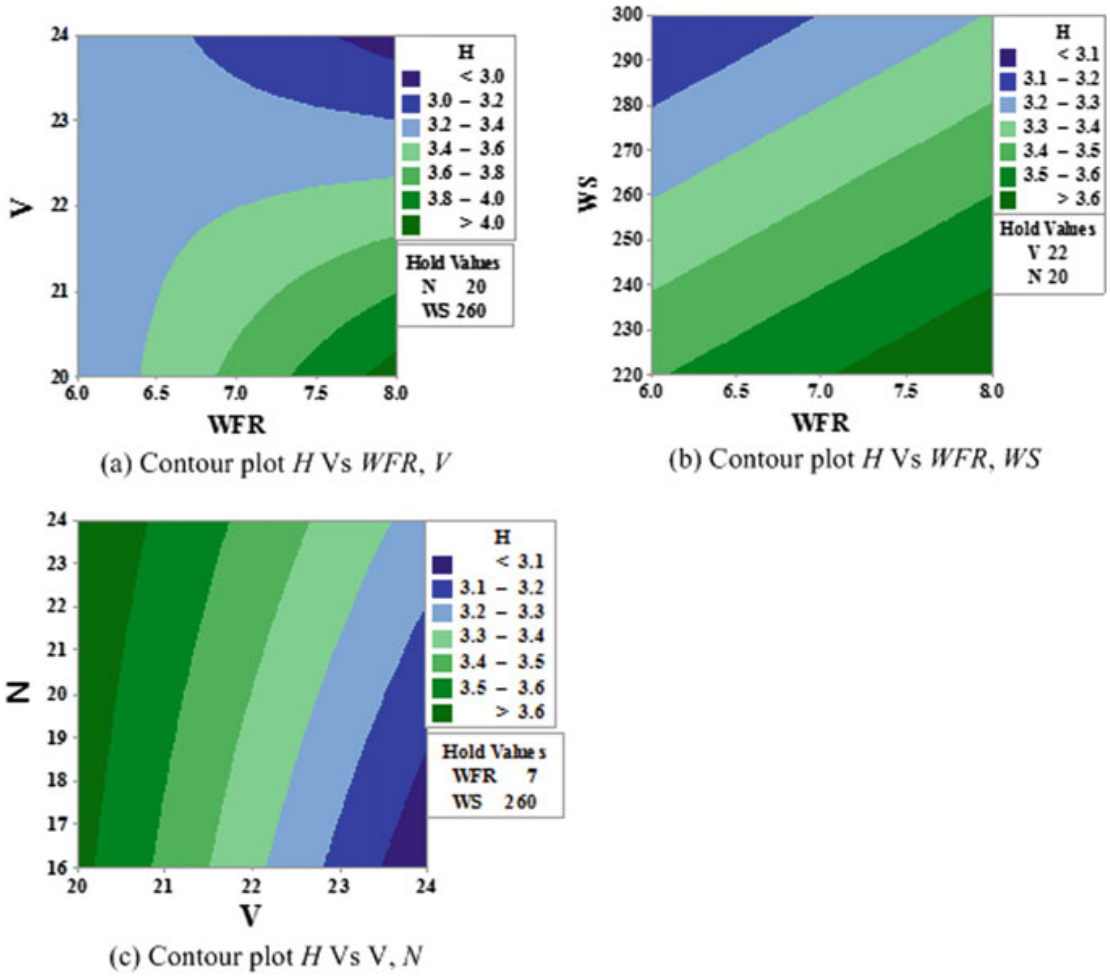


Fig. 19.6 Contour figures for bead width

19.5 Optimization of GMAW Welding

19.5.1 Single-Objective Optimization Using GA

In single-objective optimization, optimum process parameters are obtained to maximize penetration and minimize bead width and bead height individually. The optimization problem is formulated with fitness function derived from regression model equations and is optimized using GA toolbox available in MATLAB 14. The function P (penetration) is maximized and function W (bead width) and H (bead height) are minimized subjected to variable bounds of the factors. The range of the level of each factor decides the upper and lower bounds of the variables. The GA optimization is carried out for both single and multiple objectives. The problem is formulated for single-objective optimization as follows:



(a) Contour plot *H* Vs *WFR*, *V*

(b) Contour plot *H* Vs *WFR*, *WS*

(c) Contour plot *H* Vs V, *N*

Fig. 19.7 Contour figures for bead height

Objective 1

$$\textbf{Maximize } \boldsymbol{P}(\boldsymbol{WFR}, \boldsymbol{V}, \boldsymbol{N}, \boldsymbol{WS}) = -(34.11 - 5.093\,\text{WFR} - 1.358\,V - 0.427\,N + 0.00254\,\text{WS} + 0.2176\,\text{WFR} \times V + 0.0379\,\text{WFR} \times N + 0.00570\,V \times N) \tag{19.5}$$

Objective 2

$$\textbf{Minimize } \boldsymbol{W}(\boldsymbol{WFR}, \boldsymbol{V}, \boldsymbol{N}, \boldsymbol{WS}) = 187.6 - 23.6\,\text{WFR} - 9.24V - 0.58N + 0.0124\,\text{WS} + 1.180\,\text{WFR} \times V - 0.074\,\text{WFR} \times N + 0.062\,V \times N \tag{19.6}$$

Objective 3

$$\textbf{Minimize } \boldsymbol{H}(\boldsymbol{WFR}, \boldsymbol{V}, \boldsymbol{N}, \boldsymbol{WS}) = -14.6 + 3.46\,\text{WFR} + 0.896\,V - 0.185\,N - 0.00489\,\text{WS} - 0.163\,\text{WFR} \times V + 0.0112\,\text{WFR} \times N + 0.0057\,V \times N \tag{19.7}$$

Table 19.3 GA results for single-objective optimization

S. No.	Optimization objective	No. of iterations	Optimum process parameters				Results obtained (mm)
			WFR (m/min)	V (V)	N (mm)	WS (mm/min)	
1	Maximize P	121	8	24	20	258	3.468
2	Minimize W	134	6	24	16	220	4.312
3	Minimize H	109	8	24	16	300	2.483

The variable process is bounded by its lower bound (LB) and upper bound (UB). The LB and UB are given as $6 \leq \text{WFR} \leq 8$; $20 \leq V \leq 24$; $16 \leq N \leq 24$; $220 \leq \text{WS} \leq 300$. The problem can be optimized with the MATLAB 14 GA toolbox. Table 19.3 shows GA results for single optimization.

19.5.2 Simultaneous Optimization

In this work, multi-optimization is performed for two different optimization problems. Each optimization problem provides 18 Pareto-optimal fronts that provide different values of optimum parameters. Each combination of optimum parameter provides the corresponding value of maximum penetration (P), minimum bead width (W) and minimum bead height (H) obtained in the weld bead geometry. The problem is formulated as follows:

Multi-optimization objective 1

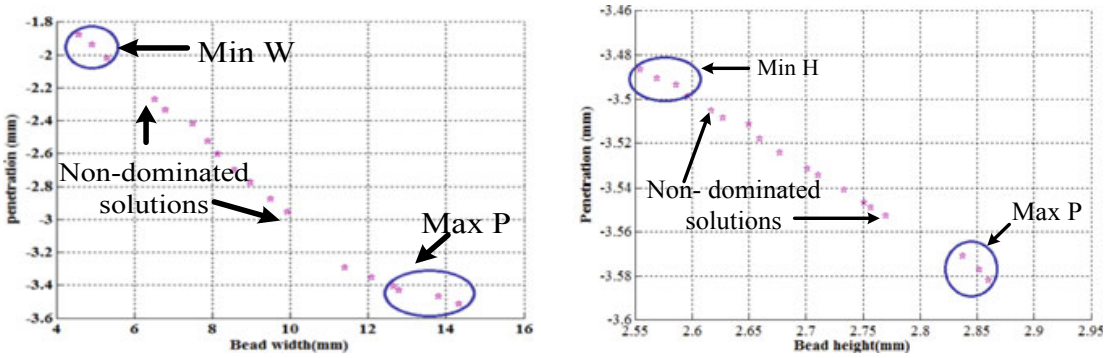
$$\textbf{Maximize } \boldsymbol{P} \textbf{ and minimize } \boldsymbol{W}(\boldsymbol{WFR}, \boldsymbol{V}, \boldsymbol{N}, \boldsymbol{WS}) \tag{19.8}$$

Multi-optimization objective 2

$$\textbf{Maximize } \boldsymbol{P} \textbf{ and minimize } \boldsymbol{H}(\boldsymbol{WFR}, \boldsymbol{V}, \boldsymbol{N}, \boldsymbol{WS}) \tag{19.9}$$

subjected to $6 \leq \text{WFR} \leq 8; 20 \leq V \leq 24; 16 \leq N \leq 24; 220 \leq \text{WS} \leq 300$.

The above multi-objective optimization problem of GMAW of AISI 201 SS is carried out using GA toolbox in MATLAB 14. Figure 19.8 shows the non-dominated Pareto front figure being a number of optimal solutions for objectives 1 and 2, respectively.



(a) For *P* and *W* (b) For *P* and *H*

Fig. 19.8 Pareto-optimal solutions

19.6 Industrial Application Optimization

Manufacturing industries require production of better quality of components having maximum penetration (*P*) with minimum bead width (*W*) and bead height (*H*) for economic production of quality components. In reference to optimization problem 1, GA solutions are given in Table 19.4. For the industrial use of this optimization

Table 19.4 Technology table (*P* and *W*)

Multi-optimization (problem 1)						
S. No.	Optimum parameters				Optimum responses	
	WFR (m/min)	*V* (V)	*N* (mm)	WS (mm/min)	*P* (mm)	*W* (mm)
1	6.001	23.998	16.310	228.799	1.852	4.571
2	6.001	23.998	16.310	228.799	1.852	4.571
3	6.058	23.961	16.530	232.176	1.888	4.954
4	6.192	23.996	16.448	230.687	1.992	5.328
5	6.265	23.991	16.796	229.464	2.025	5.727
6	6.364	23.985	16.351	240.986	2.151	6.022
7	6.576	23.990	16.359	228.908	2.278	6.612
8	6.602	23.977	16.577	242.833	2.323	6.975
9	7.051	23.986	16.627	237.872	2.649	8.490
10	7.405	23.994	17.120	246.260	2.936	10.004
11	7.557	23.991	16.632	242.902	3.047	10.315
12	7.679	23.992	16.823	240.726	3.134	10.778
13	7.735	23.989	17.895	248.331	3.198	11.420
14	7.853	23.977	18.105	251.684	3.297	11.916
15	7.928	23.928	17.484	270.552	3.377	12.144
16	7.976	23.988	18.390	264.639	3.439	12.595
17	7.977	23.885	20.216	283.906	3.460	13.269
18	7.999	23.982	20.669	286.882	3.541	13.661

results, the value of penetration P is arranged in ascending order. Considering optimum parameters to be obtained for a weld design having $P \geq 2$, the GA provides 14 possible optimal solutions (i.e. S. No. 5–18). For minimum bead width, S. No. 5 is found to be optimum providing minimum bead width W satisfying the penetration (P) in the desired level. If $P \geq 3$, GA provides 8 possible optimal solutions (i.e. S. No. 11–18) and S. No. 11 is the optimal solution providing minimum bead width (W) satisfying penetration (P).

Similarly, in reference to optimization problem 2, GA solutions are given in Table 19.5. Considering optimum parameters to be obtained for a weld design having $P \leq 3.5$, the GA provides 5 possible optimal solutions (i.e. S. No. 1–5). For minimum bead height, S. No. 1 is found to be optimum providing minimum H satisfying the P in the desired level. If $P \geq 3.5$, GA provides 13 possible optimal solutions (i.e. S. No. 6–18) and S. No. 6 is the optimal solution providing minimum H satisfying P.

Table 19.5 Technology table (P and H)

Multi-optimization (problem 2)						
S. No.	Optimum parameters				Optimum responses	
	WFR (m/min)	V (V)	N (mm)	WS (mm/min)	P (mm)	W (mm)
1	7.99	23.995	17.332	281.302	3.493	2.631
2	7.99	23.996	17.385	281.316	3.494	2.633
3	7.99	23.999	17.483	281.315	3.495	2.637
4	7.99	23.998	17.584	281.323	3.497	2.641
5	7.99	23.989	17.757	281.333	3.499	2.648
6	7.99	23.991	17.983	281.312	3.502	2.658
7	7.99	23.999	18.141	281.311	3.504	2.664
8	7.99	23.993	18.222	281.317	3.505	2.667
9	7.99	23.991	18.321	281.324	3.506	2.672
10	7.99	23.992	18.391	281.313	3.507	2.675
11	7.99	23.994	18.474	281.319	3.508	2.678
12	7.99	23.997	18.831	281.320	3.513	2.693
13	7.99	23.991	18.937	281.322	3.514	2.697
14	7.99	23.995	19.030	281.317	3.515	2.701
15	7.99	23.996	19.121	281.325	3.517	2.704
16	7.99	23.994	19.317	281.331	3.519	2.713
17	7.99	23.998	19.429	281.333	3.520	2.716
18	7.99	23.999	19.430	281.332	3.521	2.717

19.7 Conclusion

This study investigates the influence of process parameters of gas metal arc welding (GMAW) like wire feed rate, welding voltage, nozzle-to-plate distance and welding speed on weld bead geometry such as penetration, bead width and bead height during the welding process of stainless steel of AISI 201 grade. Parametric analysis and development of predictive model are carried out. The most important parameters for the geometry of weld beads are welding voltage and wire feed rates. The increase in wire feed rates and voltage increases penetration and bead width, while increasing the distance between nozzles decreases the penetration value and bead width. The deeper penetration needs high values of input factors. The developed model having R^2 value greater than 95% and predictive performance with an average error of less than 5% confirms the effectiveness of the model. The single optimization using GA provides maximum P of 3.468 mm, minimum W of 4.312 mm and minimum H of 2.483 mm corresponding to different combinations of weld parameters. The simultaneous optimization of both (i) penetration (P) and bead width (W); (ii) penetration (P) and bead height (H) using GA gives 18 combinations of Pareto-optimal solutions. Optimized results can be used as guidelines for manufacturing industry to select optimum welding parameters to produce desired penetration value (P) with either minimum width or height of weld bead. The approach is found to be an effective and have minimum effort.

Acknowledgements The authors acknowledge the financial support received from NERIST, Arunachal Pradesh, under TEQIP-II scheme and Advance Welding Training and Research Centre in Ajay Kumar Garg Engineering College, Ghaziabad, Uttar Pradesh, in carrying out the experimentation work.

References

1. David, S.B., Cheryl, A.B.: AL 201HPTM (UNS S20100) alloy A high performance, lower nickel alternative to 300 series alloy. Reprinted with the permission of KCI Publishing, Netherlands (3) 17–20 (2005)
2. Chuaiphan, W., Srijaroenpramong, L.: Effect of welding speed on microstructures, mechanical properties and corrosion behaviour of GTA-welded AISI 201 stainless steel sheets. J. Mater. Process. Technol. **214**, 402–408 (2014)
3. Houldcroft, P.T.: Welding Process Technology, 1st edn. Cambridge University Press, Cambridge (1977)
4. Choudhury, B., Chandrasekaran, M.: Investigation on welding characteristics on aerospace materials—a review. Mater. Today Proc. **4**, 7519–7526 (2017)
5. Juang, S.C., Tarng, Y.S.: Process parameters selection for optimizing the weld pool geometry in the tungsten inert gas welding of stainless steel. J. Mater. Process. Technol. **122** 33–37 (2002)
6. Ibrahim, I.A., Mohamat, S.A., Amir, A., Ghalib, A.: The effect of gas metal arc welding (GMAW) processes on different welding parameters. Procedia Eng. **41**, 1502–1506 (2012)

7. Datta, S.V., Bandyopadhyay, A.B., Kumar, P.: Grey-based Taguchi method for optimization of bead geometry in submerged arc bead-on-plate welding. Int. J. Adv. Manuf. Technol. **39**, 1136–1143 (2008)
8. Kim, D., Rhee, S., Park, H.: Modelling and optimization GMA welding by genetic algorithms and response surface methodology. Int. J. Prod. Res. **40**, 1699–1711 (2002)
9. Sathiya, P., Paneerselvam, K., Abdul Jaleel, M.Y.: Optimization of laser welding process parameters for super austenitic stainless steel using artificial neural networks and genetic algorithm. Mater. Des. **36**, 490–498 (2012)

Chapter 20
Finite Element Modelling and Experimental Verification of Dissimilar Joining Between Inconel 718 and SS 316L by Micro-plasma Arc Welding

Ajit Kumar Sahu and Swarup Bag

Abstract In the present research work, an attempt is made to successfully weld Inconel 718, a nickel-based superalloy with SS 316L, an austenitic stainless steel in autogenous mode by using constant current micro-plasma welding process. A finite element method (FEM)-based, three-dimensional (3D) thermal model is developed for butt welding configuration between the selected bimetallic combinations by using a double-ellipsoidal volumetric heat source model. A good consistency is found between the numerically obtained and experimentally obtained weld bead measurements. The numerical model is further used to extract peak temperatures, time–temperature profiles along with average cooling rates of the welding processes. Obtained cooling rates are correlated with the weld micro-structures and mechanical properties. Micro-structural study by SEM analysis has shown coarse columnar dendritic structure in the weld interior for the higher heat input welding case, whereas lower heat input during welding leads to fine weld micro-structure due to high cooling rate. An improvement in the tensile strength and hardness value is observed with increase in weld cooling rate and low heat input during the joining process.

Keywords Inconel 718 · Finite element modelling · Dissimilar welding · Cooling rate · Micro-structure

A. K. Sahu (✉) · S. Bag
Department of Mechanical Engineering, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India
e-mail: ajit.sahu@iitg.ac.in

S. Bag
e-mail: swarupbag@iitg.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_20

20.1 Introduction

Welding of dissimilar materials is always a challenging task due to the differences in material composition and physical properties which need several trials to optimize successful welding parameters. Dissimilar material welding has been in demand for various industries to reduce material cost and to improve flexibility in design. It not only satisfies the service condition but also reduces component cost by reducing the amount of high-cost materials in a component [1]. Inconel 718 is a nickel–chromium-based precipitation hardened superalloy which is strengthened mostly by γ'' (Ni_3Nb) precipitates [2]. Excellent corrosion and oxidation resistance along with extensive mechanical properties at elevated temperature makes Inconel 718 suitable for various critical condition applications. Higher resistance to strain age cracking during the welding makes it one of the most successful high temperatures nickel–chromium alloy ever designed. It is widely used to make hot section components of aero engine, liquid rocket cryogenic engines, and gas turbines blades. It is also widely used in chemical industries, process industries, and nuclear reactors in various severe working conditions [3]. Similarly, austenitic steels are widely used in petrochemical, aerospace, and power generating sector in the form of plates, coils, sheets, bars, and tubing. AISI 316L contains lower amount of carbon which makes this austenitic steel immune to the sensitization at elevated temperatures [4]. AISI 316L exhibits higher creep and tensile strength as compared to other nickel–chromium austenitic steels over a wide temperature range. It also possesses excellent toughness even at the cryogenic temperature which makes it distinct from other chromium–nickel austenitic stainless steels.

The aforesaid features of Inconel 718 and AISI 316L make them suitable to use in the form of bimetallic joint for various engineering applications to make various components in aerospace engine and gas turbine engine. This bimetallic combination is used by ALSTOM Power Sweden for making compressor rotor where high pressure compressor stages are made by Inconel 718 and low pressure stages are made of austenitic based stainless steel [5]. Fusion welding is one of the most preferred joining techniques for this bimetallic combination to establish a good metallurgical bond between the mating surfaces. Although fusion welding of Inconel 718 is not utterly free from welding defects, it can be minimized by taking required precautions during welding process. Defects like solidification cracking and micro-fissuring in the fusion zone (FZ) and heat affected zone (HAZ) are major problem associated with the Inconel 718 and AISI 316L [6]. The generation of intermetallic Laves phase by segregation of niobium (Nb) in the interdendritic region during weld metal solidification is the main cause of above welding defects in Inconel 718 [7], whereas formation of 3–9% of δ ferrite reduces tendency of hot cracking, but higher percentage of ferrite content reduces creep life along with corrosion resistance of the AISI 316L joint [8]. A higher weld cooling rate has resulted in minimization of these defects which can be achieved by using a heat source of higher power density [9]. Thavamani et al. [10] have used gas tungsten arc welding (GTAW) process for welding Inconel 718 at three different heat input

of 530, 670 and 800 J/mm and reported the size of dendrites and crack sensitivity of the weld increasing with increase in heat input. Lower value of heat input improves the weld cooling rate, whereas higher value slows down the weld cooling rate which leads to segregation of various useful alloying element into the interdendritic regions, leading to deterioration of mechanical properties of the welded component. Bansal et al. [11] have used microwave welding to join Inconel 718 and AISI316L with Inconel 718 powder as interfacing layer between the two surfaces and reported good joint strength for the bimetallic welded joint. Ramkumar et al. [12] have used multi-pass GTAW process to weld this bimetallic combination of material and reported superior mechanical properties for the joint as compared to the AISI 316L base material. Ramkumar et al. [13] inspected the dissimilar welding of Inconel 718 and SS 316L with the help of constant and pulsed current mode GTAW process. They found weld prepared by pulsed current mode has superior tensile, hardness and toughness properties as compared to constant current mode weld which is mainly due to the control of Laves phase formation due to higher cooling rate, obtained during pulsed current mode of welding. Kumar et al. [4] also employed CCGTAW and PCGTAW to weld Inconel 625 and AISI 316L and stated presence of various secondary phases in the HAZ are minimal for PCGTAW condition. From the above literature review, it can be concluded that the scope and application of joining this bimetallic combination are wider and required through investigation in terms of joint performance at ambient temperature. The selection of proper welding techniques and process condition is an important aspect of welding dissimilar materials. By using advanced welding techniques, full penetration can be achieved comparatively at a lower heat input; as a result, the weld cooing rate can be enhanced and formation of various deleterious secondary phases can be constricted [14]. From the open literature, it was evident most of the reported work were limited to GTAW process for joining this bimetallic combination. Hence, in the present study, micro-plasma arc welding technique which has higher power intensity among all the arc welding processes is used to join this bimetallic combination [15]. μ-PAW also possesses good arc stability even at very low range of current as compared to GTAW which makes this process more suitable for joining thin sheets. The present study is mainly focused on weldability, mechanical and metallurgical analysis of these bimetallic joints obtained at different condition of continuous current μ-PAW process. Another objective of the present study is to correlate the effect of heat input/cooling rate on fusion zone micro-structure and Laves phase formation by different welding condition.

Direct experimental investigations during welding are extremely expensive and often impossible because of the complexity of the process and presence of high temperature arc. Hence, a three-dimensional finite element (FE)-based numerical model was developed using ABAQUS to obtain the temperature isotherm in the fusion zone along with time–temperature history of μ-PAW welding process. The conductive heat transfer-based numerical modelling is done by using a moving double-ellipsoidal heat source that is accountable for complex physical phenomena like conduction, convection along with radiation heat transfer from the joining surfaces to the ambient. The double-ellipsoidal heat source is used in the current

study due to its unique capability to incorporate steep gradient of temperature in the front of heat source and a gentle gradient in the trailing edge of weld pool [16]. Temperature-based physical properties of the materials were considered in the model. Thereafter, the time–temperature profile of the fusion zone from the numerical model is accounted for estimation of the cooling rate. The influence of weld cooling rate/solidification rate on weld micro-structure and corresponding outcome on mechanical properties are reported in the present study.

20.2 Experimental Procedure

As received Inconel 718 and AISI 316L stainless steel sheets of thickness 0.7 mm were sheared into 100 × 60 mm coupons with welding direction normal to the rolling direction, the chemical composition of the base materials were analysed by energy dispersive X-ray (EDX) technique and results are shown in Fig. 20.1. Based on EDX analysis, the resulted composition (wt%) of Inconel 718 and AISI 316L are given in Table 20.1.

The welding surfaces were cleaned with acetone prior to the welding process for removing oxides and other foreign impurities. Sheets were clamped in a restrain fixture made up of copper to achieve high cooling rate. The welding torch of micro-plasma set-up is kept fixed and the workpiece is allowed to move at a constant velocity. Commercially available argon gas (99% pure) is used as plasma gas as well as used as a shielding gas for the protection of weld pool. Butt configuration weld joints were prepared by constant current (CC) mode using micro-plasma arc welding set-up with a DC current range of 0.05–50 A. The plasma gas and shielding gas flow rate are maintained at 0.5 and 5 L/min, respectively, with a nozzle standoff distance of 2 mm. The other welding parameters employed for making defect free welds are depicted in Table 20.2 which is obtained by a wide range of welding trials.

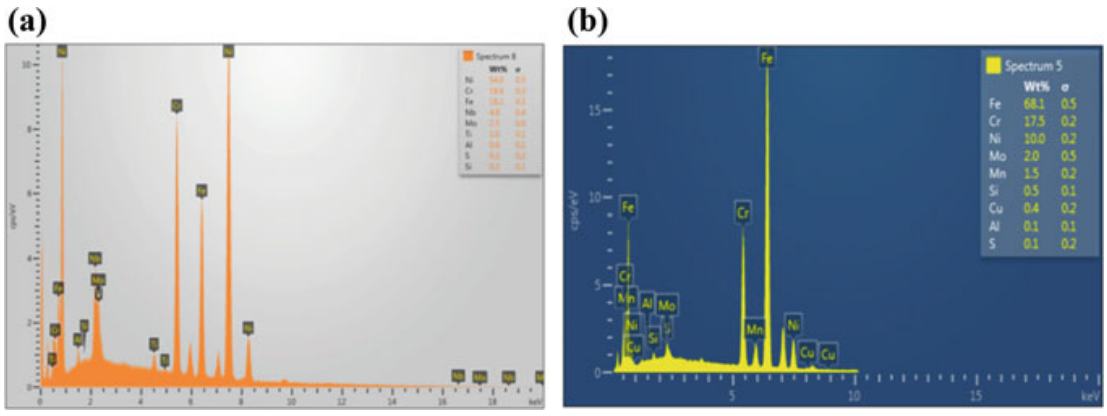


Fig. 20.1 EDX elemental analysis of **a** Inconel 718 and **b** AISI 316L base material

Table 20.1 Chemical composition (wt%) of Inconel 718 and AISI 316L

Elements	Ni	Cr	Nb	Mo	Ti	Al	Si	S	Fe
Inconel 718	54.0	18.6	4.8	2.5	1.0	0.6	0.2	0.2	Bal.
Elements	Ni	Cr	Mo	Mn	Si	Cu	Al	S	Fe
AISI 316L	10.0	17.5	2.0	1.5	0.5	0.4	0.1	0.1	Bal.

Table 20.2 Process parameters for μ-PAW of Inconel 718 and SS 316L weldments

Sample	Current (A)	Voltage (V)	Welding speed (mm/s)	Heat input (J/mm)
S1	14	24	3.4	99
S2	15	24	3.2	112

The welded samples were sheared normal to the welding direction and polished for macro-structural analysis. Then, the mirror polished samples were etched in a mixture of HCl, HF and CH_3COOH in the ratio of 1.5:1:1, respectively, for the micro-structural analysis. The micro-structural analyses were done by the help of optical microscope and scanning electron microscope. Vickers's micro-hardness measurements were done across the weld bead for each specimen to analyse the hardness behaviour. Tensile test of the welded specimens and base materials are conducted at room temperature to analyse the yield strength, ultimate tensile strength and percentage elongation. For each condition, two sub-size transverse specimens were prepared according to ASTM E8 M standard with a gauge length of 25 mm. The tests were conducted by a computer controlled servo-hydraulic universal tensile testing machine with a cross-head speed of 1 mm/min.

20.3 Numerical Model

Experimental investigations have certain limitations due the involvement of high temperature heat source during the welding process. Hence, a finite element-based conduction heat transfer model is developed for the micro-plasma arc welding process to obtain temperature distribution of the entire weldment. The moving heat source is incorporated in the model by using a DFLUX subroutine with the numerical model. To approximate the influence of convective transport of heat in weld pool, a Gaussian distributed double-ellipsoidal volumetric heat source is used to distribute the heat flux into the joining surfaces. Temperature-based thermo-physical properties for SS 316L and Inconel 718 were considered in the present analysis [17, 18]. For simplification, following assumptions were considered in the model: Top surface of the fusion zone is assumed as flat and neglecting the effect of plasma arc and shielding gas. The initial temperature of workpiece was taken as uniform ambient temperature (T_0) of 300 K. The latent heat of melting and

cooling is incorporated, whereas material evaporation from the fusion zone is neglected. The surface of the base plates which is exposed to heat flux and other surfaces is subjected to heat losses due to convection and radiation. The surfaces which are directly exposed to air were assigned a convective heat transfer coefficient (h) of 35 W/m^2 K [16]. However, surfaces which are in direct contact with the copper fixture and backing plate, which have high thermal conductivity, are assigned with a high heat transfer coefficient of 1000 W/m^2 K [19]. The fundamental governing equation of heat transfer or conservation of energy in 3D Cartesian coordinate is expressed as

$$\rho C\left(\frac{\partial T}{\partial t}-v\frac{\partial T}{\partial x}\right)=\frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(k\frac{\partial T}{\partial y}\right)+\frac{\partial}{\partial z}\left(k\frac{\partial T}{\partial z}\right)+\dot{q} \qquad (20.1)$$

where (x, y, z) is the coordinate system associated with the moving heat source. ρ, k and C refer to density, thermal conductivity and specific heat of the base material, respectively. T and t represents temperature and time variable, respectively. $\dot{q}$ is the internal heat generation per unit time and per unit volume and 'v' is considered as the velocity of the plasma arc in welding direction (x-axis) or moving coordinate axis. The meshing arrangement of base plate is shown in Fig. 20.2. Very fine meshing is provided in the fusion zone area to capture thermal history precisely as it is exposed to a very concentric heat source where the temperature gradient becomes steep. While coarser meshing was generated in the area far from the heat flux to reduce computational cost, Continuum solid eight nodded brick elements (DC3D8 type) were selected in the thermal analysis for diffusive heat transfer.

The double-ellipsoidal parameters like width (b) and depth (c) of the double-ellipsoidal heat source were directly taken from the experimental results. The length of front (c_f) and rear (c_r) quadrant were estimated as a function of velocity [20]. In dissimilar welding, the fusion zone becomes a mixture of both base materials which will produce various intermetallic compounds during the

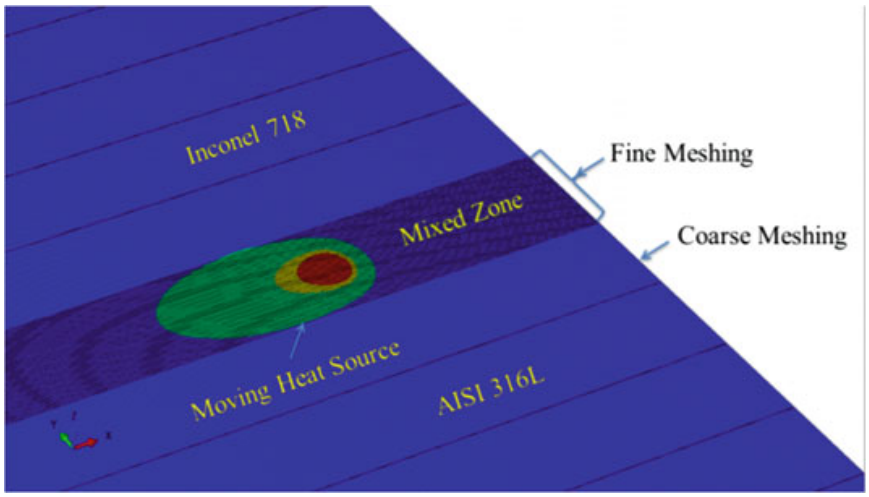


Fig. 20.2 Finite element meshing of solution domain

solidification and will have an impact the joint efficiency. Therefore, to correlate the weld cooling rate with intermetallic phase formation precisely, the temperature dependent thermo-physical properties of fusion zone were considered as the average of both base materials which is named as mixed zone as shown in Fig. 20.2.

20.4 Results and Discussion

20.4.1 Temperature Field Characteristics

Temperature distribution corresponding to S1 welding condition along with the moving heat source is shown in Fig. 20.3. A peak temperature of 2169 K is achieved at the weld centre line. The fusion zone (FZ) is defined by 1730 K which is the liquidus temperature of AISI 316L, i.e. maximum between the base materials [1]. And the heat affected zone (HAZ) was defined below 1547 K which is the solidus temperature of Inconel 718, i.e. minimum between the base materials [21]. The region between these two isotherm contours is considered as mushy zone where liquid and solid phase coexists together.

The comparison between simulated and experimentally estimated weld macrographs in transverse cross section for both welding conditions is shown in Fig. 20.4, where FZ, mushy zone, and HAZ are well defined. The shape and size of the numerically simulated outcomes were found to be in good agreement with the experimentally obtained weld macrographs. With increase in welding heat input per unit length, the size of weld bead increases and the fusion zone becomes wider. The weld cross section of S2 welding condition is little bit of straight type as compared

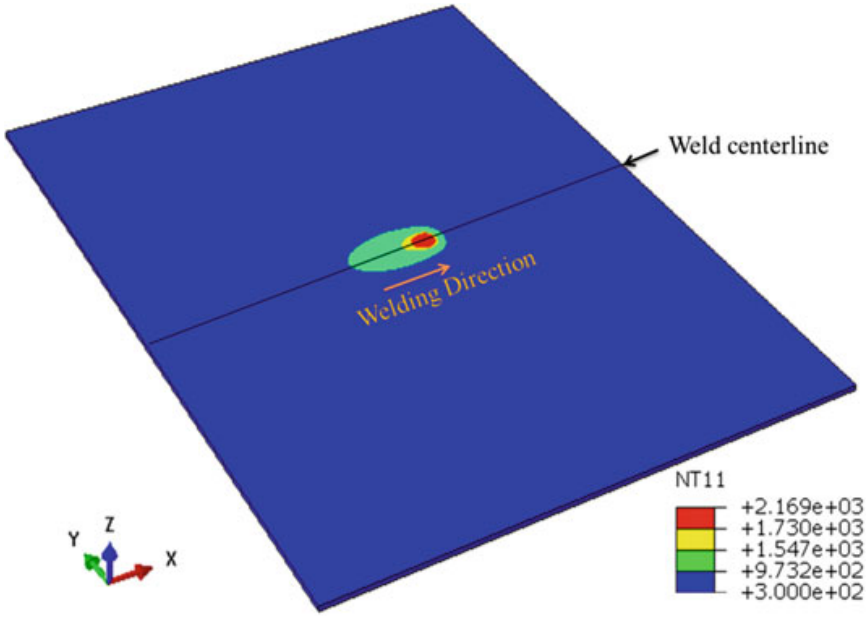


Fig. 20.3 Temperature distribution for S1 welding condition

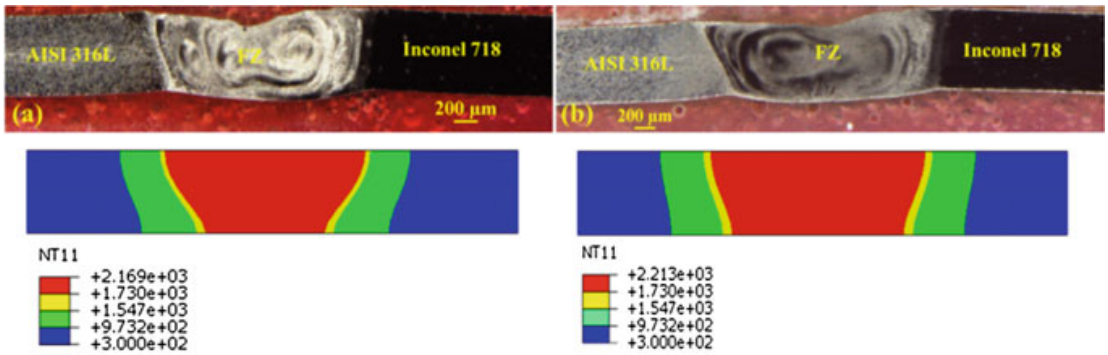


Fig. 20.4 Comparison of weld macrograph for **a** S1 and **b** S2 welding condition

to S1 welding condition which is due to higher heat input during welding process. The maximum error found to be 6% in the weld bead width comparison at the cap and root between experimentally and numerically measured value (Table 20.3).

Figure 20.5 shows the time–temperature profiles of S1 and S2 welding condition at weld centre along with the average weld cooling rate. The temperature rises rapidly as the heat source approaches and drops steadily to ambient temperature when the heat source passes away from the node. The quality of weld is greatly dependent on heat input per unit length during the welding process, peak

Table 20.3 Weld bead dimensions at different location

Location	S1		S2	
	Experimental	Numerical	Experimental	Numerical
Weld cap (mm)	2.46	2.38	2.6	2.44
Weld root (mm)	1.91	1.8	1.96	1.94

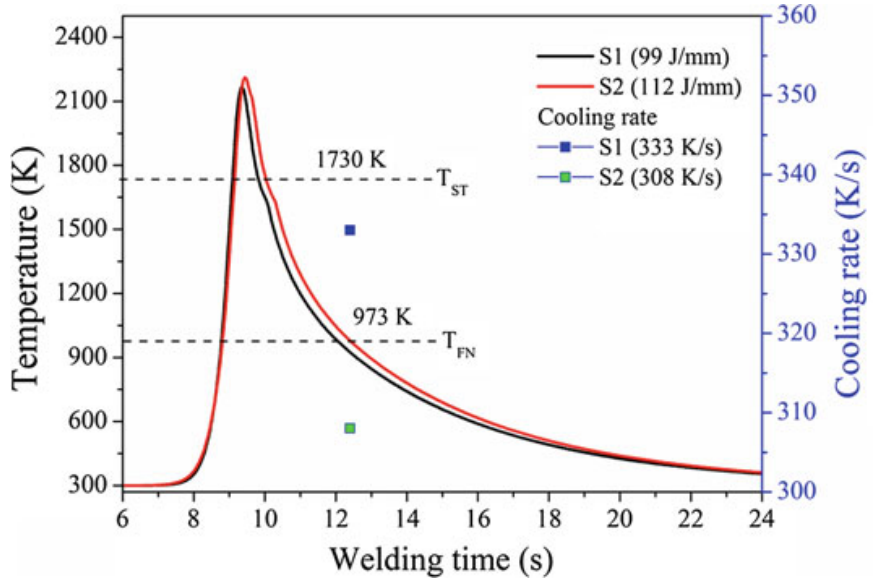


Fig. 20.5 Time–temperature profile and average cooling rate of S1 and S2 welding conditions

temperature, and cooling rate. The peak temperature of S1 and S2 welding conditions were found to be 2169 and 2213 K, respectively. Higher peak temperature for S2 welding condition is attributed to the higher heat input during welding process. The average weld cooling rate was calculated by taking the slope of time–temperature curve between transformation start (T_{ST}) and finish temperature (T_{FN}). Higher liquidus temperature among Inconel 718 and SS 316L was considered as T_{ST} as below this temperature phase transformation initiated. Minimum value of solid phase transformation end temperature among the bimetallic is taken as T_{FN} as below this temperature phase transformation does not occurs. The calculated average cooling rate for S1 and S2 welding condition were 333 and 308 K/s. When the weld peak temperature increases, weld pool lasts for longer period of time resulted in lower average cooling rate.

20.4.2 Effect of Cooling Rate on Weld Micro-structure and Mechanical Properties

Formation of various intermetallic phases during weld cooling period is the major challenge of fusion welding process. Segregation of Mo and Nb in the interdendritic region leads to formation of deleterious Laves phases in Inconel 718 which promotes liquation cracks and micro-fissuring in the FZ and HAZ [14]. In AISI 316L, formation of higher percentage of δ ferrite in interdendritic region has deleterious effect on mechanical properties and corrosion resistance of the joint. The scale of solidification structure is dominated by the cooling rate; higher cooling rate will produce finer dendrites, whereas lower value will produce coarse dendrites. In Fig. 20.6, higher magnification SEM micro-structure of weld interior for both the welding conditions is shown. Higher cooling rate of S1 weld condition (333 K/s) leads to fine micro-structure as compared to S2 weld condition (308 K/s). Coarse columnar dendrites were observed in the weld interior for S2 welding condition, whereas fine interconnected equiaxed type of micro-structure was observed in S1 welding condition.

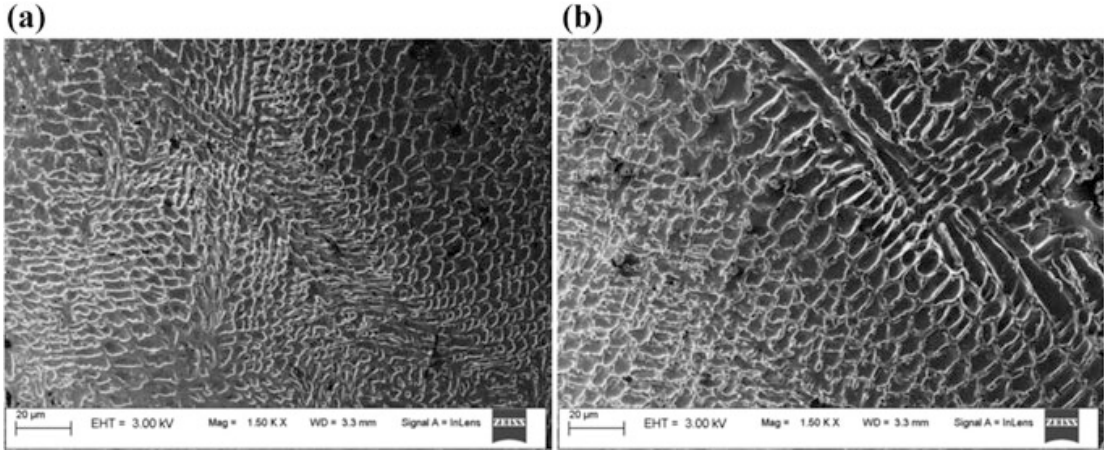


Fig. 20.6 Fusion zone micro-structure of **a** S1 and **b** S2 welding condition

Table 20.4 Results of tensile test

	Inconel 718 (base material)	AISI 316L (base material)	S1 (weld joint)	S2 (weld joint)
YS (MPa)	594	408	389	337
UTS (MPa)	878	679	653	596
Elongation (%)	23	39	26	21

The weld cooling rate plays a crucial role in controlling interdendritic segregation and formation of various intermetallic phases in the fusion zone. Lower heat input assures a higher cooling rate during the solidification that hinders the segregational tendency of solutes in the interdendritic regions, resulting in a fined micro-structure in the fusion zone. Formation of various intermetallic phases consumes a notable amount of beneficial alloying elements from the base material which results deterioration of weld quality. Hence, control of intermetallic Laves and δ-ferrite phases in the solidified zones improves mechanical properties of the joint. The tensile properties of both base and welded materials are given in Table 20.4. S1 welded samples were failed in the AISI 316L base material sides, whereas S2 welded samples were failed in the fusion zone as shown in Fig. 20.7. The tensile strength of S1 weld metal were found superior to the S2 weld material which is mainly due to the fined micro-structure of the fusion zone that reduces segregation and formation of intermetallic in the grain boundary. S1 welded sample has a tensile strength of 96% to that of the AISI 316L base material and a tensile strength of 74% to that of Inconel 718. The improvement in tensile strength of S1 welded condition is mainly attributed to the fine weld micro-structure which is mainly due to the higher cooling rate achieved during the welding process.

Cross-weld micro-hardness profile of the S1 and S2 weldments are shown in Fig. 20.8. The average micro-hardness values of base material are 219 HV and 134 HV for Inconel 718 and AISI 316L, respectively. Fusion zone of S1 welding condition has an average micro-hardness value of 154 HV, whereas for S2 welding condition, it has a lower value of 148 HV. Higher hardness of S1 fusion zone is mainly due to the fine micro-structure because of higher cooling rate. The hardness is found to be less in the HAZ region as compared to the FZ and BM, which is mainly attributed to the grain coarsening from FZ towards BM.

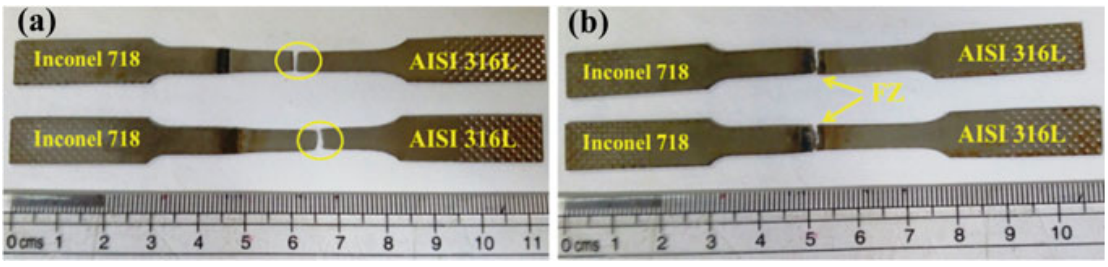


Fig. 20.7 Location of tensile fracture for **a** S1 and **b** S2 welding conditions

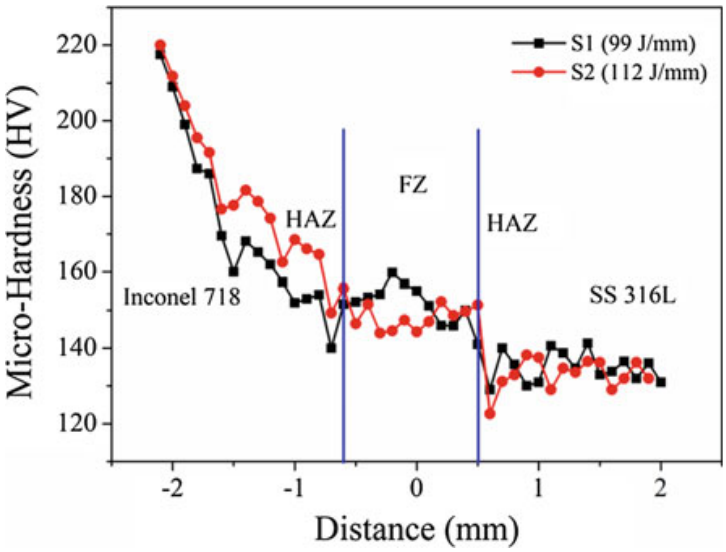


Fig. 20.8 Cross-weld micro-hardness distribution

20.5 Conclusion

In the present study micro-plasma arc welding between Inconel 718 with AISI 316L sheet materials with butt joint configuration has been undertaken. A three-dimensional FE-based thermal model is developed to investigate the thermal cycle during the welding process. The following conclusions were drawn from the current study

- A good agreement was obtained between the shape and size of computed and experimentally obtained weld bead.
- Full weld penetration for the bimetallic combination was achieved at a heat input of 99 J/mm.
- A maximum cooling rate of 333 K/s was calculated from the numerical model for the low heat input of 99 J/mm.
- Refinement in weld micro-structure is observed at 333 K/s cooling rate as compared to 308 K/s which results in improved joint efficiency (96.17% of AISI 316L base material) and ductility (26%) of the welded joint.
- A maximum micro-hardness difference of 6 HV is reported between the FZ of 99 and 112 J/mm welding condition.

Acknowledgements The authors would like to acknowledge the Department of Mechanical Engineering, Central Instrument Facility (CIF), and Central Workshop of IIT Guwahati, India for providing all facilities to carry out the experimental and numerical work.

References

1. Lee, H.T., Jeng, S.L., Yen, C.H., Kuo, T.Y.: Dissimilar welding of nickel-based Alloy 690 to SUS 304L with Ti addition. J. Nucl. Mater. **335**, 59–69 (2004)
2. Ye, X., Hua, X., Wang, M., Lou, S.: Controlling hot cracking in Ni-based Inconel-718 superalloy cast sheets during tungsten inert gas welding. J. Mater. Process. Technol. **222**, 381–390 (2015)
3. Huang, C.A., Wang, T.H., Lee, C.H., Han, W.C.: A study of the heat-affected zone (HAZ) of an Inconel 718 sheet welded with electron-beam welding (EBW). Mater. Sci. Eng., A **398**, 275–281 (2005)
4. Kumar, K.G., Ramkumar, K.D., Arivazhagan, N.: Characterization of metallurgical and mechanical properties on the multi-pass welding of Inconel 625 and AISI 316L. J. Mech. Sci. Technol. **29**, 1039–1047 (2015)
5. Henderson, M.B., Arrell, D., Larsson, R., Heobel, M., Marchant, G.: Nickel based superalloy welding practices for industrial gas turbine applications. Sci. Technol. Weld. Joining **9**, 13–21 (2004)
6. Radhakrishnan, B., Thompson, R.G.: A model for the formation and solidification of grain boundary liquid in the heat-affected zone (HAZ) of welds. Metall. Mater. Trans. A **23**, 1783–1799 (1992)
7. Janaki Ram, G.D., Venugopal Reddy, A., Prasad Rao, K., Reddy, G.M., Sarin Sundar, J.K.: Microstructure and tensile properties of Inconel 718 pulsed Nd-YAG laser welds. J. Mater. Process. Technol. **167**, 73–82 (2005)
8. Kar, J., Roy, S.K., Roy, G.G.: Effect of beam oscillation on microstructure and mechanical properties of AISI 316L electron beam welds. Metall. Mater. Trans. A **48**, 1759–1770 (2017)
9. Radhakrishna, C.H., Rao, K.P.: The formation and control of Laves phase in superalloy 718 welds. J. Mater. Sci. **32**, 1977–1984 (1997)
10. Thavamani, R., Balusamy, V., Nampoothiri, J., Subramanian, R., Ravi, K.R.: Mitigation of hot cracking in Inconel 718 superalloy by ultrasonic vibration during gas tungsten arc welding. J. Alloy. Compd. **740**, 870–878 (2018)
11. Bansal, A., Sharma, A.K., Das, S., Kumar, P.: On microstructure and strength properties of microwave welded Inconel 718/stainless steel (SS-316L). Proc. Inst. Mech. Eng., Part L: J. Mater. Des. Appl. **230**, 939–948 (2016)
12. Ramkumar, T., Selvakumar, M., Narayanasamy, P., Begam, A.A., Mathavan, P., Raj, A.A.: Studies on the structural property, mechanical relationships and corrosion behaviour of Inconel 718 and SS 316L dissimilar joints by TIG welding without using activated flux. J. Manuf. Process. **30**, 290–298 (2017)
13. Ramkumar, K.D., Patel, S.D., Praveen, S.S., Choudhury, D.J., Prabaharan, P., Arivazhagan, N., Xavior, M.A.: Influence of filler metals and welding techniques on the structure–property relationships of Inconel 718 and AISI 316L dissimilar weldments. Mater. Des. (1980–2015) **62**, 175–188 (2014)
14. Manikandan, S.G.K., Sivakumar, D., Rao, K.P., Kamaraj, M.: Effect of weld cooling rate on Laves phase formation in Inconel 718 fusion zone. J. Mater. Process. Technol. **214**, 358–364 (2014)
15. Kou, S.: Welding Metallurgy. Wiley, New Jersey, USA (2003)
16. Goldak, J., Chakravarti, A., Bibby, M.: A new finite element model for welding heat sources. MTB **15**, 299–305 (1984)
17. Rahman Chukkan, J., Vasudevan, M., Muthukumaran, S., Ravi Kumar, R., Chandrasekhar, N.: Simulation of laser butt welding of AISI 316L stainless steel sheet using various heat sources and experimental validation. J. Mater. Process. Technol. **219**, 48–59 (2015)
18. Dye, D., Hunziker, O., Roberts, S.M., Reed, R.C.: Modeling of the mechanical effects induced by the tungsten inert-gas welding of the IN718 superalloy. Metall. Mater. Trans. A **32**, 1713–1725 (2001)

19. Baruah, M., Bag, S.: Influence of heat input in microwelding of titanium alloy by micro plasma arc. J. Mater. Process. Technol. **231**, 100–112 (2016)
20. Yadaiah, N., Bag, S.: Effect of heat source parameters in thermal and mechanical analysis of linear GTA welding process. ISIJ Int. **52**, 2069–2075 (2012)
21. Aktaş Çelik, G., Polat, Ş., Atapek, Ş.H., Haidemenopoulos, G.N.: Microstructural and mechanical characterization of solidified austenitic stainless steels. Arch. Foundry Eng. **17**, 163–167 (2017)

Chapter 21
Numerical Simulation and Experimental Validation of A-TIG Welding of 2.25Cr–1Mo Steel

A. R. Pavan, B. Arivazhagan, S. Arun Kumar, M. Vasudevan and S. Mahadevan

Abstract Localized heating and cooling during welding introduces residual stress in the weld joints due to the thermal gradients. The microstructure formation in 2.25Cr–1Mo steel is very much sensitive to the rate of cooling and eventually, the austenite may transform to pearlite/bainite/martensite. A variant of Tungsten Inert Gas (TIG) welding called Activated-TIG (A-TIG) welding was made use of in this study. Here, we have tried to develop a numerical model based on the finite element method to predict the thermo-mechanical behavior of 6 mm thick 2.25Cr–1Mo steel weld joints, with accounting solid-state phase transformation. The heat energy supplied to the workpiece is determined and calibrated in the heat source model. Simulated thermal cycle showed a peak temperature of 1377 °C at 8 mm distance away from the weld centerline, whereas, the experimental result showed 1371 °C. A *k*-type thermocouple was used to measure the temperature distribution during welding. The simulated thermal distributions were sequentially coupled to mechanical analysis. The evolution of stress and ultimately the locked-in residual stresses were determined. X-ray diffraction studies showed the peak residual stress near heat affected zone of 592 MPa and the vertical height gauge was used to measure the distortion before and after welding. The predicted residual stresses and distortion showed good agreement with the experimental measurements.

Keywords 2.25Cr–1Mo steel · A-TIG welding · FEM · Solid-state phase transformation · Welding residual stress and distortion

A. R. Pavan · B. Arivazhagan · S. Arun Kumar · M. Vasudevan (✉) · S. Mahadevan
Metallurgy and Materials Group, Indira Gandhi Centre for Atomic Research, Kalpakkam, Tamil Nadu, India
e-mail: dev@igcar.gov.in

A. R. Pavan · B. Arivazhagan · M. Vasudevan · S. Mahadevan
Homi Bhabha National Institute, Indira Gandhi Centre for Atomic Research, Kalpakkam, Tamil Nadu, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_21

21.1 Introduction

A higher level of welding residual stress or locked-in stress in the weld joint is a life-limiting factor under service conditions. Steep thermal gradient and shrinkage of sizeable molten metal volume generate tensile stress near the weld joint and is inevitable. This, in turn, leads to stress corrosion cracking, causes brittle fracture and premature failure of engineering components [1]. But, compressive residual stresses are beneficial in certain circumstances [2]. Therefore, by knowing the magnitude of residual stress, one can estimate the remaining life of weld joints with safety measures. However, the locked-in stresses can be controlled and reduced by proper choice of the welding process and use of restraints. Thus, the safety of the structure and design can be increased.

The 2.25Cr–1Mo steel has been most widely used in petrochemical, modern power generation industry in the form of boilers, heaters, heat exchangers, reactors, and hydrocrackers usually built as heavy wall pressure vessels in very high temperature range of 350–550 °C because of its excellent strength at elevated temperatures [3, 4]. The excellent resistance to creep and corrosion, high toughness and good weldability makes it as a candidate material in service conditions. Alongside, these materials have good physical properties such as a low thermal expansion coefficient and a high thermal conductivity. Hence, Cr–Mo steels are promising to operate at high temperatures and under low stresses. Moreover, the integrity of the weld joint decides the performance of any components. Further, the mechanical properties of 2.25Cr–1Mo steel depend on the microstructure, which is sensitive to the cooling rate which in turn depends on heat input during welding [5]. Normally air-cooled steel forms bainitic microstructure at room temperature. This type of steel is usually air cooled after austenitization for service in power plant applications [6]. For 2.25Cr–1Mo joint fabrication, Shielded Metal Arc Welding (SMAW) and TIG welding have been widely adopted [7]. Moreover, TIG has lower productivity while joining the thick (>3 mm) components due to a large number of weld passes utilized to create a joint which eventually introduces high tensile residual stress near the joint with very high distortion. Recently, the variant of TIG welding known as A-TIG welding is widely applied due to its higher productivity [7], low cost, lower micro inclusion, low residual stress and distortion and also because of autogenous welding. An activating flux containing multi-component oxides, applied on the joint alters the surface tension of the weld pool and controls the arc constriction during TIG welding and eventually enhances the depth of penetration up to 300% compared to conventional TIG welding process [8].

However, experimentally assessing the evolution of welding residual stress and optimizing the welding parameters to achieve better weld geometry is time-consuming and costly. Thus, a computational approach is utilized to envisage the thermo-mechanical behavior of the weld joint. The welding thermal cycle, residual stress and distortion could be determined by simulation and appropriate measures or changes in the welding process and conditions could be adopted during

the fabrication to reduce the level of residual stresses and distortion in the actual component.

The computational simulation tool has been used from the past few decades to understand the evolution of welding residual stress and distortion in the welded structures. The commonly used approach is one way coupled approach, viz. sequential coupling of thermal history to mechanical analysis. Various studies [9–13] are available on the prediction of welding residual stress and distortion. Further, the complexity in the numerical model is, to consider the solid-state phase transformations during welding. Many studies have successfully incorporated phase transformation in the numerical model [6, 14, 15]. In addition, very few literatures predicted the welding residual stress and distortion in 2.25Cr–1Mo steel [16–18]. However, there is a limited number of literatures on understanding the evolution of welding residual stress and distortion in A-TIG welded 2.25Cr–1Mo steel joints. Therefore, the motivation of this research work is to predict welding residual stress and distortion by considering solid-state phase transformation in 2.25Cr–1Mo steel weld joint fabricated by autogenous A-TIG welding process by numerical simulation.

21.2 Finite Element Modeling

The dimension of 2.25Cr–1Mo steel plate is 300 × 120 × 6 mm^3. Straight-sided square butt joint configuration was used for A-TIG welding process. Numerical simulations were performed using commercially available SYSWELD software. The optimized meshing is necessary to resemble the exact experimental result in the numerical model. Hence, optimization of mesh size, with the right choice of mesh dimensions is required for improving the accuracy of the simulation. In this work, around 30 mm near to the weld line, fine mesh to a size of 1 mm × 0.8 mm (quadrilateral) was chosen, since the effect of the temperature gradient is more near the weld region and heat affected zone as shown in Fig. 21.1.

Goldak double ellipsoidal heat source model was used for defining a heat source [19]. The thermomechanical analysis is done by sequential coupling of thermal and mechanical analysis. Here the thermal and phase transformation strains are coupled

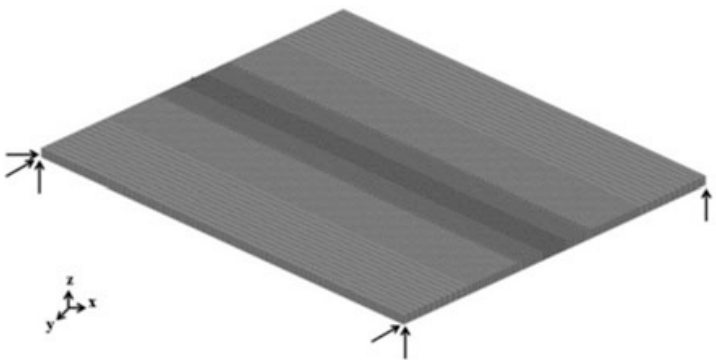


Fig. 21.1 Weld pad model

with non-linear mechanical analysis to understand the evolution of stress and distortion in the weld joint. It was assumed that there is no influence of any structural changes on the thermal analysis part.

21.2.1 Material Properties

The residual stresses and the dimensional change in the weld joint originate owing to the changes in thermo-physical, mechanical properties and microstructural characteristics of materials. Therefore, to get good accuracy in the model, it is necessary to choose appropriate material properties. For 2.25Cr–1Mo steel, the material database was taken from the literature and SYSWELD database [16] as shown in Fig. 21.2. Heat transfer due to convective flow in molten metal is considered in the model by increasing the thermal conductivity value by 2–3 times above the melting temperature.

The heat source fitting tool is used to calibrate the double ellipsoidal heat source and compared with the experimental weld bead macrograph. In addition, the experimental and predicted thermal profiles were compared to know significant heat supply into the material. Further, the obtained thermal stains are sequentially coupled to the mechanical analysis.

21.2.2 Governing Equations

Welding introduces significant heat energy in a localized region. Thus, the heat transfer during welding is primarily due to conduction mode. Heat conduction can be described as shown in the following Eq. (21.1).

$$\nabla(k\nabla T) + Q = \rho C(\partial T/\partial t) \tag{21.1}$$

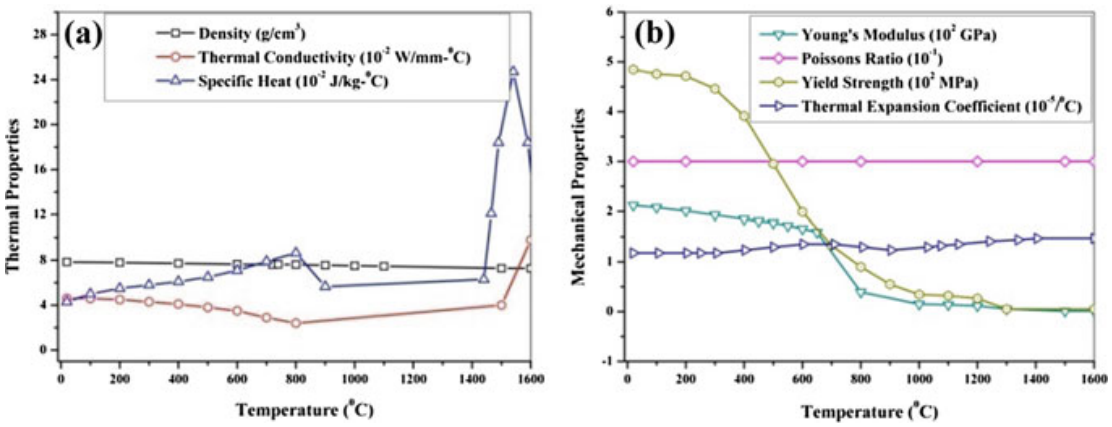


Fig. 21.2 2.25Cr–1Mo steel material properties **a** thermal physical and **b** mechanical properties

where k is the thermal conductivity (W/m °C), Q is the volumetric heat flux (W/m^3), ρ is the density (kg/m^3), and C is the specific heat (J/kg °C).

After welding, the heated material starts to cool to ambient temperature due to convection and radiation. The rate of convection heat transfer is governed by Newton's law of cooling as Eq. (21.2).

$$Q_c - h_c(T_S - T_0) = 0 \tag{21.2}$$

From, Stefan–Boltzmann law the radiation heat transfer is expressed as Eq. (21.3),

$$Q_r - \varepsilon\sigma(T_S^4 - T_0^4) = 0 \tag{21.3}$$

where h_c is the film coefficient for convection [$\sim$25 W/(m^2 K)], T_S is the surface temperature (K), T_0 is the ambient temperature (K), ε is the emissivity (0.8), and σ is the Stefan–Boltzmann constant (5.67 $\times$ 10^{-8} W/m^2 K^4). The total heat transfer can be written as,

$$Q_t = Q_c + Q_r \tag{21.4}$$

$$Q_t = h_c(T_S - T_0) + \varepsilon\sigma\,(T_S^4 - T_0^4) \tag{21.5}$$

By solving Eq. (21.5)

$$\left.\begin{aligned} Q_t &= h_c(T_S - T_0) + \varepsilon\sigma(T_S - T_0)(T_S + T_0)(T_S^2 + T_0^2) \\ Q_t &= [h_c + \varepsilon\sigma(T_S + T_0)(T_S^2 + T_0^2)](T_S - T_0) \end{aligned}\right\}$$

$$h_{eq} = h_c + \varepsilon\sigma(T_S + T_0)\,(T_S^2 + T_0^2) \tag{21.6}$$

$$Q_t = h_{eq}(T_S - T_0) \tag{21.7}$$

where Q_t is the total surface heat transfer and h_{eq} is the overall heat transfer coefficient. To account the phase transformation from parent phase to austenite and austenite to bainite phase transformation in the analysis, Leblond's equation and Johnson–Mehl–Avrami (JMA) were used (Eqs. 21.8 and 21.9).

$$dz/dt = \big(z_{eq}(T) - z/\tau(T)\big)f(\dot{T}) \tag{21.8}$$

$$p = 1 - \exp(-kt^n) \tag{21.9}$$

where z_{eq} refers to austenite proportion at phase equilibrium, z is austenite proportion, p is the fraction of bainite, k and n are kinetic parameters and t is time. The parameters of the JMA equation should be obtained from the Continuous Cooling Transformation (CCT) diagram. Here, we have assumed in the model that there is no formation of martensite upon cooling in the weld metal.

In addition, after the completion of the thermal analysis, by solving Eqs. (21.1) and (21.7) the obtained temperature distributions are coupled to a mechanical analysis by using Eqs. (21.10) and (21.11). In simple terms, the thermal analysis is sequentially coupled to mechanical analysis. Hence, the possible total strain induced on the material during welding is by elastic, thermal, plastic and due to phase transformation as shown in Eq. (21.10).

$$\Delta\varepsilon^{\text{total}} = \Delta\varepsilon^{\text{elastic}} + \Delta\varepsilon^{\text{thermal}} + \Delta\varepsilon^{\text{plastic}} + \Delta\varepsilon^{\text{phase}} \tag{21.10}$$

For, mechanical analysis, the force equilibrium is given as Eq. (21.11)

$$\sigma_{ij} + p_j = 0 \tag{21.11}$$

where σ_{ij} is stress tensor and p_j is body force at any point within the boundary. For the finite element formulation, the basic equilibrium equation can be used to achieve the virtual induced strain and displacement.

21.2.3 Boundary Condition

During the experiment, the weld joint was clamped to avoid vertical movement. Therefore, to match with experimental conditions, the apt external constraints were provided in the model. The rigid clamps were made use of to avoid rotational motion and translational motion of the weld joint. The elastic clamp with a finite stiffness of 10,000 N/mm is used to achieve the proper deformation. Isotropic hardening model was used to consider elastic-plastic material behavior.

21.3 Experimental

As received 2.25Cr–1Mo plates with normalized at 950 °C for 20 min and tempered at 730 °C for 60 min was used with dimension 300 mm × 120 mm × 6 mm having a straight square but joint configuration. Weld plate is fabricated by using A-TIG welding with heat input 2.38 kJ/mm. The transient temperature during the welding process was captured by Chromel-Alumel thermocouples placed 8 mm away from the center of the weld as shown in Fig. 21.3. Vertical height gauge was used to obtain the difference in height of the plate before and after welding (least count 0.006 mm) to estimate the level of distortion.

X-ray diffraction was utilized to estimate the magnitude of welding residual stress with Cr–Kα radiation. The $\sin^2\psi$ method was performed to evaluate the magnitude of residual strain and later strain is converted into stress. Initially, the

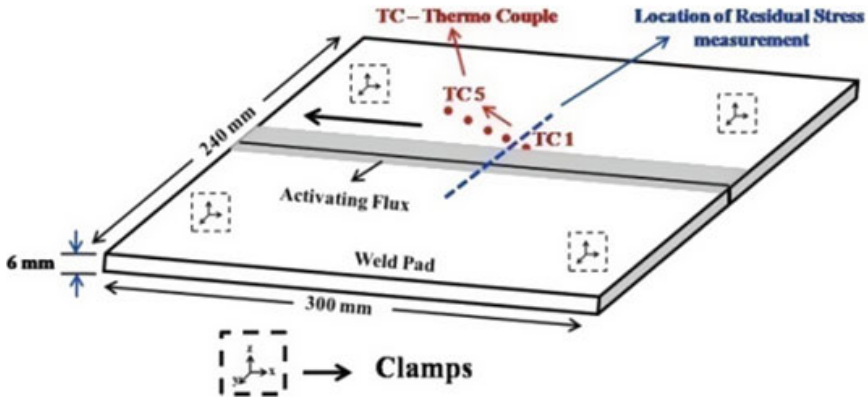


Fig. 21.3 Weld pad configuration

change in d-spacing is characterized at different ψ angles. The slope of a linear plot of d-spacing versus $\sin^2\psi$ gives the residual strain in the material, where, d represents the lattice spacing and ψ is the angle between the surface normal and the bisector of the incident and diffracted beam which is normally varied in the angular range of 0° to 45°.

21.4 Results and Discussion

21.4.1 Thermal Analysis

The thermo-mechanical simulation was carried out by utilizing FEM based SYSWELD software. The heat source fitting tool is employed to calibrate the heat source by comparing with the experimental weld macrograph to ensure the significant amount of heat energy supplied into the material (Fig. 21.4). Initially, heat energy supplied into the weld joint was calibrated using double ellipsoidal heat

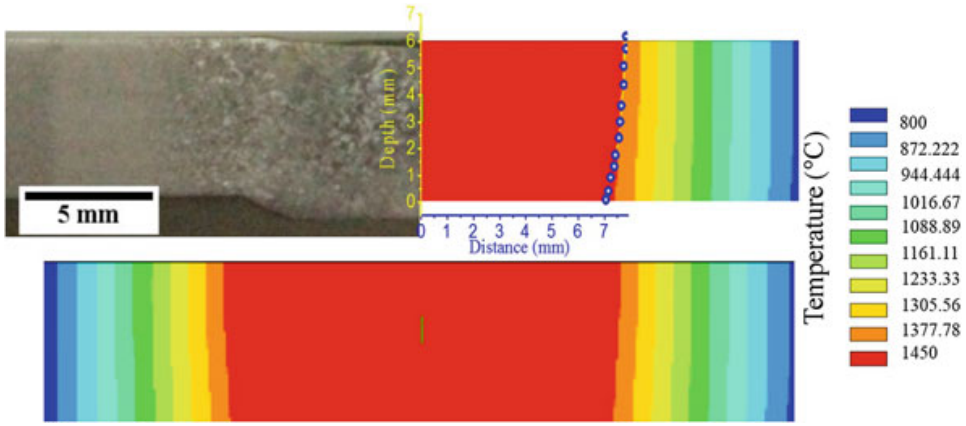


Fig. 21.4 Heat source fitting

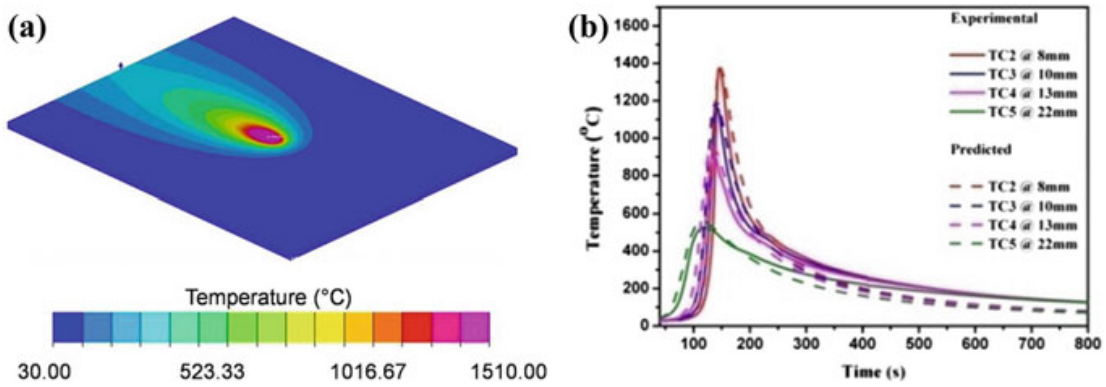


Fig. 21.5 **a** Thermal analysis of weld plate **b** thermal cycles near the weld center

source function parameters as shown in Eqs. (21.12) and (21.13) and fitted with the experimental result as shown in Fig. 21.5b.

Heat input, $Q = \left(\frac{I \times V \times \eta}{v}\right)$

$$q_f(x,y,z,t) = \frac{6\sqrt{3}f_r Q}{a_f bc\pi\sqrt{\pi}} e^{-3x^2/a_f^2} e^{-3y^2/b^2} e^{-3[z+v(\tau-t)]^2/c^2} \tag{21.12}$$

$$q_r(x,y,z,t) = \frac{6\sqrt{3}f_r Q}{a_r bc\pi\sqrt{\pi}} e^{-3x^2/a_r^2} e^{-3y^2/b^2} e^{-3[z+v(\tau-t)]^2/c^2} \tag{21.13}$$

I is the current, V is voltage, η is the efficiency of the welding process and v is the torch velocity in m/s. Efficiency is assumed to be 80% for A-TIG welding. Where a_f and a_r is the front and rear length of the weld pool, respectively, c is the weld penetration depth and b is half width of the weld bead. f_f and f_r ($f_r + f_f = 2$) is the front and rear fractions of the double ellipsoidal heat source, which is considered as 0.4 and 1.6 respectively. q_r and q_f are power densities in W/m^3. The first thermocouple fixed near (8 mm) to the weld bead shows highest peak temperature of 1371 °C and the fifth thermocouple is fixed away (22 mm) from the center of the weld showed lowest peak temperature of 520 °C. This change in peak temperature shows the Gaussian heat distribution, where, the heat flux density is high at the center and gradually decreases away from the center.

During the welding of 2.25Cr–1Mo steel, the microstructures in the weld zone and near the heat affected zone alters according to the applied heat input and mainly due to cooling rate. It is well known that phase transformation affects the magnitude of residual stress due to volume change (body-centered cubic structure to face-centered cubic structure). However, the magnitude of residual stress depends on the yield strength of the material and in turn, yield strength, directly depends on the microstructure of the weld material. Hence it is important to consider yield strength and phase transformation effect on the numerical model for good accuracy. Figure 21.6 shows the granular bainitic microstructure of base metal and bainitic lath structure of as-welded sample; this displayed the typical characteristics of the

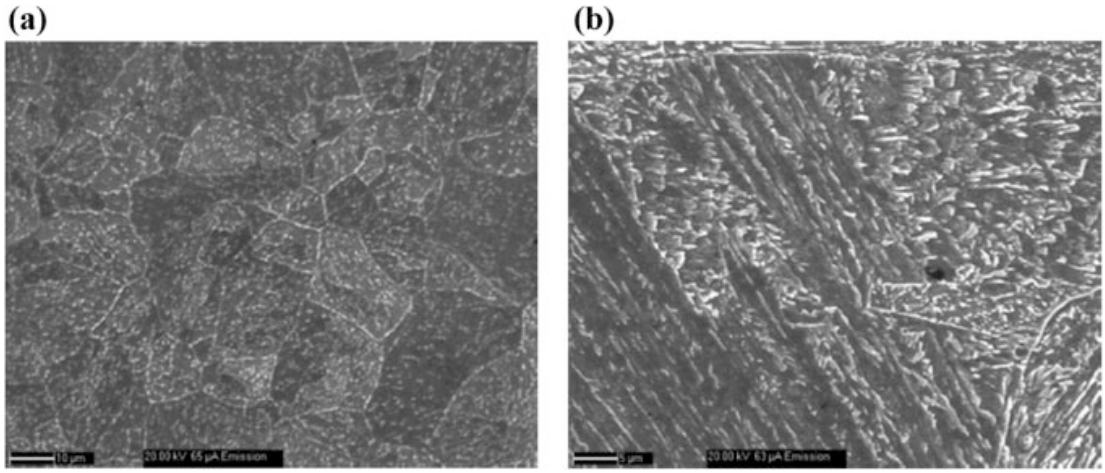


Fig. 21.6 Microstructure **a** base metal **b** weld metal

bainitic microstructure. To consider ferrite/bainite to austenite and austenite to bainite phase transformation, Eqs. (21.4) and (21.5) were used in the model respectively.

21.4.2 Mechanical Analysis

The temperature histories were then sequentially coupled for the mechanical analysis to determine the level of welding residual stress, shrinkage effect and distortion. The strains produced by the thermal energy are coupled to mechanical analysis. The thermal strains, plastic and elastic strains along with volume strains produced due to solid-state phase transformation effect, induce stress in the material. To consider all these factors in the numerical model, temperature dependent material properties must be incorporated, that is, Young's modulus and Poisson's ratio was taken into account for elastic strains. The coefficient of thermal expansion was used for achieving thermal strains. For plastic and phase transformation induced strains, temperature dependent stain hardening and change in volume viz., phase transformation properties were considered in the material database, respectively. A dip in the residual stress distribution across the fusion zone indicates volume relaxation of the weld metal due to phase transformation effect as shown in the contours (Fig. 21.7a, b) of longitudinal residual stress in 2.25Cr–1Mo weld joints.

Figure 21.7 shows the comparison of measured and predicted longitudinal welding residual stress profiles at the mid-cross-section of the weld joint. Measured peak tensile residual stress is 592 MPa while the predicted value was 568 MPa. The weld-induced residual stress has high magnitude at heat affected zone (HAZ) and at the weld center; there was a relaxation in the stress value. This is mainly due to localized heating and cooling, the only a small area near the joint is affected by the

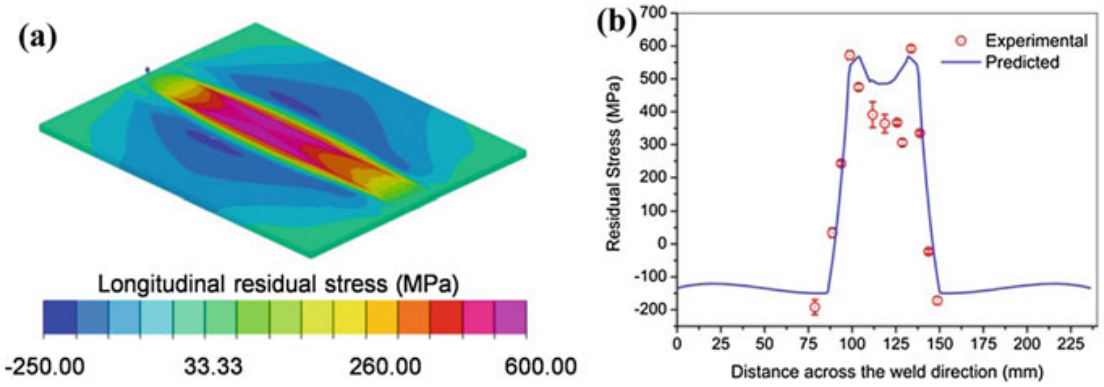


Fig. 21.7 Longitudinal residual stress **a** contour image **b** validation

intense heat source which leads to the steep thermal gradient. Hence, the contraction of the molten metal pool is hindered by the cooler region. Therefore, usually in a weld metal peak tensile stress near the interface (heat affected zone) can be found (Fig. 21.7b). Further, volume change during phase transformation relieves the built-up stress and eventually leads to the stress relaxation at the weld center. The predicted and measured tensile residual stresses near the vicinity of the joint are of the order of yield strength of the material. Thus, the excess generation of stress will drive the deformation in the structure, called distortion (Fig. 21.8a, b).

The formation of stress and eventually the residual stress during heating and cooling cycles at a spot can be understood from Fig. 21.9a, b. As the heating starts, compressive stress increases to a peak value and the stress is limited to materials yield strength. Further, compressive stress starts to decrease as the metal starts to soften. Eventually drops down to zero at the melting point of the material. During cooling, tensile stress starts to build slowly due to thermal strains and at the bainitic start temperature (Bs (T)) there was small stress relieving effect due to the phase transformation from austenite to bainite and also due to change in yield strength. For 2.25Cr–1Mo steel, the Bs (T) varies from 490 to 520 °C. Also, it can be predicted by using the below equation [7].

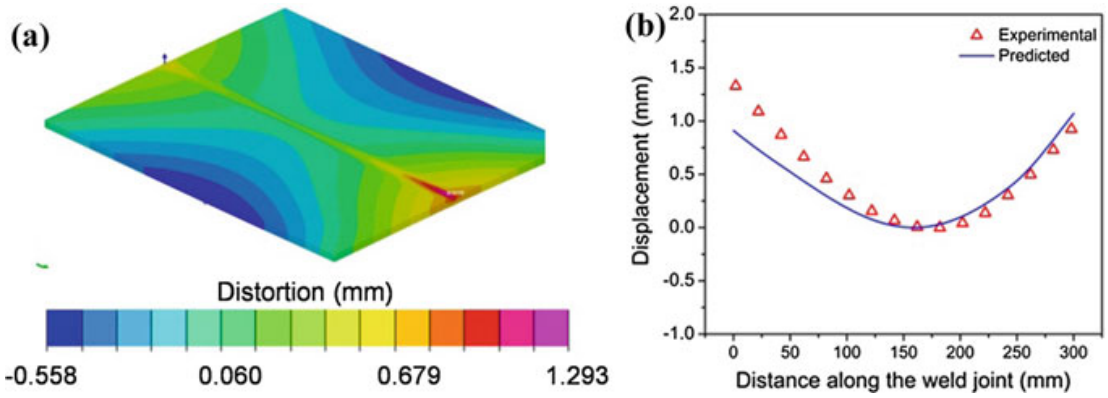


Fig. 21.8 Distortion **a** contour image **b** validation

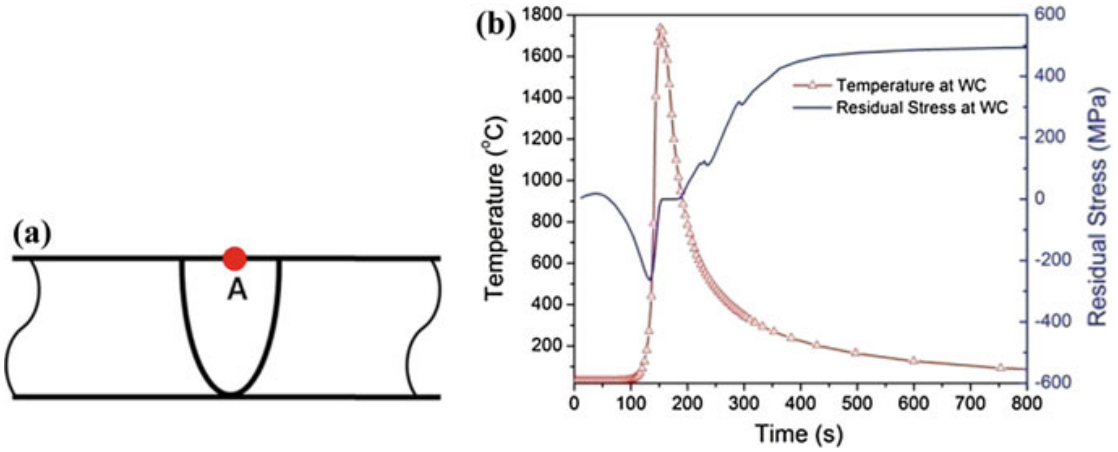


Fig. 21.9 **a** Schematic sketch showing the location of a spot for predicting stress evolution. **b** Evolution of stress during welding thermal cycle at a spot

$$\mathrm{Bs}(^{\circ}\mathrm{C}) = 630 - 45\,\mathrm{Mn} - 40\,\mathrm{V} - 35\,\mathrm{Si} - 30\,\mathrm{Cr} - 25\,\mathrm{Mo} - 20\,\mathrm{Ni} - 15\,\mathrm{W}. \tag{21.14}$$

Meanwhile, cooling beyond bainitic start temperature causes the generation of very high tensile stress because of the dominating effect of thermal strains over volume strain. Eventually, the tensile stress reaches the maximum value of the order of yield strength of base/weld metal.

21.5 Conclusions

Thermo-Mechanical analysis of A-TIG welding of 2.25Cr–1Mo steel has been successfully carried out by the numerical model. The predictions of the numerical model are validated by experimental tools. The following conclusions are drawn from this study:

1. The peak temperature at 8 mm distance was measured as 1371 °C using thermocouples and the thermal cycle exhibited a sharp temperature gradient. Numerical prediction showed 1377 °C at the same spot and therefore, there was good agreement between the predicted and measured temperatures.
2. The predicted and the measured peak tensile residual stress were found to be 592 and 568 MPa respectively near the heat affected region. The welding residual stress profile across the weld joint displayed a typical “M” shape curve which implied that austenite to bainitic phase transformation had an influence on the final residual stress.

3. The observed microstructure at room temperature showed typical bainitic microstructure. Also, the model successfully predicted the distortion with good accuracy. The predicted thermal cycle, residual stress and distortion showed good agreement with the experimentally measured values.

References

1. Mochizuki, M.: Control of welding residual stress for ensuring integrity against fatigue and stress–corrosion cracking. Nucl. Eng. Des. **237**(2), 107–123 (2007)
2. Vasantharaja, P., Maduarimuthu, V., Vasudevan, M., Palanichamy, P.: Assessment of residual stresses and distortion in stainless steel weld joints. Mater. Manuf. Processes **27**(12), 1376–1381 (2012)
3. Na, H., Lee, S., Kang, C.: Effect of micro-segregation on impact toughness of 2.25 Cr-1Mo steel after post weld heat treatment. Metals **8**(6), 373 (2018)
4. Trautwein, A., Mayer, H., Gysel, W., Walser, B.: Structure and mechanical properties of 2¼Cr-1Mo cast steel for pressure components with wall thicknesses up to 500 mm. In: Application of 2¼Cr-1 Mo Steel for Thick-Wall Pressure Vessels, ASTM International (1982)
5. Gojic, M., Kosec, L., Matkovic, P.: The effect of tempering temperature on mechanical properties and microstructure of low alloy Cr and CrMo steel. J. Mater. Sci. **33**(2), 395–403 (1998)
6. Wada, T., Eldis, G.: Transformation characteristics of 2¼Cr-1Mo steel. In: Application of 2¼Cr-1 Mo Steel for Thick-Wall Pressure Vessels, ASTM International (1982)
7. Arivazhagan, B., Vasudevan, M.: Studies on A-TIG welding of 2.25 Cr-1Mo (P22) steel. J. Manuf. Processes **18**, 55–59 (2015)
8. Mills, K., Keene, B., Brooks, R., Shirali, A.: Marangoni effects in welding. Philos. Trans. R. Soc. Lond., Ser. A Math. Phys. Eng. Sci. 911–926 (1998)
9. Friedman, E.: Thermomechanical analysis of the welding process using the finite element method. J. Pressure Vessel Technol. **97**(3), 206–213 (1975)
10. Lindgren, L.-E.: Numerical modelling of welding. Comput. Methods Appl. Mech. Eng. **195** (48–49), 6710–6736 (2006)
11. Hyde, T., Yaghi, A., Tanner, D., Bennett, C., Becker, A., Williams, E., Sun, W.: Current capabilities of the thermo-mechanical modelling of welding processes. J. Multiscale Model. **1**, 451–478 (2009)
12. Deshpande, A., Xu, L., Sun, W., McCartney, D., Hyde, T.: Finite-element-based parametric study on welding-induced distortion of TIG-welded stainless steel 304 sheets. J. Strain Anal. Eng. Des. **46**(4), 267–279 (2011)
13. Lindgren, L.-E.: Computational Welding Mechanics. Elsevier (2014)
14. Leblond, J., Devaux, J.: A new kinetic model for anisothermal metallurgical transformations in steels including effect of austenite grain size. Acta Metall. **32**(1), 137–146 (1984)
15. Heinze, C., Pittner, A., Rethmeier, M., Babu, S.: Dependency of martensite start temperature on prior austenite grain size and its influence on welding-induced residual stresses. Comput. Mater. Sci. **69**, 251–260 (2013)
16. Deng, D., Murakawa, H.: Finite element analysis of temperature field, microstructure and residual stress in multi-pass butt-welded 2.25 Cr–1Mo steel pipes. Comput. Mater. Sci. **43**(4), 681–695 (2008)

17. Vakili-Tahami, F., Daei-Sorkhabi, A., Saeimi-S, M., Homayounfar, A.: 3D finite element analysis of the residual stresses in butt-welded plates with modeling of the electrode-movement. J. Zhejiang Univ. Sci. A **10**(1), 37–43 (2009)
18. Deng, D., Tong, Y., Ma, N., Murakawa, H.: Prediction of the residual welding stress in 2.25 Cr-1Mo steel by taking into account the effect of the solid-state phase transformations. Acta Metall. Sin. (English Letters) **26**(3), 333–339 (2013)
19. Goldak, J., Chakravarti, A., Bibby, M.: A new finite element model for welding heat sources. Metall. Trans. B **15**(2), 299–305 (1984)

Chapter 22
Finite Element Modeling of Hybrid Laser-TIG Welding of Type 316L(N) Stainless Steel

M. Ragavendran, M. Vasudevan and M. Menaka

Abstract In the present study, thermomechanical analysis of autogenous hybrid laser-TIG welding of type 316L(N) austenitic stainless steel has been carried out. It has been realized that the residual stress due to welding plays a vital role in the mechanical properties of the joint and its performance during service. Therefore, it is important to study the residual stress of type 316L(N) stainless steel weld joint produced by hybrid welding process. First, the thermal analysis was carried using a hybrid heat source consisting of conical and double ellipsoidal models. The heat source was calibrated by comparing the simulated weld bead profile with that of the experimentally obtained weld bead profile. Then, the obtained temperature distribution was sequentially coupled to the mechanical analysis. The simulated thermal cycle was validated by temperature measurements using noncontact real-time IR thermography. The simulated residual profile was validated by an ultrasonic technique employing critically refracted longitudinal waves. The ferrite content of the weld metal was measured using ferrite scope. Distortion in the weld joint is measured using digital vertical height gauge. Optical microscopy is employed for microstructural characterization of the weld joint. The weld metal exhibited a peak tensile residual stress value of 320 MPa. The predicted distortion values were very low. There was a good agreement found between the predicted and experimentally measured thermal cycles, residual stresses, and distortion.

Keywords Hybrid laser-TIG welding · 316(L)N stainless steels · Thermomechanical analysis

M. Ragavendran (✉)
Homi Bhabha National Institute, Indira Gandhi Centre for Atomic Research, Kalpakkam, Tamil Nadu, India
e-mail: mragavendran@igcar.gov.in

M. Vasudevan
Metallurgy and Materials Group, Indira Gandhi Centre for Atomic Research, Kalpakkam, Tamil Nadu, India

M. Menaka
Health Safety and Environmental Group, Indira Gandhi Centre for Atomic Research, Kalpakkam, Tamil Nadu, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_22

22.1 Introduction

Type 316L(N) austenitic stainless steel is the main structural material used in the construction of nuclear reactors. It is the preferred nuclear grade structural material for reactor due to its good mechanical properties, corrosion, and irradiation resistance [1]. Welding is the predominant fabrication method employed during their construction. Therefore, the utilization of an advanced welding process for welding of type 316L(N) stainless steel is of high importance. Laser welding has high-energy density and the advantages are lower residual stress and distortion that are due to low heat input, high welding speeds, and ease of automation. But for welding high-thickness plates using a laser, keyhole instability occurs which produces a lot of porosity in the welds. Arc welding is most widely used, but it has a lot of limitations namely high residual stress and distortion due to high heat input. Slow welding speed and need of groove preparation limit the productivity.

For welding, heavier components, an advanced welding process such as hybrid laser-TIG welding process could be used. In our earlier study, hybrid laser-TIG welding process parameters were optimized for welding of type 316L(N) austenitic stainless steel using RSM and GA [2, 3]. The advantages of hybrid welding are good gap bridging ability and enhanced productivity; high penetration capability with faster welding speeds due to more energy delivered into the material [4]. The process improves the weld quality with reduction of porosity and cracks due to slow cooling rate and enhances ductility [5]. The accurate measurement of temperature during hybrid welding is difficult using conventional thermocouples. It cannot measure the weld pool temperature; it will measure only the nearby base metal temperature. Therefore, in this study, these limitations will be overcome by using noncontact real-time IR thermography for precise measurement of temperature. It is also employed for authenticating, fine-tuning, and validating the FEM model.

The novelty in this investigation was the use of IR thermography to measure the temperature distribution in the weld pool and its surroundings to validate the FEM predictions. Further, the thermal model was sequentially coupled to the mechanical analysis for determination of residual stress. The predictions of the numerical analysis presented in this investigation on the temperature distribution, residual stress, and distortion were confirmed by measurements by the use of an experimental tool.

22.2 Experimental

The 316L(N) plates of size 300 mm × 125 mm × 5.6 mm were used in straight square butt joint configuration. The setup is given in Fig. 22.1. The laser-TIG hybrid welding was carried out with the determined optimal parameters with a heat input of 0.466 kJ/mm. Weld joint was made autogenously. After the weld joint was fabricated, the residual stress studies were made by utilizing an ultrasonic technique

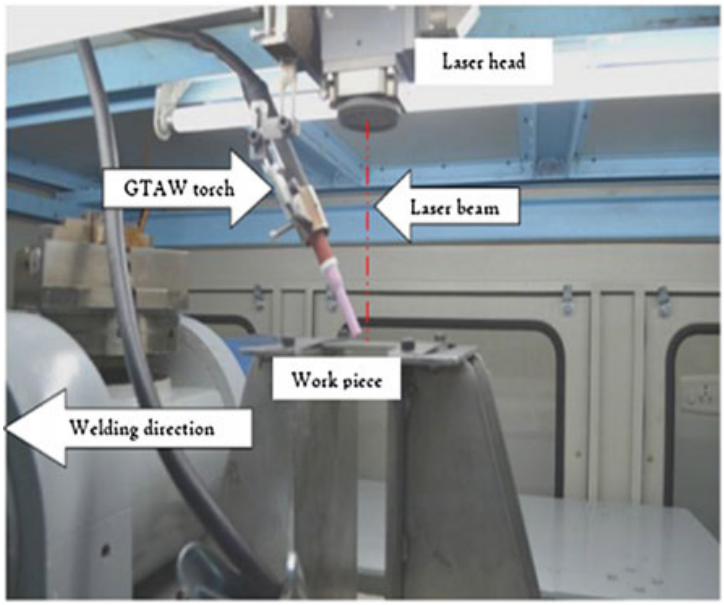


Fig. 22.1 Experimental setup

Table 22.1 Laser welding heat source specification

TRUMPF laser welding specification	
Max average power	530 W
Pulse power	0.5–10 kW
Pulse duration	0.3–50 ms
Max pulse energy	60–100 J
Max rep rate	0.1–833 Hz
λ of laser light	1064 nm

using L_{CR} waves. The weld joint was cross-sectioned, polished, and etched electrolytically using equal amounts of concentrated nitric acid and distilled water. The temperature was measured in real time using IR thermography. Distortion was measured using digital height gauge. Tables 22.1 and 22.2 show the heat source

Table 22.2 Welding parameters

Welding parameters	
Leading heat source	TIG
Welding speed	20 cm/min
Defocusing	−2 mm
Gap between heat sources	3 mm
Torch angle	45°
Arc gap	1 mm
Shielding gas/flow rate	Argon, 10 L/min
Power	3750 W
Pulse frequency	12 Hz
Pulse duration	10 ms
TIG current	120 A

Fig. 22.2 Weld joint fabricated by hybrid laser-TIG welding process

details and welding parameters used in hybrid laser welding. Figure 22.2 shows the joint fabricated by the hybrid welding process.

22.3 Numerical Modeling

The thermal analysis of hybrid laser-TIG welding simulation was executed using commercially available SYSWELD software which is exclusively used for the modeling and simulation of fusion welding processes. The hybrid heat source model consisting of double ellipsoid for the arc welding and three-dimensional conical for the laser welding is used in the analysis. The geometrical parameters of the heat source models are taken from the experimental weld bead. Figure 22.3 schematically represents the combined heat source model employed in this study.

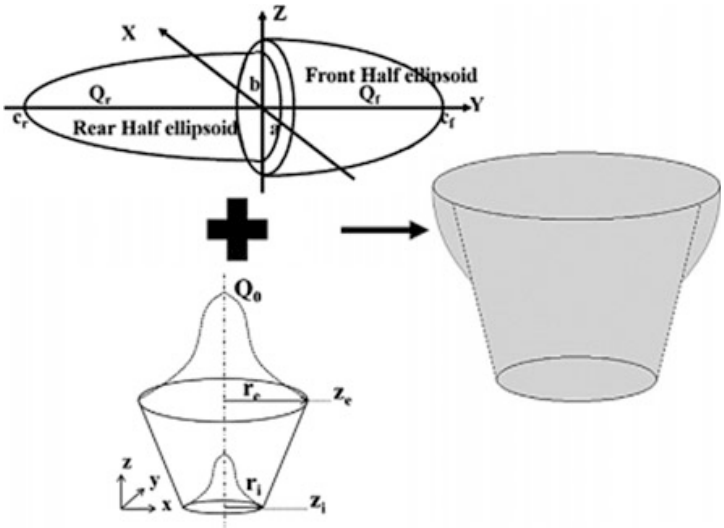


Fig. 22.3 Coupled double ellipsoidal with three-dimensional conical heat source model

22.3.1 Governing Equations

The heat transfer due to confined heating while welding is mainly due to conduction mode. The amount of heat conduction relies on the materials thermophysical properties, the amount of weld metal volume produced, and weld morphology. Heat conduction defined by finding the solution for the following equation.

$$\nabla(k\nabla T) + Q(x, y, z, t) = \rho \mathrm{Cp}(\partial T/\partial t) \tag{22.1}$$

where Q = Heat flux volume (W m^{-3}), k = Thermal conductivity (W m^{-1} °C^{-1}), ρ = Density (kg m^{-3}), Cp = Specific heat capacity (J kg^{-1} °C^{-1}), and T = Transient temperature (°C), t = time (s). The heat loss due to radiation and conduction throughout the welding is given by Stefan–Boltzmann relation and Newton's law of cooling.

22.3.2 Boundary Condition

The hybrid laser-TIG hybrid was conducted using various welding parameters. The initial temperature was 30 °C. All the heat losses namely radiation heat loss which is predominant in the weld pool, convective and conductive heat losses that are prevalent away from the weld pool are taken into account. In the model, convective and radiative losses were considered on all its surfaces except the symmetry plane. Similar to the experimental conditions, the plates to be welded are clamped appropriately to avoid the translational and rotational movements.

22.3.3 Finite Element Modeling

The dimension of the weld plate used in finite element modeling was 300 × 125 × 5.6 mm. The butt joint configuration with a square edge preparation, without a groove, was used for welding. To improve the precision and effectiveness of the simulation, optimum mesh size should be used. The effect of heat input due to welding will be more in the weld center and heat-affected zone (HAZ). So to get the accurate weld bead morphology, finer mesh size was used in these zones and the other regions are given little coarser mesh to reduce the computation time. The heat source moves in Y-axis. The model was formulated such as the welding direction was along the Y-axis and the model was about the XY plane. The thermophysical properties of the type 316L(N) austenitic stainless steel were considered from room temperature to melting point are taken from IGCAR databook [6].

22.4 Results and Discussion

22.4.1 Temperature Distribution and Weld Bead Comparison

Temperature distribution in the transverse direction is plotted from the weld centerline. The induced peak temperature is almost the same near the regions next to the weldment. The distance far away from the weldment is not affected by two heat sources. The peak temperature changes if the heat source traverses in its path throughout the welding. All the nodes are heated till its peak temperature because the heat source strikes a corresponding node and it begins to cool when it travels further. The profiles are obtained in the quasi-state condition. The weld bead produced by the hybrid welding process is perfectly matching with the predicted finite element model using SYSWELD. If the predicted weld bead geometry is matching well with experiments, then the results given by the predicted model are reliable and almost accurate. Figure 22.4 shows the weld macrograph which is compared with the FEM model.

The thermal heat source profiles of IR thermography are compared with FEM model. Temperature profile matches well in the quasi-state condition. There was a good agreement found within the measurements and the predicted temperature using the model. The peak temperature measured also almost found equal to that of the predicted by the model. The variations seen may be due to assumptions during the material modeling beyond the melting point of the material and experimental constraints. These deviations are reduced by improving the numerical modeling analysis. Figure 22.5 shows the IR temperature distribution profile which is compared with that of the FEM model. Figure 22.6a presents the numerical modeling results, in which the induced peak temperature due to welding which keeps on

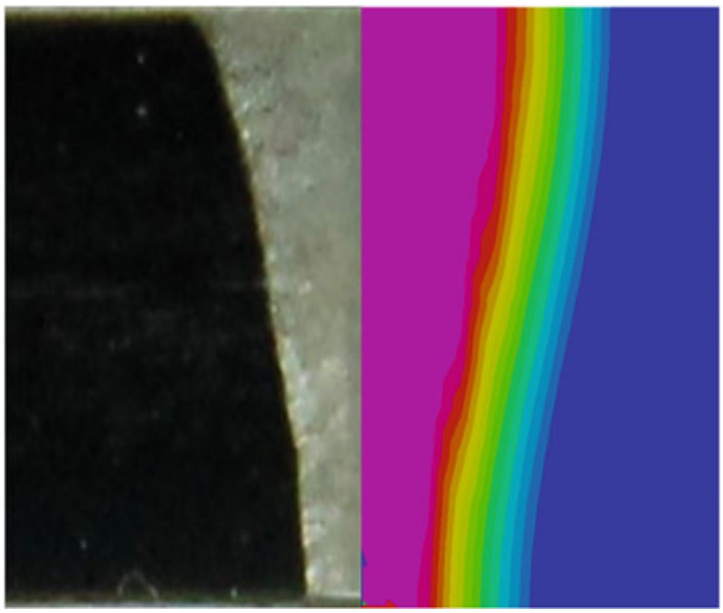


Fig. 22.4 Weld macrograph compared with FEM model

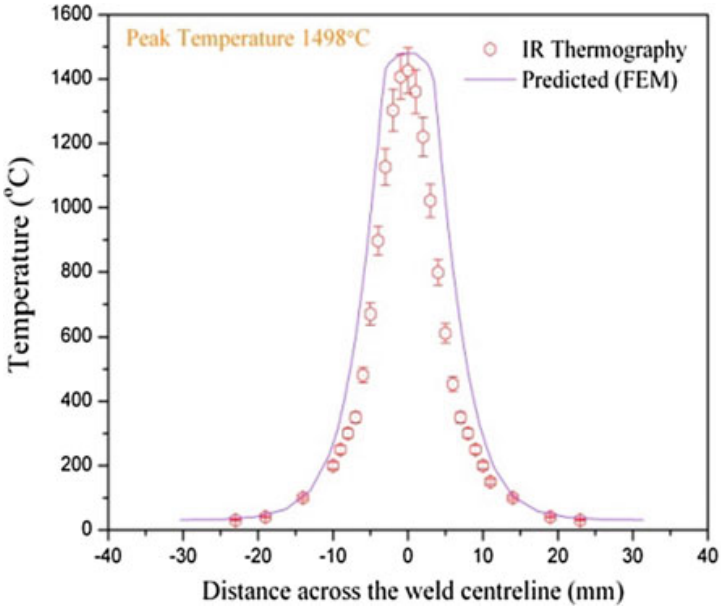


Fig. 22.5 Temperature measured by IR thermography compared with FEM at 50 s

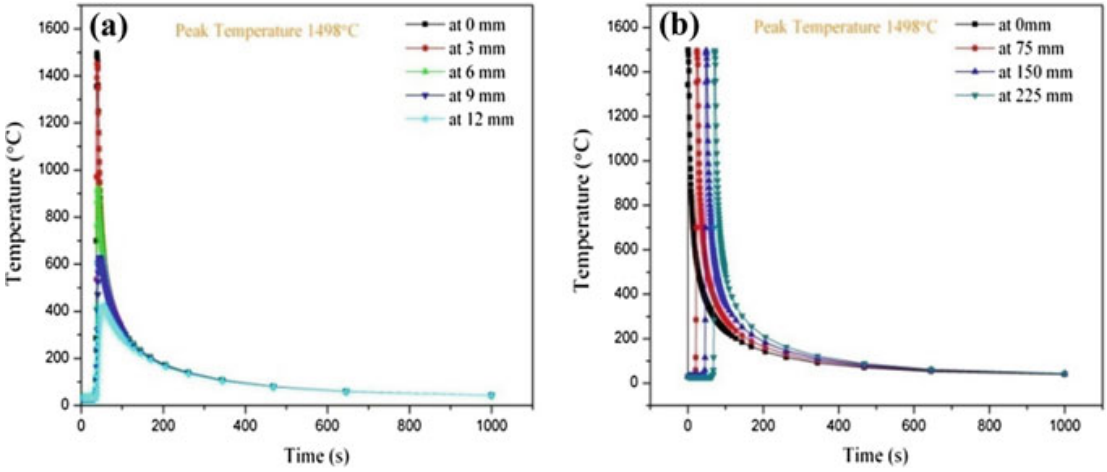


Fig. 22.6 Thermal cycle from welding centerline: **a** transverse direction **b** longitudinal direction

decreasing in the transverse direction. Thermal cycle along the weld center line for various locations is shown in Fig. 22.6b.

22.4.2 Residual Stress Analysis

The residual stress measurement was conducted by employing ultrasonic testing. The predicted residual stress profiles were confirmed by longitudinal residual stress measurements taken at the center of the weld joints. The residual stress values evaluated are compared with the model and it is presented in Fig. 22.7a.

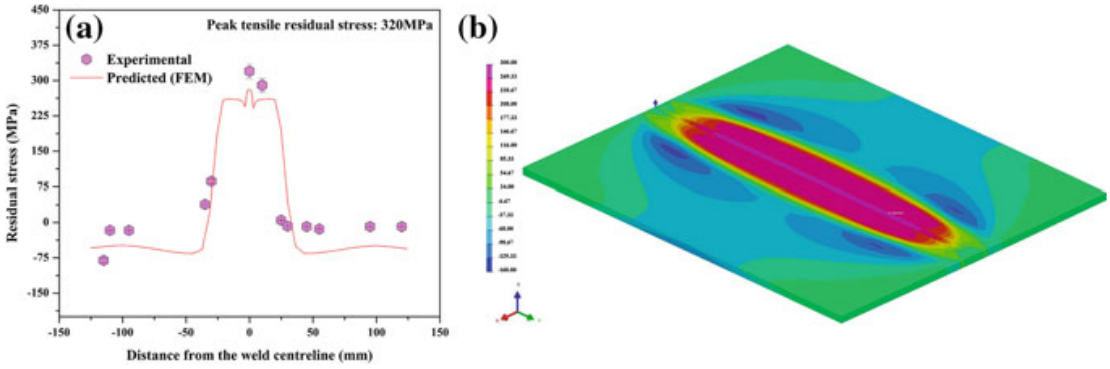


Fig. 22.7 Residual stress **a** comparison of FEM versus ultrasonic testing **b** distribution in welded plate

Figure 22.7b shows the residual stress distribution in the welded plate. The peak tensile stress is at the weld center and the value is around 280 MPa as predicted by the model. The maximum tensile stress measured by the ultrasonic technique is 320 MPa. The flat region observed next to the peak tensile residual stress is due to the autogenous hybrid heat sources. That is, if the filler wire is used, heat input will be utilized for melting the filler wire. But in the case of autogenous welding, the heat input is spread over a more area next to weld metal. Moreover, the type 316L (N) austenitic stainless steel has low thermal conductivity, so it retains high temperature over a more extended period due to which the residual stress profile is flat near to the weldment. There was a good agreement found within the predicted model and the experimentally calculated residual stress.

The evolution of stress and finally the residual stress during heating and cooling cycles due to welding can be explained by spot analysis given in Fig. 22.8a, b as a function of time. When welding starts, compressive stress rises to a peak value and the stress is restricted to materials yield strength. Then, the compressive stress begins to decline as the metal commences softening. Finally, stress falls to zero at the melting point of the base metal. During cooling, the tensile stress rises to

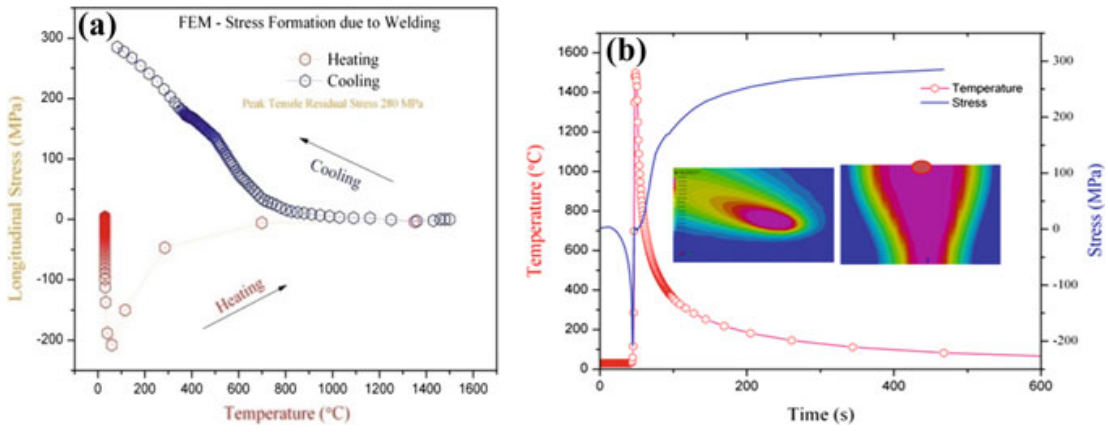


Fig. 22.8 **a** Stress evolution during welding—spot analysis **b** stress evolution as a function of time

develop slowly due to the thermal strains. Due to welding, there is no phase transformation effect in austenitic stainless steel. Henceforth, no signature dips or stress relaxation changes in the temperature versus stress plot on cooling side.

22.4.3 Distortion Analysis

The distortion in weld joints is caused by the heat input amount and the associated weld metal volume. Since the welding parameters are employed in such a way to produce low weld metal volume with full penetration without any defects, the distortion is less in the joint fabricated by the hybrid welding process. The heating and melting efficiency are improved in hybrid welding which results in lower weld metal volume. Therefore, low distortion is accompanied with lower heat input. Distortion is too low because of the weld metal volume produced by the two heat sources is limited when compared with arc welding alone which is possessing a higher heat input. The predicted model values are found to agree with the experimental values. Figure 22.9a exhibits the distortion comparison of the FEM model with the experimental data. Figure 22.9b shows the distortion in the welded plate.

22.4.4 Microstructure

The weld joint produced by hybrid laser-TIG welding process shows a primary austenitic solidification mode. Weld metal consists of the fewer amount of delta ferrite (0.9FN) and austenite phases. The hybrid welding process has higher welding speed which increases the cooling rate that results in primary austenitic solidification mode with the low ferrite number [7]. Figure 22.10a shows the weld macrostructure and Fig. 22.10b shows microstructure of the weld. The weld macrostructure exhibited a typical wine glass shape. Because the TIG heat source

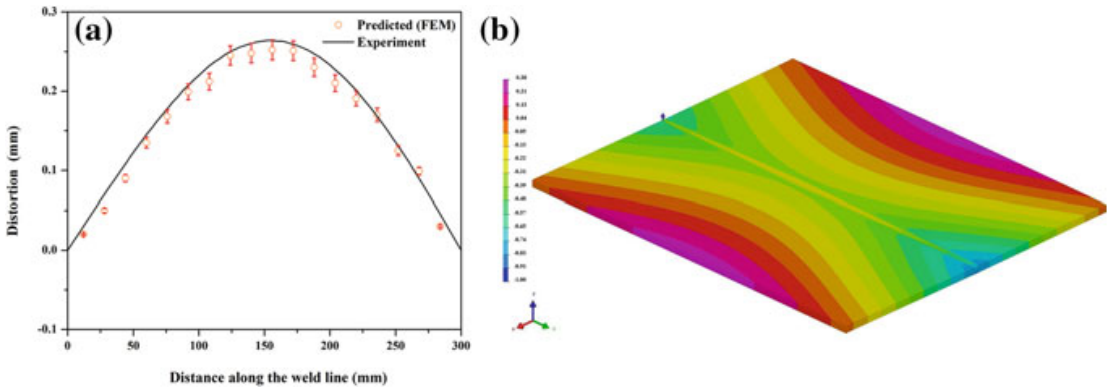


Fig. 22.9 Distortion **a** comparison of FEM with experiment **b** weld plate—FEM

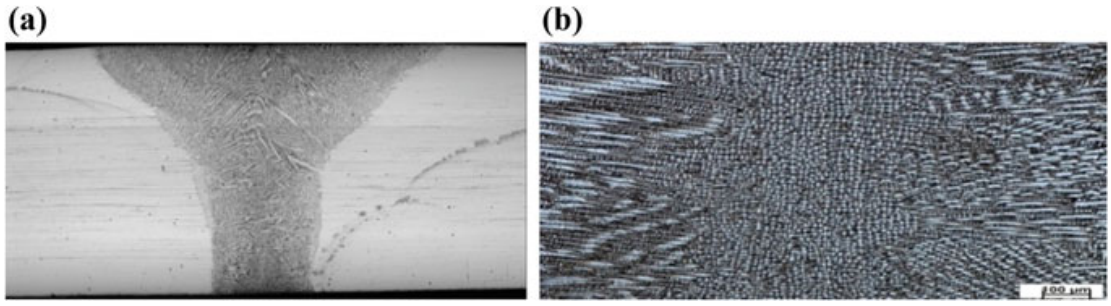


Fig. 22.10 Hybrid laser-TIG **a** weld macrograph and **b** microstructure

acts on the upper part of the weld and widens the bead width. The laser acts on the bottom part of the weld pool through the keyhole mode. This synergistic action of the combined heat sources gives the wine cup-shaped weld bead [8]. All through the entire weld, the central equiaxed grains were sandwiched between columnar dendrites.

22.5 Conclusions

In this paper, the thermomechanical analysis of hybrid laser-TIG welding of type 316L(N) austenitic stainless steel is done and validated by the experiments. The conclusions are given below.

- The calibration of the heat source was done on the basis of the weld bead morphology obtained by welding. The weld macrograph matches perfectly with the model reflecting the accuracy of the heat source chosen.
- The thermal cycle observed using FEM and IR thermography is almost equal and the variations are due to experimental limitations while capturing the thermal data.
- The maximum tensile residual stress predicted is 280 MPa and the experimentally measured is 320 MPa. The residual stress predictions are found to be in a good agreement with the predicted finite element model using SYSWELD.
- The distortion prediction is also in correspondence with validated experiments.

References

1. Vogt, J.B., Foct, J., Regnard, C., Robert, G., Dhers, J.: Low-temperature fatigue of 316L and 316LN austenitic stainless steels. Metall. Trans. A **22**(10), 2385–2392 (1991)
2. Ragavendran, M., Chandrasekhar, N., Ravikumar, R., Saxena, R., Vasudevan, M., Bhaduri, A. K.: Optimization of hybrid laser—TIG welding of 316LN steel using response surface methodology (RSM). Opt. Lasers Eng. **94**, 27–36 (2017)

3. Chandrasekhar, N., Ragavendran, M., Ravikumar, R., Vasudevan, M., Murugan, S.: Optimization of hybrid laser–TIG welding of 316LN stainless steel using genetic algorithm. Mater. Manuf. Processes **32**(10), 1094–1100 (2017)
4. Bagger, C., Olsen, F.O.: Review of laser hybrid welding. J. Laser Appl. **17**(2), 2–14 (2005)
5. Ribic, B., Palmer, T.A., DebRoy, T.: Problems and issues in laser-arc hybrid welding. Int. Mater. Rev. **54**(4), 223–244 (2009)
6. Data Book (1999) Material Properties for Design. IGCAR
7. Suutala, N.: Effect of solidification conditions on the solidification mode in austenitic stainless steels. Metall. Trans. A **14**(1), 191–197 (1983)
8. Gao, M., Zeng, X.Y., Hu, Q.W.: Effects of welding parameters on melting energy of CO_2 laser–GMA hybrid welding. Sci. Technol. Weld. Joining **11**(5), 517–522 (2006)

Chapter 23
Numerical Study on Electromagnetic Crimping of Aluminum Tube and Steel Profiled Rod

Getu Tilahun Areda and Sachin Dnyandeo Kore

Abstract Electromagnetic crimping (EMC) is one of the advanced manufacturing technologies used to join lightweight material including metals with metallurgical and thermal property differences. This paper investigates EMC process numerically to deform aluminum tube (Al6061-T6) to the steel profiled rod using the electromagnetic (EM) module of LS-DYNA^{TM}. Electromagnetic crimping methods can join multiple material systems without melting. A longitudinal groove was designed on steel rod for mechanical interlocking to resist the torsional load. Groove geometrical parameters like groove length, depth, edge radius, and width were kept constant. Effects of voltage on an effective plastic strain, resultant impact velocity, resultant displacement, and temperature were studied in detail. Moreover, the simulation was carried out also to analyze the effect of capacitance on process parameters, and results are discussed in detail. The developed model is validated and found in a good agreement with experiment. Based on the results found, it is possible to predict essential process parameters for enhanced joint strength.

Keywords Electromagnetic crimping (EMC) · Numerical modeling · Groove geometrical parameters · Maxwell's equation · LS-DYNA^{TM} · Johnson-Cook

23.1 Introduction

A new era of a manufacturing process is looking for developing advanced techniques for joining lightweight materials that can overcome problems associated with the conventional approach. Joining methods like fusion welding has several

G. T. Areda (✉)
Mechanical Engineering Department, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India
e-mail: a.getu@iitg.ac.in

S. D. Kore
Mechanical Engineering Department, Indian Institute of Technology Goa, Farmagudi 403401, Goa, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_23

limitations like the creation of intermetallic compound, crack, voids, etc. that needs to be addressed. An advanced technological solution like solid-state joining becomes best options to overcome the existed limitations. Joining metals which have high aspect ratio for lightweight applications requires a method which can provide enhanced joint strength without being affected by thermal property differences between the mating parts. EMC process is a punchless, an environmentally friendly, green, and able to complete the work within a microsecond. This method becomes best suited to resolve defects associated with a thermal-based joining technique like arc welding. Automotive and space frame industries are exceedingly demanding the advantage of this process.

Thermal, mechanical, and electromagnetic are phenomenon's which needed to be addressed using numerical analysis in case of an electromagnetic forming (EMF) process. LSTC^{TM} has introduced an electromagnetic module to simulate the EMF process [1]. In the electromagnetic module of LS-DYNA^{TM}, magnetic field, Lorentz force, current in the conductor, etc. are calculated. The mechanical solver deals with the structural analysis of the deformation. The electromagnetic module solves Maxwell's equations. The thermal solver calculates internal energy, generation of heat and its flow. As the process involves high strain rates, it has all advantages of high velocity forming like improved FLD, minimized spring back, and less wrinkling [2]. Metal forming is one of the LS-DYNA^{TM} main applications with capabilities that allow one to simulate conventional and high-speed forming process [3].

In this chapter, the electromagnetic crimping of an aluminum tube on profiled steel rod was modeled and simulated with 3D finite element method. Simulations are carried out using LS-DYNA^{TM}. Effects of voltage and capacitance on process parameters of electromagnetic crimping were studied. Resultant velocity, resultant displacement, effective plastic strain, and temperature are predicted, and the effect of high voltage on deformation was analyzed using the EM module of LS-DYNA^{TM} software. The compression of the aluminum tube on profiled carbon steel rod has been predicted with 3D finite element simulation at various voltage and capacitance levels.

Capacitor bank in an electromagnetic forming machine is used to store the required amount of energy that will be discharged into the coil, while the high current switch is closed. Induced current starts to flow on the surface of the tube which is caused by a current in the coil. The magnetic field generated due to primary and induced current results radial force known as the Lorentz force used to launch the workpiece away from the coil. The impact that produced due to high speed will make the workpieces to deform beyond its elastic limit for different EMF applications like joining, cutting, welding, and shearing. The total duration of the process takes no longer than 100 microseconds. Electromagnetic crimping process has several applications such that medical (pharmaceutical glass bottles), automotive, space frame structures, wire crimping, double seaming, hemming, and clinching [4, 5]. Capacitors, resistors, and inductors are the main components of the EMF system as shown in Fig. 23.1. The electrical properties of the considered workpieces are exceedingly significant in increasing or decreasing deformation

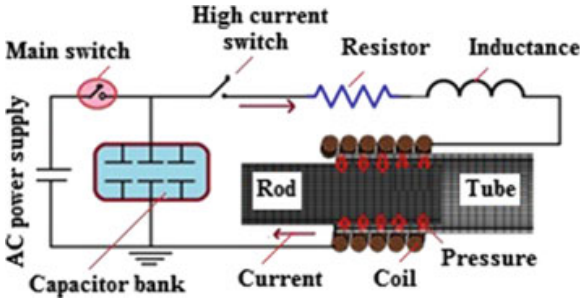


Fig. 23.1 Circuit diagram of EMC process

levels. Direct application of electromagnetic energy can deform metals like magnesium and aluminum which has low formability in the conventional forming process. Applying magnetic energy directly to deform lower electrically conductive metals like stainless steel is too difficult.

23.2 Governing Equations of EMF Process

The required electrical energy which is stored in the bank of the capacitor is given by Eq. 23.1, and the resulting current $I(t)$ by Eq. 23.2.

$$E_c(t) = \frac{1}{2} CV^2(t) \tag{23.1}$$

$$I(t) = V_0 \sin(wt)\mathrm{e}^{-\beta t} \tag{23.2}$$

where $V(t)$ voltage across the capacitor, C capacitance, ω frequency, β damping coefficient, and t time. Set of Maxwell's equations form the basis of numerical computation in any FEM software. The generation of magnetic pressure due to the interaction of two opposite magnetic field is computed using Maxwell's equations. The resultant Lorentz force which is a function of current density and magnetic field density also calculated in the same approach. Differential forms of Maxwell's equations (Eqs. 23.3–23.6) are shown below.

$$\nabla \times B = \mu_0 (J + \varepsilon_0 \partial E / \partial t) \tag{23.3}$$

$$\nabla \times E = -\partial B / \partial t \tag{23.4}$$

$$\nabla \cdot E = \rho / \varepsilon_0 \tag{23.5}$$

$$\nabla \cdot B = 0 \tag{23.6}$$

where E electric flux intensity, B magnetic field density, μ_0 magnetic permeability, J current density, ρ net charge density, ε_0 permittivity of vacuum, and t is for time.

23.3 Geometrical Dimensions of Workpiece and Material Properties

Geometrical dimensions tube, copper coil, and steel profiled rod are shown in Fig. 23.2a–c. Total of four longitudinal grooves was designed, and its geometrical parameters kept constant for all simulation. Axisymmetric coil with dimension and shape as shown in Fig. 23.2b was used.

Table 23.1 shows the properties of the aluminum tube, copper oil, and steel rod used in this study.

23.4 3D Finite Element Modeling

Simulations were carried out for different voltages levels using identical coil geometry which is modeled as a rigid body to make it simple for calculation in the boundary element method. Hence, deformation was analyzed only on the tube at the joining zone. The ranges of voltage were varied from 1.5 to 4 kV with an increasing rate of 0.5 kV. The arrangement of workpieces and tool coil for simulation is shown in Fig. 23.3. The model parts have meshed with 3290, 6580, and 2394 number of elements for the coil, rod, and tube, respectively.

Electromagnetic crimping is a high-speed joining process using a magnetic pulse, and the material model which take care of strain rate hardening and softening due to temperature rise is Johnson cook (JC). In order to evaluate the constitutive response of the tube, the flow stress which is determined as a function of the plastic

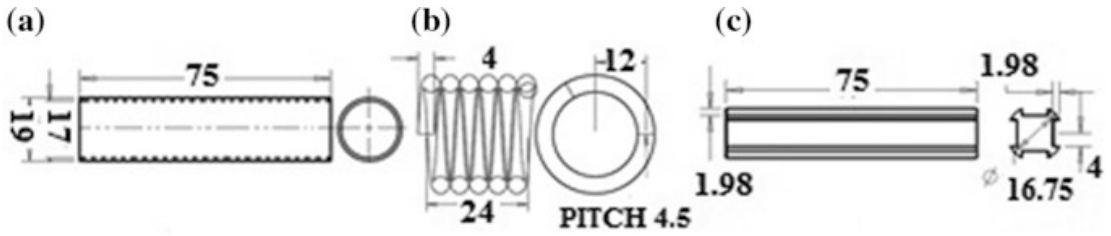


Fig. 23.2 Geometrical dimension of **a** tube **b** coil, and **c** rod (dimensions are all in mm)

Table 23.1 Material properties of tube, rod, and coil used in the simulation [6, 7]

Material	ρ (kg/m^3)	G (GPa)	E (GPa)	γ	C (J/kg K)	K (W/m k)	σ (MS/m)	T_m (K)
Al6061-T6	2700	26	68.9	0.31	896	167	25	925
Steel	7850	–	207	0.29	–	–	–	–
Cu1001	8960	–	124	0.27	–	–	59.6	–

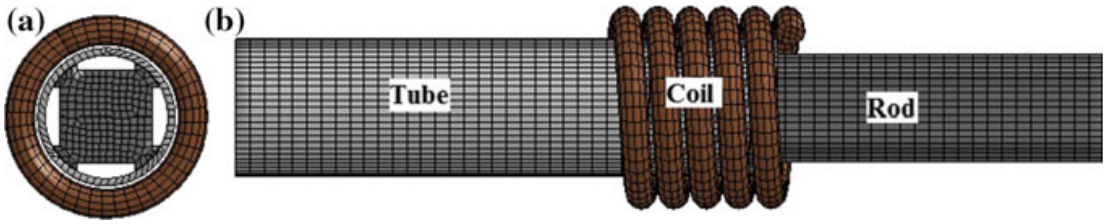


Fig. 23.3 Arrangement of rod, coil, and tube **a** side and **b** front view

Table 23.2 Parameters of JC strength model for Al6061-T6 and steel [9, 10]

Materials	A (MPa)	B (MPa)	C	n	m	T_m (K)	$\dot{\varepsilon}^*$ (s^{-1})
Al6061-T6	324	114	0.002	0.42	1.34	925	1
4340 Steel	729.2	5.095	0.014	0.26	–	1793	–

Table 23.3 Parameters for linear polynomial EOS for tube [10]

C_0	C_1	C_2	C_3	C_4	C_5	C_6	E_0	V_0
0	74.2 GPa	60.5 GPa	36.5 GPa	1.96	0	0	0	1

strain, strain rate, and temperature is related in JC constitutive equation as discussed by Schwer [8].

$$\sigma_y = \left(A + B\,\bar{\varepsilon}^{p^n}\right)\left(1 + c\ln\dot{\varepsilon}^*\right)\left(1 - \left(\frac{T - T_R}{T_m - T_R}\right)^m\right) \tag{23.7}$$

where A is an initial yield strength, B and C are material constant, $\bar{\varepsilon}^{p^n}$ reference plastic strain, $\dot{\varepsilon}^*$ equivalent plastic strain rate of the workpiece, n strain hardening index, m thermal softening index, T_R room temperature, and T_m melting temperature of the workpiece, T operating temperature.

Parameters of JC strength model for Al6061-T6 and steel and parameters for linear polynomial EOS for Al6061-T6 are shown in Tables 23.2 and 23.3.

23.5 Results and Discussion

Numerical results are discussed and analyzed in detail as follows. Effects of voltage and capacitance on other process parameters were explained based on the selected element in the joining zone. Moreover, the transient magnetic field is determined for different values of capacitance.

23.5.1 Effect of Voltage on Effective Plastic Strain

Effective plastic strain result at different voltage levels is shown in Fig. 23.4. Maximum of 1.038 and minimum of 0.0127 effective plastic strain were obtained using 4.0 and 1.5 kV for an element (H-24494) which is found in the deformation zone. These results were obtained from the element found at maximum deformation which is parallel to the groove surface. There were rebounding effects when we considered voltage more than 4 kV that leads to having minimum radial displacement and effective plastic strain.

23.5.2 Effect of Voltage on Resultant Displacement and Velocity

Figure 23.5a shows the resultant displacement of an element (H-24494) at different voltage levels. Keep increasing voltage levels causes the resultant displacement to reach its maximum due to increasing discharge energies. Increasing voltage to 4.0 kV leads an element to displace by 2.29 and 0.085 mm was found while using 1.5 kV voltage. Resultant displacement reaches a peak at a different time for each voltage level but almost equal to for more than 2.5 kV with small variations. If we go beyond 4.0 kV, radial displacements will significantly decrease due to rebounding effect. Measuring parameters like a radial or resultant displacement during an experiment is a challenging task due to the involving speed, but it is possible to predict using 3D simulation as shown in Fig. 23.5a. Besides that, resultant velocity is one of the critical process parameters needs to be determined by the electromagnetic crimping process. Increasing voltage from 1.5 to 4.0 kV causes the velocity to rise from 15.075 to 438.44 m/s. The percentage of increment of

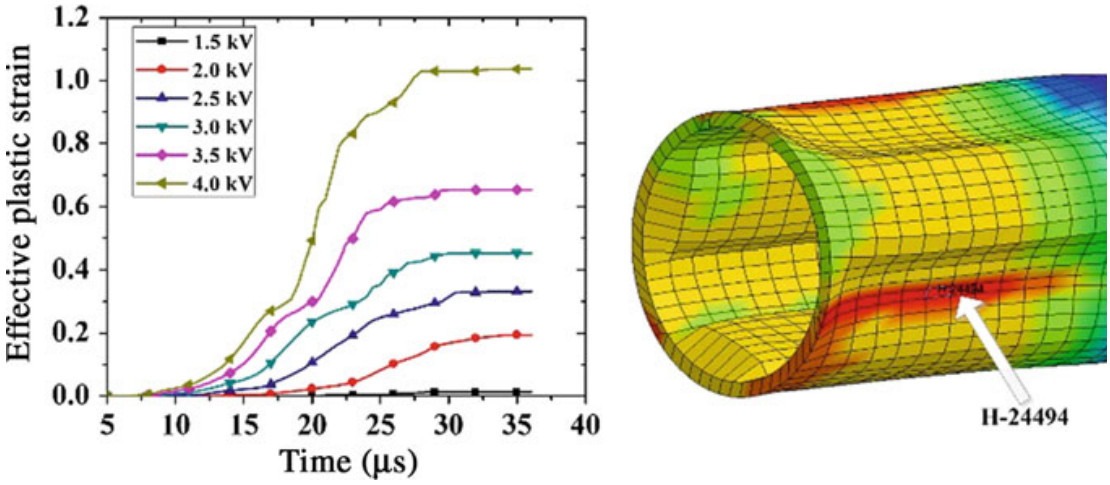


Fig. 23.4 Effect of voltage on effective plastic strain

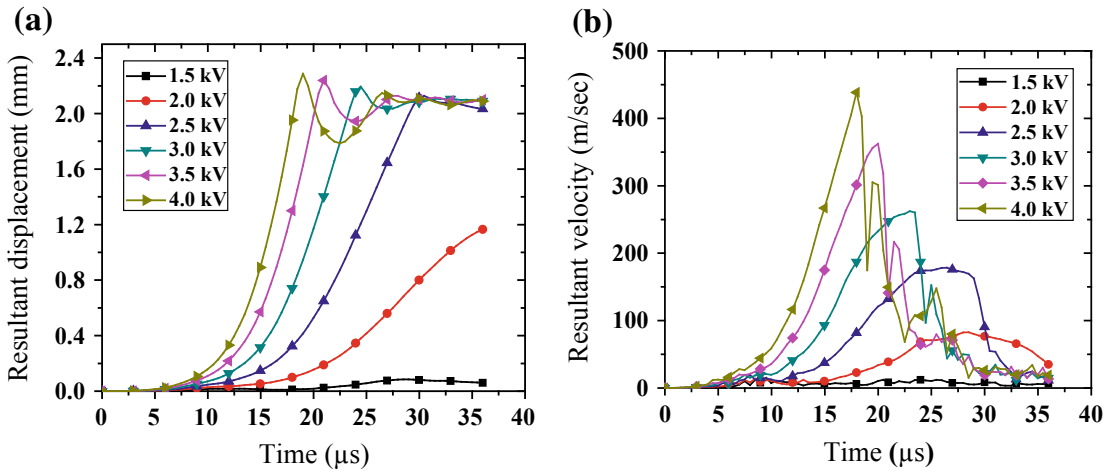


Fig. 23.5 Effect of voltage on **a** resultant displacement and **b** resultant velocity

velocity compared to variation in voltage shows some difference, but it has direct influence as shown in Fig. 23.5b. Minimum velocity of 15.075 m/s was found while applying a voltage of 1.5 kV.

23.5.3 Effect of Voltage on Temperature and Current

As a solid-state process, increments of temperature up to a melting point of any of the metal under consideration were not expected. The rising of temperature in electromagnetic crimping process is not sufficient to cause phase change during deformation, but it may soften the joining zone. Maximum of temperature 568.19 K was obtained using 4.0 kV as shown in Fig. 23.6a. The result shows an apparent variation of temperature caused by different voltage level. Minimum increment of 27.15 K from room temperature was found while using 1.5 kV compared to

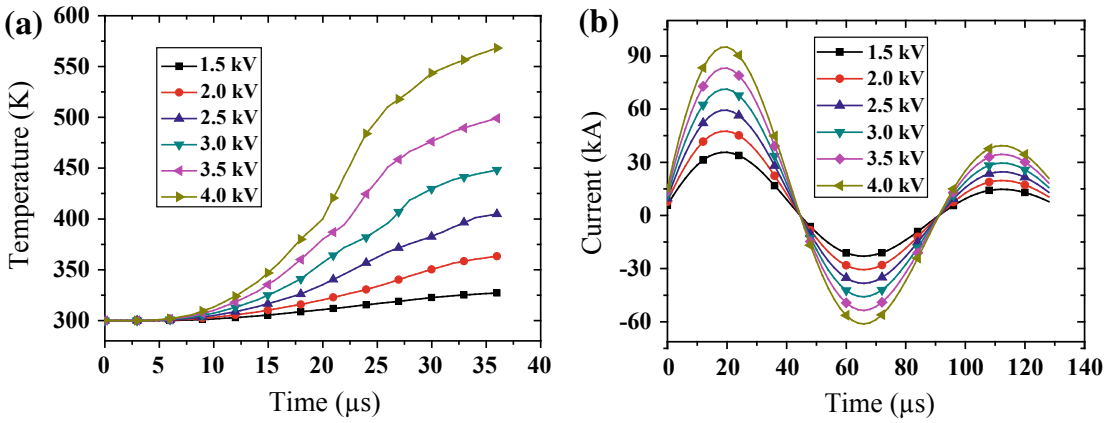


Fig. 23.6 Effect of voltage on **a** temperature and **b** current

different voltage levels. Variations of current at different voltage levels for identical RLC circuit are shown in Fig. 23.6b. The amplitude of the current curve for all voltage levels reaches a maximum at 20 microseconds. Maximum current was achieved while using 4.0 kV that is 95 kA, and a minimum of 35.6 kA was found while using 1.5 kV. The direct proportionality between voltage and current obtains the pattern as shown in Fig. 23.6b for different voltage levels. Current tends to increase as voltage increases, and variation of 0.5 kV has a significant effect on increasing or decreasing its amplitude.

23.5.4 *Effect of Capacitance on Process Parameters (V and ε)*

Simulation result obtained using 3.0 kV was referred to analyze the effect of capacitance on resultant velocity and effective plastic strain. Increasing 22.68% of capacitance from 426 to 551 μF leads to an increment of a 10.66% resultant velocity. This increment is depicted graphically in Fig. 23.7a. It is required to have a significant difference in capacitance to obtain great variation in resultant velocity. Stored energy will be high if the capacitance is increased and that leads to having an increased effective plastic strain as shown in Fig. 23.7b.

An element (H-24494) on the outside surface of the tube and at the middle of the coil was selected to measure the strain. The effective plastic strain was found uniformly increased as capacitance varies from 426 to 551 μF. Maximum and minimum effective plastic strain of 0.530 and 0.454 were obtained while rising values of capacitance from 426 to 551 μF. It is essential to increase capacitance significantly to attain a meaningful variation in resultant velocity and plastic strain.

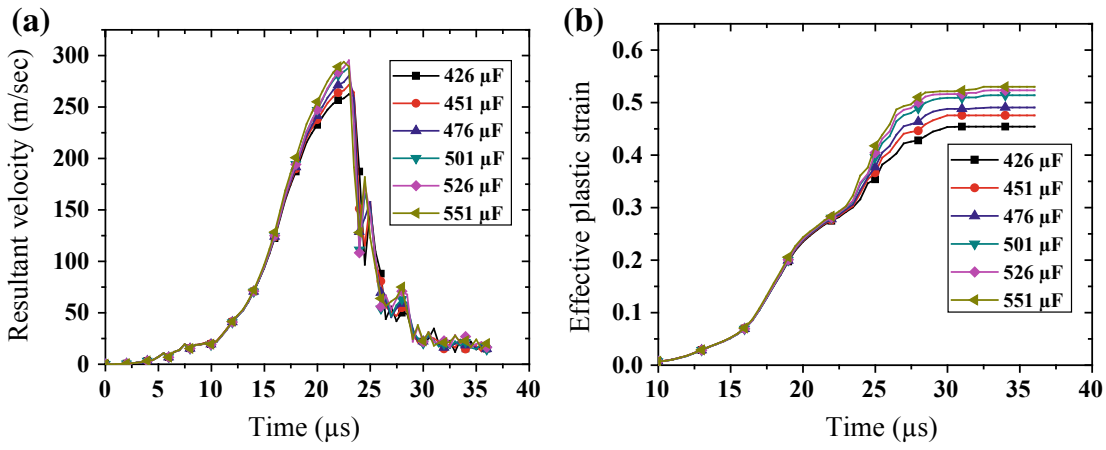


Fig. 23.7 Effect of capacitance on **a** resultant velocity and **b** effective plastic strain

23.5.5 Effect of Capacitance on Resultant Displacement and Temperature

Increasing capacitance by 22.6% leads the resultant displacement to increase only by 3.22% as demonstrated in Fig. 23.8a. The proportional variation between capacitance and resultant displacement shows great differences. Beside that Fig. 23.8b shows how temperature varies due to different levels of capacitance where inductance, resistance, and voltage kept constant. A temperature increment of 4.9% was obtained by comparing the maximum and the minimum result found for increasing capacitance by 22.6%. Room temperature was considered as simulation condition which is 300 K.

23.5.6 Effect of Capacitance on Magnetic Field

Magnetic field increases as capacitance increases by keeping other circuit parameters constant. The minimum magnetic field of 10.7 T was found using 426 µF, and a maximum of 14.62 T obtained using 551 µF capacitance. Lorentz force which is a function of the magnetic field also increases accordingly. The result revealed that increasing the capacitance by 29.34% leads to increasing the magnetic field by 36.6%. It is vital to add capacitance of the EMF (electromagnetic forming) system based on the required level of deformation, which gives a significant difference in the resulted magnetic field.

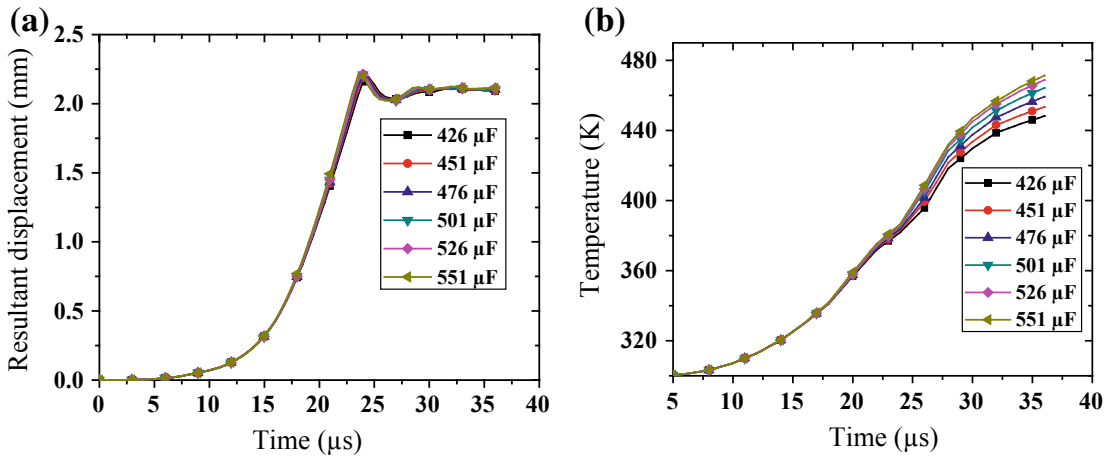


Fig. 23.8 Effect of capacitance on **a** resultant displacement and **b** temperature

23.5.7 Failure Prediction

Figure 23.9 shows von Mises stress distribution on the joining zone at a different time value using 4 kV. The magnitude of stress is found maximum at groove edge that possibly leads to generate necking and thinning which cause joint failure. One of the critical geometrical parameters which can directly affect the joint strength by acting as a cutting edge is the groove edge radius which is 0.5 mm in this study. Moreover, Fig. 23.10 confirms the failure region by showing where shear strain, von Mises stress, and shear stress become maximum.

Results are taken at the end of 36 microseconds for different energy levels. A minimum and maximum von Mises stress of 341.6 and 382.6 MPa were found at the various voltage levels. A better groove filling is observed while using 4 kV that leads to resist higher torque load. Therefore, 4 kV is found optimum voltage for the given material system and geometrical parameters.

23.5.8 Experimental Work

The experiment is carried out using a setup shown in Fig. 23.11 which has a similar geometrical parameter to the FE model. The same geometrical and process parameters are used to conduct the experiment. Discharge energy generated using

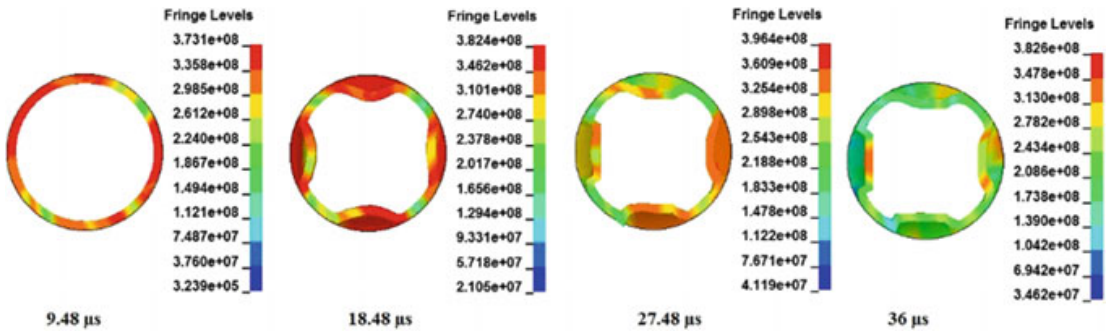


Fig. 23.9 Distribution of von Mises stress on joining zone for 4 kV at the different time value

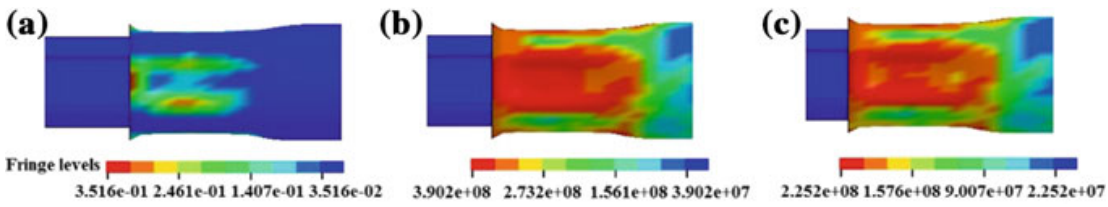


Fig. 23.10 Contours of **a** maximum shear strain **b** von Mises stress, and **c** Tresca (maximum shear stress)

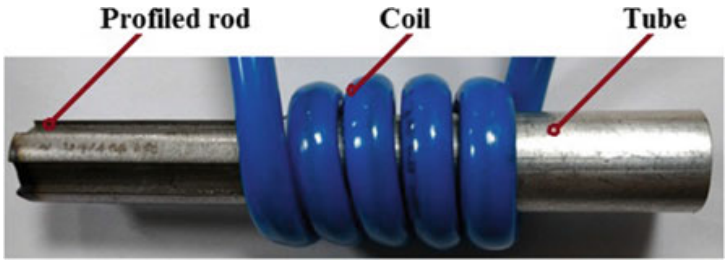


Fig. 23.11 Experimental setup

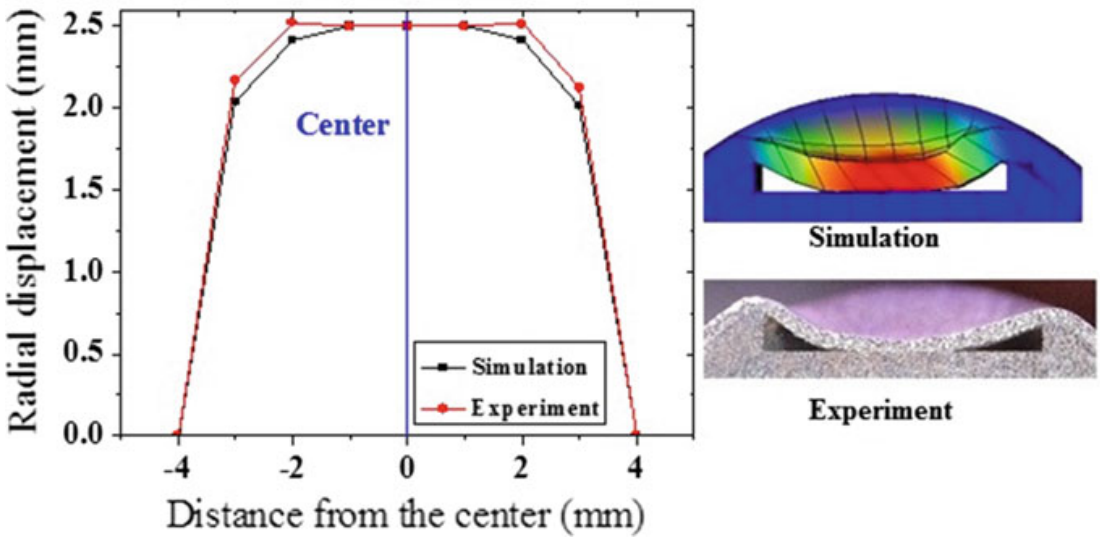


Fig. 23.12 Radial displacement in simulation and experiment

4 kV is applied, and the radial displacement is measured to compare with the simulation.

The radial displacement was measured using cross section of crimped sample and compared with simulation as depicted in Fig. 23.12. Measured gap between the tube and groove surface at the deformation zone was found in good agreement with the simulation. The deformation at the center of the coil is higher that leads to having the maximum radial displacement. The radial displacement was measured with 1 mm gap starting from the center to ends of the groove.

23.6 Conclusions

In this study, numerical modeling and simulations of electromagnetic crimping for transferring torsional load were demonstrated. The aluminum tube gets crimped on the groove of the steel rod successfully, while electromagnetic energy is applied. Process parameters which are difficult to measure and the possible failure region during the experiment were predicted successfully using the EM module of

LS-DYNA™. The increasing voltage causes the effective plastic strain, radial displacement, resultant velocity, temperature, and output current to increase accordingly. Moreover, the effects of increasing capacitance on resultant velocity, effective plastic strain, radial displacement, temperature, and magnetic field were found less significant. The result revealed that increasing the capacitance by 29.34% leads to increasing the magnetic field by 36.6%. Further, the resultant displacement is getting decreased as voltage increases beyond 4 kV due to re-bouncing effect. Therefore, 4 kV is found optimum voltage for the given material system and geometrical parameters. A mean percentage error which is less than 7% was found while comparing radial displacement obtained from simulation and experiment. The developed model was validated and shows a good agreement with experiment.

References

1. L'Eplattenier, P., Çaldichoury, I.: Update on the electromagnetism module in LS-DYNA. In: 12th LS-DYNA Users Conference, Detroit (2012)
2. LS-DYNA Theory Manual, LSTC
3. L'Eplattenier, P., Cook, G., Ashcraft, C., Burger, M., Shapiro, A., Daehn, G., Seth, M.: Introduction of an electromagnetism module in LS-DYNA for coupled mechanical-thermal-electromagnetic simulations. Steel Res. Int. **80**(5), 38–56 (2009)
4. Psyk, V., Risch, D., Kinsey, B.L., Tekkaya, A.E., Kleiner, M.: Electromagnetic forming—a review. J. Mater. Process. Technol. **211**(5), 787–829 (2011)
5. Zittel, G.: A historical review of high speed metal forming. In: Proceedings of the 4th International Conference on High Speed Forming—ICHSF, pp. 2–15 (2010)
6. Patel, C., Kore, S.D.: Effect of frequency on electromagnetic expansion of thin tubes. J. Mach. Form. Technol. **7**(1/2), 91(2015)
7. Rajak, A., Kore, S.: Electromagnetic hemming of aluminum sheets using FEM. In: ICMMD-2016 Advances in Intelligent Systems Research, pp. 77–82. Atlantis Press (2017)
8. Schwer, L.: Optional strain-rate forms for the Johnson-Cook constitutive model and the role of the parameter. Epsilon_0, LS-DYNA Anwenderforum, LS-DYNA Frankenthal, pp. 1–14 (2007)
9. Johnson, G.R., Cook, W.H.: A constitutive model and data for metals subjected to large strains, high strain rates and high temperatures. In: Proceedings of the 7th International Symposium on Ballistics, vol. 2, pp. 541–547 (1983)
10. Schwer, L.E., Windsor, C.A.: Aluminum plate perforation: a comparative case study using Lagrange with erosion, multi-material ALE, and smooth particle hydrodynamics. In: 7th European LS-DYNA Conference, May (2009)

Chapter 24
Temperature Prediction During Self-pierce Riveting of Sheets by FEA-ANN Hybrid Model

Deepak Mylavarapu, R. Ganesh Narayanan and Manas Das

Abstract Nowadays, the automobile industries are more concerned about reducing automobile weight for improving fuel efficiency and reduced vehicle emission. Hence, there is a need for relatively lighter weight materials like aluminium alloy replacing steel parts for automobiles. However, joining of these lightweight materials is very difficult using conventional spot welding technique, which is a slow process also. In the present study, an alternative high-speed advanced mechanical fastening technique, namely self-pierce riveting is used for joining of these lightweight sheets of different materials. In this process, a semi-tubular rivet is pressed by a punch into two or more sheets which are supported on a die. Due to the special design of die shape, the rivet flares inside bottom sheet to form a mechanical interlock. Experimental measurement of temperature inside the sheets during the process is a difficult task, as the process takes place for a fraction of a second within a span of 10 mm. In the present study, the temperature rise during SPR process is predicted by finite element simulation using Abaqus® for any combination of sheet materials, process conditions and tool dimensions to understand the process. In addition, an ANN model is developed to predict the temperature rise during joint formation in sheets for any combination of process parameters. The ANN model accurately predicted the temperature evolution during the process. Hence, with the hybrid ANN-FEM strategy one can design and optimize the process, tool and material conditions, efficiently, which otherwise will be time and resource-intensive.

Keywords Self-pierce rivet · Temperature · Neural network · Finite element simulation

D. Mylavarapu · R. Ganesh Narayanan (✉) · M. Das
Department of Mechanical Engineering, IIT Guwahati, Guwahati 781039, India
e-mail: ganu@iitg.ac.in

M. Das
e-mail: manasdas@iitg.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_24

24.1 Introduction

The overall weight of an automobile governs the fuel efficiency and vehicle emissions. Polmear [1] pointed out that the fuel consumption can be reduced by 5.5% for each 10% reduction in weight. This is possible when lightweight structures and lightweight materials are used in automotive sectors. Use of aluminium alloys in place of steel is one of the possible solutions. Hirsch [2] highlighted that, as part of Super Light Car project, about 53% of the parts contribution is from Aluminium alloy in the form of sheet panels, extrusion in front rails, bumper, crash parts, in the rear underbody rail, and in wheelhouse structure as high-pressure die-cast. The formability of Aluminium sheets is also acceptable and can be improved by materials processing. Joining of lightweight sheets including Aluminium alloys is sometimes difficult with traditional welding processes. In this context, friction stir welding, adhesive bonding, mechanical clinching, laser welding and some hybrid joining methods are used successfully with support from numerical simulations and parametric optimization [3].

Self-pierce riveting (SPR) is a high-speed cold joining method used for joining similar and dissimilar materials including non-metals and metals. In SPR, a rivet is made to deform the sheets arranged in lap configuration above a die with groove. A punch is used to displace the rivet downwards. Depending on the die design, the sheets are mechanically locked. In the process of joining, the upper sheet is pierced and the rivet is flared, while the lower sheet is just deformed.

The stages in SPR along with the schematic of load evolution are shown in Fig. 24.1. During clamping, the rivet is pushed by a punch perpendicular to the sheet surface. In this stage, the rivet clamps the sheets. In piercing stage, the rivet causes severe plastic deformation of the upper sheet resulting in sheet piercing. The lower end of the rivet flares in the third stage and follows the contours of the die. Sheet overlap and interlock distance that are responsible for the required mechanical resistance are achieved during the stage. In the compression stage, the punch continues to push the rivet inside the sheets to be joined. The punch movement stops

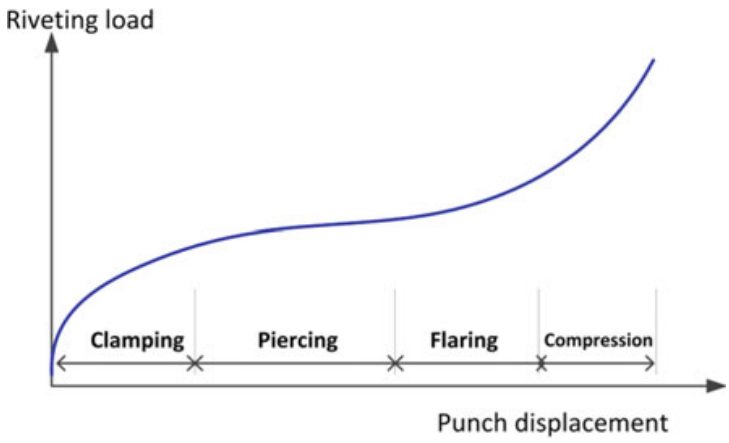


Fig. 24.1 Load evolution in SPR

when it reaches a predetermined load or displacement. The rivet head surface is levelled with the sheet surface at this stage.

Some important contributions in SPR of variety of materials including those involving modelling is presented here. Xu [4] studied the influence of three important factors, namely die type (height of projected region is varied), rivet length and sheet combination on the SPR quality indices like bottom sheet thickness, undercut and rivet spread. The relation between SPR variables and outputs is not unique. The die requirements change with the outputs monitored. The rivet length has negative effect on bottom thickness, but positive effect on undercut and spread. Some of these parameters have interaction effect as well. The experimental analyses by Sun et al. [5] showed superiority of SPR joints (made of aluminium and steel grade sheets) as compared to resistance spot welded joints. The adhesive bonding also improved the static and fatigue behaviour, specifically in lap shear mode. In a similar attempt by Mori et al., it has been observed that the static and fatigue strengths of SPR joints are better than that of clinched joints and resistance spot welded joints. There exists a slight slip at the joint interface in the case of SPR and clinched joints that have improved the fatigue behaviour as compared to resistance spot welding [6]. Johnson et al. [7] attempted to evaluate the joint quality by computer image processing method. With this system, one can identify the number of rivets present in the machine jaw and the material thickness after joining. A comparative analysis of SPR joints made by steel rivet and aluminium rivet by mechanical testing revealed that the rivet material has no influence on the initial stiffness of the mechanical response [8]. But the maximum force and the displacement at maximum force get affected significantly. It is also observed that the SPR joint made by steel rivet can tolerate larger normalized maximum load as compared to the one made by aluminium rivet at varying loading angle.

Several authors have attempted to study the SPR process through numerical simulations. Mori et al. [9] through finite element analyses (FEA) optimized the die shape to produce a successful SPR joint made of multi-layered steel-Al sheets. Porcaro et al. [10] performed numerical simulations using a 2D axisymmetric model in LS-DYNA for joining two sheets using rivet and tools. The effect of mesh size, adaptive time interval, friction and material failure on the joint prediction is studied. Bouchard et al. [11] investigated the SPR process using Forge 2005, an FE code. A Lemaitre-coupled damage model has been applied to predict failure during the process. In the same manner, Casalino et al. [12] implemented fracture conditions through plastic strain threshold during SPR numerical simulations to understand the piercing of rivet leg into the sheets. He et al. [13] through numerical simulations and experiments showed that the SPR joints and hybrid SPR joints (SPR + adhesive bonded) performed better than adhesive bonded joints. The article on recent developments in FE simulations of SPR process presented by He et al. [14] is noteworthy. Mylavarapu et al. [15] predicted critical thinning during SPR through FE simulations. They proposed a strain mapping method to evaluate the thinning of sheets. Later, they used a trained artificial neural network (ANN) model for which FE simulations were conducted to predict the temperature evolution during SPR. The friction coefficient, upper sheet materials properties and lower

sheet material properties like strain hardening exponent and strength coefficient are the input parameters for temperature prediction [16]. The present work is an extension of the previous attempt made by the authors.

The main aim of the present work is to predict the temperature evolution during SPR by FEA-ANN hybrid model. For this case, ten different SPR parameters are considered for temperature prediction. SPR process is modelled and simulated using a FE code ABAQUS® using explicit-temperature-displacement module. Temperature prediction during SPR is not attempted until now. Such analyses will help in selecting the materials and process conditions appropriately during SPR.

24.2 FE Simulation of SPR

FE simulation of SPR is performed using a 2D-axisymmetric model that is generated in ABAQUS® using explicit-temperature-displacement module. Four-node 2D elements with four Gauss points and hourglass control are selected for modelling. An average element size of 0.5 mm × 0.5 mm is used for meshing. The element type of CPS4RT, which is a four-node plane-strain quadrilateral element with reduced integration along with temperature DOF, is used. Reduced integration corresponds to lowering of the order of integration as compared to that of full integration. The meshing in the sheets and the blank holder is done in structured form, but the meshing of the rivet, die and the punch are done in free meshing form. This causes the rivet to have varying element dimensions. The FE model contains a rivet, bottom plate, top plate, punch, blank holder and die. Among this, the punch, blank holder and die are assumed as rigid parts. The rivet and sheets are considered plastically deformable during riveting process and are modelled as elastic–plastic materials. The rivet geometry is assumed to have a nominal diameter of 5 mm and is made of high-strength steel. The punch is of 5 mm diameter with 2 mm as its corner radius. The blank holder has inner radius of 10 mm and outer radius of 20 mm and is modelled as an annular ring. The die geometry has a projection at the centre that controls the rivet flaring. The blank holder and punch geometries are modelled such that sufficient clearance is provided between the surfaces. The dimensions are based on standards available practically. Figure 24.2 shows the FE model for SPR analyses and one of the simulated cases.

The sheet material properties represent a 5xxx Al alloy, and the stress–strain behaviour of the material is shown in Fig. 24.3. The stress–strain curve was obtained from uniaxial tensile tests following standard procedure. The rivet material was considered much stronger than the sheet. Since its material properties were not evaluated, a multiplying factor was considered with respect to the sheet to obtain its stress–strain behaviour. However, both are modelled as elasto–plastic materials. The other tools (die, punch and the blank holder) were modelled as rigid by considering properties of tool steel which has very high elastic modulus. The other material properties required to simulate the FE explicit-temperature-displacement model are given in Table 24.1.

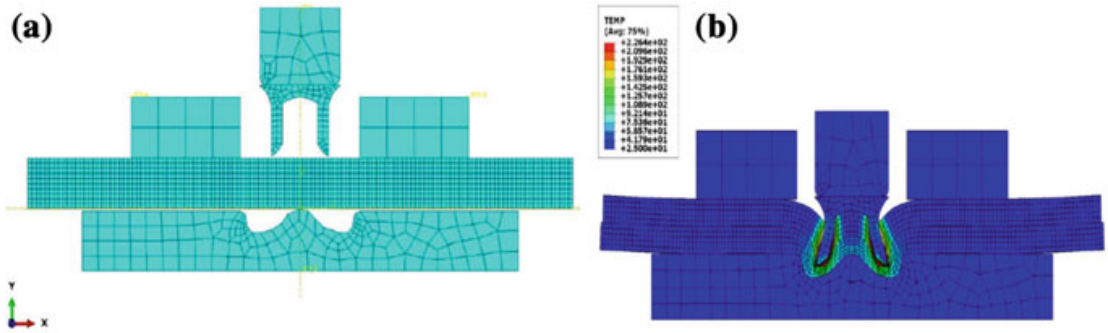


Fig. 24.2 FE model of SPR **a** initial stage for 0.21 tip height, **b** FE simulation of an SPR case

Fig. 24.3 True stress–strain behaviour of 5xxx Al sheet

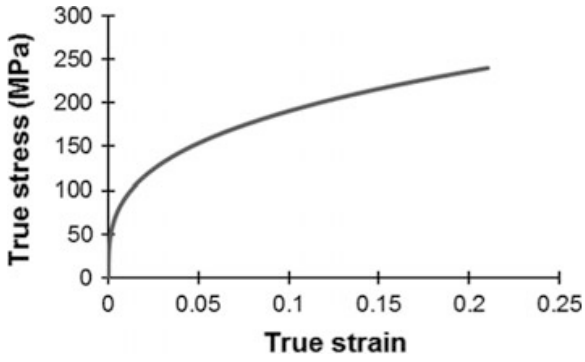


Table 24.1 Material properties used for FE simulations

Material	Properties					
	Density (kg/m^3)	Specific heat (J/kg K)	Thermal conductivity (W/m K)	Inelastic heat factor	Elasticity (GPa)	Poisson's ratio
5xxx sheet	2700	900	205	0.9	68	0.33
Rivet	7830	490	50.2	0.9	180	0.3
Tool	7830	490	50.2	0.9	210	0.3

In the FE model, the interactions are created between the rivet and the sheets, the upper and the lower sheets, the blank holder and the upper sheet, the rivet and the punch, and the die and the lower sheet. The contact was modelled using surface-to-surface penalty formulation. Constant Coulomb's friction coefficient was set as per the levels decided (given later). The sheets are clamped between the die and the blank holder with a holding force of about 5 kN. The die and the blank holder are fixed during SPR simulations. The punch is provided with a prescribed displacement of 6.5 mm to push the rivet through the sheets for joint formation. The rate of punch movement is fixed such that the total displacement occurs in 1 s in real practice. After the completion of each simulation, temperature at an element that is in contact with the rivet tip was evaluated with respect to punch displacement.

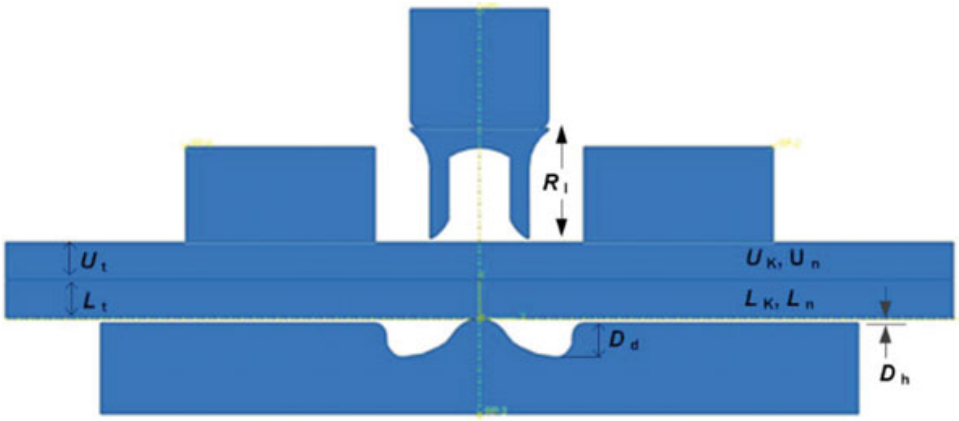


Fig. 24.4 Definition of SPR parameters

24.3 SPR Parameters

The SPR parameters considered for the ANN modelling are Coulomb's coefficience of friction (μ), die tip height (D_h), die depth (D_d), rivet length (R_l), upper sheet thickness (U_t), lower sheet thickness (L_t), upper sheet K (U_K), upper sheet n (U_n), lower sheet K (L_K), lower sheet n (L_n). Here, K and n are the strength coefficient and strain hardening exponent of the sheet materials. The parameters are defined in Fig. 24.4. The levels for these parameters (shown in Table 24.2) were selected based on available SPR data and material properties. The range was chosen to satisfy the practical range available for these parameters.

24.4 ANN Model

In the present work, ANN model was developed to predict the temperature evolution during SPR. A total of 440 FE simulations were carried out for ten parameters (Table 24.2) with several levels for each parameter. To predict the temperature evolution during SPR, the temperature-displacement curves from FE simulations were divided into segments at equal intervals. For example, in Fig. 24.5, the plot for four combinations out of 440 combinations is shown with the segments made at different displacements. Here, for 6-mm-punch displacement, 10 segments are considered. Hence, for every 0.6 mm a partition line is drawn. For one partition line, the curves intersect at four different points. These four temperatures are considered as output for a particular displacement at four different SPR conditions. Similar exercise has been followed for all the segments resulting in large data set for temperature evolution. A sample table for four combinations at one partition line is given in Table 24.3. The eleven different temperature values for a condition will be trained to generate an ANN model.

Table 24.2 Levels of SPR variables used for FE simulations and ANN model

P	L																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
μ	0.1	0.12	0.14	0.16	0.18	0.2	0.22	0.24	0.26	0.28	0.30	0.32	0.38	0.4	0.43	–	–
U_K	100	125	150	175	200	225	250	275	300	310	330	350	370	390	410	430	460
U_n	0.1	0.14	0.16	0.17	0.21	0.23	0.25	0.27	0.29	0.31	0.33	0.34	0.37	0.39	0.4	0.43	0.47
L_K	100	125	150	175	200	225	250	275	300	310	330	350	370	390	410	430	460
L_n	0.1	0.14	0.16	0.17	0.21	0.23	0.25	0.27	0.29	0.31	0.33	0.34	0.37	0.39	0.4	0.43	0.47
D_h	0.21	0.1	0	−0.1	−0.2	−0.29	−0.4	−0.5	−0.6	−0.79	−1	−1.5	−2	–	–	–	–
D_d	1.8	1.85	1.9	1.95	2	2.05	2.15	2.2	–	–	–	–	–	–	–	–	–
U_t	1	1.25	1.5	1.75	2	2.25	2.5	2.75	3	–	–	–	–	–	–	–	–
L_t	1	1.25	1.5	1.75	2	2.25	2.5	2.75	3	–	–	–	–	–	–	–	–
R_l	5.5	5.8	6	6.25	6.5	6.75	7	–	–	–	–	–	–	–	–	–	–

L: Levels; P: Parameters; U_K, L_K in MPa; D_h, D_d, U_t, L_t and R_l in mm

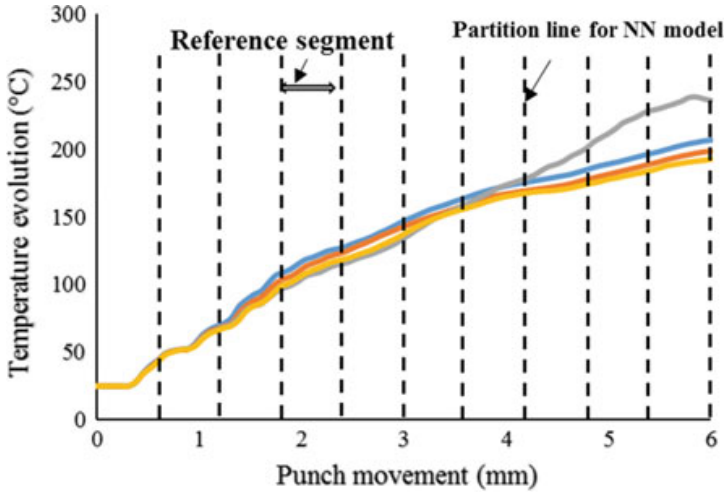


Fig. 24.5 Temperature evolution for four combinations of parameters

Table 24.3 Sample input table for ANN model

SPR combinations	SPR parameters										Temperature (°C)
	U_K	U_n	L_K	L_n	μ	D_h	D_d	U_t	L_t	R_l	
Combination 1	390	0.31	390	0.31	0.47	0.21	2	2	2	6	227
Combination 2	390	0.22	390	0.31	0.10	0.21	2	2	2	6	196
Combination 3	390	0.26	390	0.31	0.10	0.21	2	2	2	6	188
Combination 4	390	0.31	390	0.31	0.17	0.21	2	2	2	6	183

The ANN model was trained by using back-propagation algorithm using an in-house C++ code. The algorithm is based on the error-correction learning rule. In the current application, the objective is to use the network to learn mapping between input and output patterns. The components of the input pattern consist of the control variables of the SPR operation (Table 24.2), whereas the output pattern components represent the temperature rise during the process. The nodes in the hidden layer are necessary to implement the nonlinear mapping between the input and output patterns. Each neuron is connected to all the other neurons in the adjacent layer through the weighted connections. In the forward pass, the synaptic weights remain unaltered throughout the network and the function signals of the network are computed on a neuron-by-neuron basis.

Training and prediction are performed separately for each output temperature, i.e., for each partition line. Thus, the eleven outputs are trained separately, generating eleven ANN architectures. The final ANN architecture contains five input neurons, one hidden layer with ten neurons, and one output neuron corresponding to temperature at each displacement. The final architecture has been decided based on numerous trial exercises. For the first eight displacement levels, tan-sigmoid transfer function has been used between input and hidden layer, while log-sigmoid between hidden layer and output layer. However, for the last three displacement levels, tan-sigmoid transfer function has been used between input and hidden layer and hidden layer and output layer. The coefficient of transfer function between the

input and hidden neurons and the coefficient of transfer function between the hidden neurons and output, learning constant, momentum constant are decided based on trial and error and accuracy of the network. The trained ANN has been validated with the FE simulation results at 44 intermediate levels (Table 24.4) for checking the adequacy of the ANN model.

Table 24.4 Validation cases for checking the adequacy of ANN model

Validation cases	U_K	U_n	L_K	L_n	μ	D_h	D_d	U_t	L_t	R_l
Case 1	390	0.31	390	0.31	0.17	0.21	2	2	2	6
Case 2	390	0.31	390	0.31	0.25	0.21	2	2	2	6
Case 3	390	0.31	390	0.31	0.39	0.21	2	2	2	6
Case 4	**390**	**0.31**	**390**	**0.31**	**0.47**	**0.21**	**2**	**2**	**2**	**6**
Case 5	112	0.31	390	0.31	0.1	0.21	2	2	2	6
Case 6	167	0.31	390	0.31	0.1	0.21	2	2	2	6
Case 7	**237**	**0.31**	**390**	**0.31**	**0.1**	**0.21**	**2**	**2**	**2**	**6**
Case 8	350	0.31	390	0.31	0.1	0.21	2	2	2	6
Case 9	390	0.12	390	0.31	0.1	0.21	2	2	2	6
Case 10	390	0.26	390	0.31	0.1	0.21	2	2	2	6
Case 11	390	0.35	390	0.31	0.1	0.21	2	2	2	6
Case 12	**390**	**0.22**	**390**	**0.31**	**0.1**	**0.21**	**2**	**2**	**2**	**6**
Case 13	**390**	**0.31**	**112**	**0.31**	**0.1**	**0.21**	**2**	**2**	**2**	**6**
Case 14	390	0.31	167	0.31	0.1	0.21	2	2	2	6
Case 15	390	0.31	237	0.31	0.1	0.21	2	2	2	6
Case 16	390	0.31	350	0.12	0.1	0.21	2	2	2	6
Case 17	390	0.31	390	0.29	0.1	0.21	2	2	2	6
Case 18	390	0.31	390	0.35	0.1	0.21	2	2	2	6
Case 19	**390**	**0.31**	**390**	**0.22**	**0.1**	**0.21**	**2**	**2**	**2**	**6**
Case 20	112	0.31	390	0.31	0.15	0.21	2	2	2	6
Case 21	167	0.31	390	0.31	0.25	0.21	2	2	2	6
Case 22	390	0.12	390	0.31	0.15	0.21	2	2	2	6
Case 23	390	0.26	390	0.31	0.25	0.21	2	2	2	6
Case 24	390	0.31	112	0.31	0.15	0.21	2	2	2	6
Case 25	390	0.31	167	0.31	0.25	0.21	2	2	2	6
Case 26	390	0.31	390	0.29	0.15	0.21	2	2	2	6
Case 27	390	0.31	390	0.35	0.25	0.21	2	2	2	6
Case 28	**390**	**0.31**	**390**	**0.31**	**0.1**	**−0.3**	**2**	**2**	**2**	**6**
Case 29	390	0.31	390	0.31	0.1	−0.7	2	2	2	6
Case 30	390	0.31	390	0.31	0.1	0.05	2	2	2	6
Case 31	390	0.31	390	0.31	0.1	−0.05	2	2	2	6
Case 32	390	0.31	390	0.31	0.1	−0.15	2	2	2	6
Case 33	**390**	**0.31**	**390**	**0.31**	**0.1**	**0.21**	**1.97**	**2**	**2**	**6**
Case 34	390	0.31	390	0.31	0.1	0.21	2.07	2	2	6

(continued)

Table 24.4 (continued)

Validation cases	U_K	U_n	L_K	L_n	μ	D_h	D_d	U_t	L_t	R_l
Case 35	390	0.31	390	0.31	0.1	0.21	2.18	2	2	6
Case 36	390	0.31	390	0.31	0.1	0.21	2.02	2	2	6
Case 37	390	0.31	390	0.31	0.1	0.21	2	1.12	2	6
Case 38	**390**	**0.31**	**390**	**0.31**	**0.1**	**0.21**	**2**	**2.12**	**2**	**6**
Case 39	390	0.31	390	0.31	0.1	0.21	2	2.07	2	6
Case 40	390	0.31	390	0.31	0.1	0.21	2	2	1.12	6
Case 41	**390**	**0.31**	**390**	**0.31**	**0.1**	**0.21**	**2**	**2**	**2.12**	**6**
Case 42	390	0.31	390	0.31	0.1	0.21	2	2	2.07	6
Case 43	**390**	**0.31**	**390**	**0.31**	**0.1**	**0.21**	**2**	**2**	**2**	**6.12**
Case 44	390	0.31	390	0.31	0.1	0.21	2	2	2	5.92

24.5 ANN Prediction

The trained ANN model has been validated for 44 intermediate cases (given in Table 24.4) that were not part of the training set. Out of these, validation for ten cases, namely case 4, case 7, case 12, case 13, case 19, case 28, case 33, case 38, case 41, case 43 (indicated in Table 24.4), is shown in Fig. 24.6. It is observed that the temperature evolution predictions from ANN agree well with that of from FE simulations. There is about 5% deviation. Similar accuracy levels are observed in all other cases as well. Figure 24.6k shows the accurate ANN predictions for all 44 SPR cases. With the FEA-ANN hybrid model, one can design and optimize the SPR process, tool and material conditions, efficiently, which otherwise will be time and resource intensive.

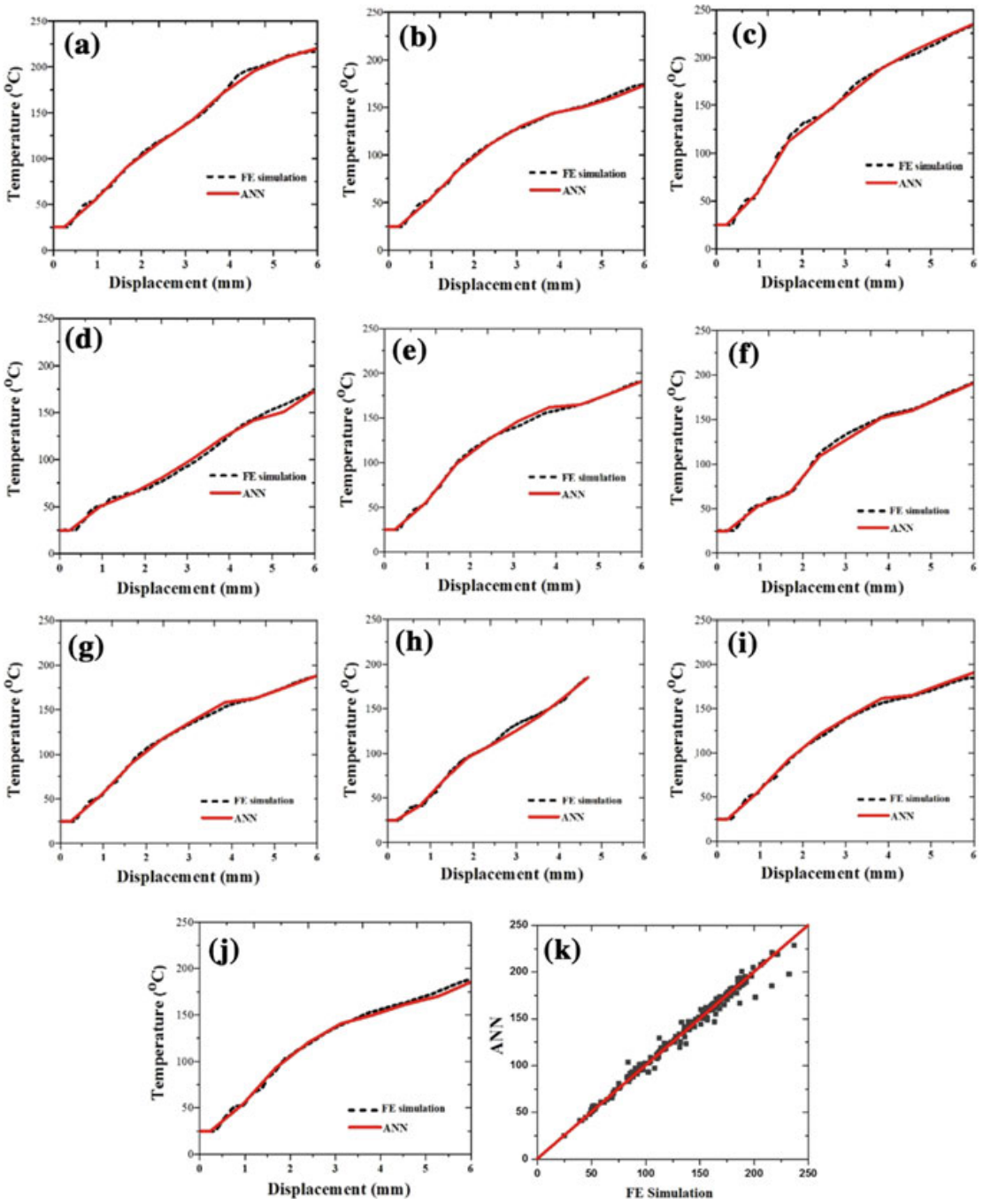


Fig. 24.6 ANN model validation with FE simulation results **a** case 4, **b** case 7, **c** case 12, **d** case 13, **e** case 19, **f** case 28, **g** case 33, **h** case 38, **i** case 41, **j** case 43 (indicated in Table 24.4), **k** ANN versus FE simulation results for all 44 cases

24.6 Conclusions

The temperature measurement during SPR is challenging and difficult task to perform in shop floor. The temperature has to be monitored in the deforming region of the sheet which is inside the die with a diameter of about 10 mm and the process occurs for a small duration. The present work demonstrates that the temperature rise

during SPR process can be predicted using a robust FEA-ANN strategy for any combination of sheet materials, process conditions and tool dimensions. The ANN model is sufficiently accurate to predict the temperature evolution during SPR and hence can be used as a 'virtual machine' for analyses.

References

1. Polmear, J.: Light Metals: Metallurgy of the Light Metals, vol. 3, 3rd edn. Wiley, USA (1995)
2. Hirsch, J.: Recent development in aluminium for automotive applications. Trans. Nonferrous Met. Soc. China **24**, 1995–2002 (2014)
3. Mori, K., Bay, N., Fratini, L., Micari, F., Tekkaya, A.E.: Joining by plastic deformation. CIRP Ann. **62**(2), 673–694 (2013)
4. Xu, Y.: Effects of factors on physical attributes of self-piercing riveted joints. Sci. Technol. Weld. Joining **11**(6), 666–671 (2006)
5. Sun, X., Stephens, E.V., Khaleel, M.A.: Fatigue behaviors of self-piercing rivets joining similar and dissimilar sheet metals. Int. J. Fatigue **29**, 370–386 (2007)
6. Mori, K., Abe, Y., Kato, T.: Mechanism of superiority of fatigue strength for aluminium alloy sheets joined by mechanical clinching and self-pierce riveting. J. Mater. Process. Technol. **212**, 1900–1905 (2012)
7. Johnson, P., Cullen, J.D., Sharples, L., Shaw, A., Al-Shamma'a, A.I.: Online visual measurement of self-pierce riveting systems to help determine the quality of the mechanical interlock. Measurement **42**, 661–667 (2009)
8. Hoang, N.H., Porcaro, R., Langseth, M., Hanssen, A.G.: Self-piercing riveting connections using aluminium rivets. Int. J. Solids Struct. **47**, 427–439 (2010)
9. Mori, K., Abe, Y., Kato, T.: Self-pierce riveting of multiple steel and aluminium alloy sheets. J. Mater. Process. Technol. **214**, 2002–2008 (2014)
10. Porcaro, R., Hanssen, A.G., Langseth, M., Aalberg, A.: Self-piercing riveting process: an experimental and numerical investigation. J. Mater. Process. Technol. **171**, 10–20 (2006)
11. Bouchard, P.O., Laurenta, T., Tollier, L.: Numerical modeling of self-pierce riveting-From riveting process modeling down to structural analysis. J. Mater. Process. Technol. **202**, 290–300 (2008)
12. Casalino, G., Rotondo, A., Ludovico, A.: On the numerical modelling of the multiphysics self-piercing riveting process based on the finite element technique. Adv. Eng. Softw. **39**, 787–795 (2008)
13. He, X., Xing, B., Zeng, K., Gu, F., Ball, A.: Numerical and experimental investigations of self-piercing riveting. Int. J. Adv. Manuf. Technol. **69**, 715–721 (2013)
14. He, X., Gu, F., Ball, A.: Recent development in finite element analysis of self-piercing riveted joints. Int. J. Adv. Manuf. Technol. **58**, 643–649 (2012)
15. Mylavarapu, D., Narayanan, R.G., Das, M.: Prediction of critical thinning during self-pierced riveting of sheets. In: Proceedings of 6th International & 27th All India Manufacturing Technology, Design and Research Conference (AIMTDR-2016), pp. 1017–1020, College of Engineering, Pune, India (2016)
16. Mylavarapu, D., Das, M., Narayanan, R.G.: Prediction of temperature evolution during self-pierced riveting of sheets. In: Handbook of Research on Manufacturing Process Modeling and Optimization Strategies (Chapter 18), pp. 381–398, IGI Global, USA (2017)

Part III
Subtractive Manufacturing

Chapter 25
Determination of Optimal Cutting and Tool Geometry Parameters for Better Surface Integrity of Hard Turned AISI 52100 Steel-Hybrid GRA-PCA

P. Umamaheswarrao, D. Ranga Raju, K. N. S. Suman and B. Ravi Sankar

Abstract In the present work, AISI 52100 steel hard turning has been performed using PCBN tools. Cutting speed, feed, depth of cut, nose radius and negative rake angle are the input parameters, and the measured responses are surface roughness and workpiece surface temperature. Experiments are planned as per central composite rotatable design (CCD) of response surface methodology (RSM). The effect of input parameters and their interactions are discussed with main effects plot and response surface plots. Further, the multi-objective optimization scheme is proposed by adopting grey relational analysis (GRA) coupled with the principal component analysis (PCA). Results demonstrated that responses considerably affected by speed followed by nose radius, feed, depth of cut and negative rake angle. Cutting speed 1000 rpm, feed 0.04 mm/rev, depth of cut 0.4 mm, nose radius 1 mm and negative rake angle 15° are the obtained optimum cutting parameters.

Keywords Hard turning · AISI 52100 steel · Surface roughness · Workpiece surface temperature

P. Umamaheswarrao (✉) · B. Ravi Sankar
Department of Mechanical Engineering, Bapatla Engineering College, Bapatla 522102, AP, India
e-mail: maheshponugoti@gmail.com

D. Ranga Raju
Department of Mechanical Engineering, Srinivasa Institute of Engineering and Technology, Amalapuram 533216, AP, India

K. N. S. Suman
Department of Mechanical Engineering, College of Engineering, Andhra University, Visakhapatnam 530003, AP, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_25

25.1 Introduction

Optimum machining parameter's determination in manufacturing industry leads to the reduction in production cost and time, increases the production rate, improves the product quality and reduces environment impact which in turn leads to productivity enhancement [1, 2]. The comparison of conventional machining versus hard turning is shown in Fig. 25.1. [3]. Higher material removal rate, less machining time, less set-up time and greater flexibility are the merits in hard machining in contrast to grinding. Complex part fabrication using hard turning saves 30% manufacturing costs [4]. During AISI 52100 steel hard turning with PCBN tool, feed rate has a significant influence on the surface finish, while cutting speed and depth of cut had marginal effect [5, 6]. Ravi Sankar et al. [7] revealed surface roughness greatly influenced by the speed and nose radius in AISI 52100 hard turning with PCBN tools.

Suhail Adeel et al. [8, 9] obtained better surface roughness at higher the workpiece surface temperature for AISI 1020 steel. The cutting performance could be effectively sensed and controlled by the workpiece surface temperature, while AISI D2 steel hard turning cutting speed, depth of cut followed by feed were the major contributors to workpiece surface temperature development [10]. During turning of EN-9 steel, the workpiece surface temperature was considerably affected by feed subsequently cutting speed and depth of cut [11]. Multi-objective

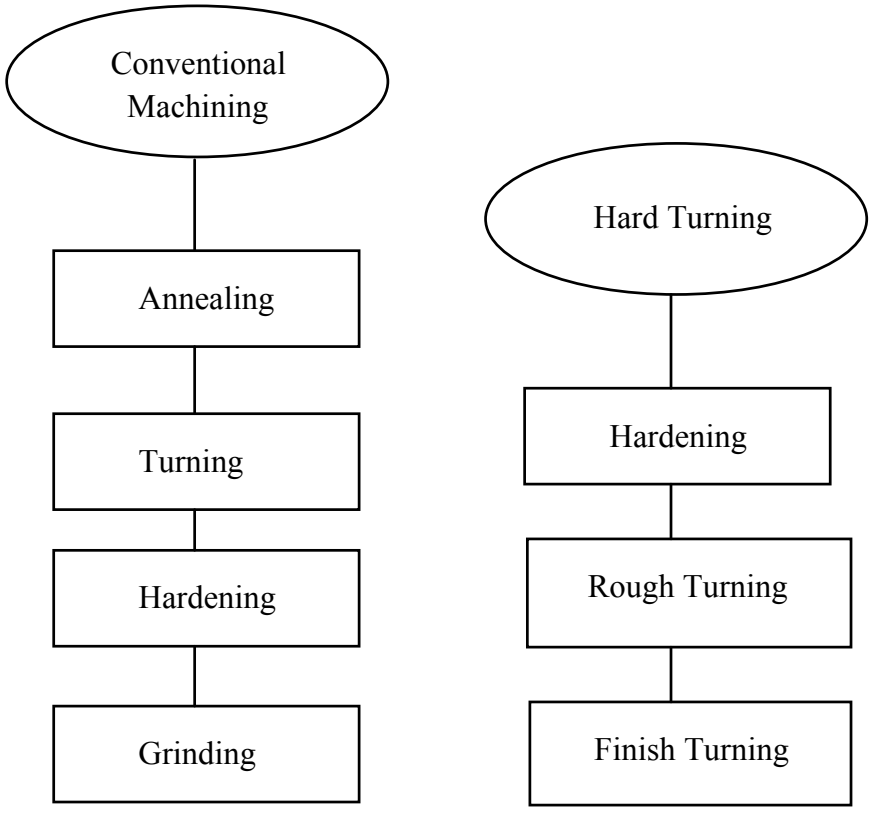


Fig. 25.1 Conventional machining and hard turning comparison [4]

optimization of parameters was performed by deploying GRA coupled with PCA and gained wider accomplishment [12–14].

From the literature, it was elucidated that hard turning was the best alternative to grinding owing to its merits. Much emphasis was made on AISI 52100 steel hard tuning by several researchers due to its applications in various parts of the industry. The past studies made a large amount of interest to investigate the effect of cutting parameters on the responses. Scanty literature was available on the effect of tool parameters on the output responses, and the much interest was focused on nose radius. Most of the studies focused on the machining forces developed and surface roughness during hard turning.

However, the workpiece surface temperature reported to be the sensible and controlling factor for turning performance. Hence, the present study was aimed to conduct hard turning of AISI 52100 steel using PCBN tools with cutting speed, feed rate and depth of cut as cutting conditions whereas nose radius and negative rake angle as tool geometry parameters. Surface roughness and workpiece surface temperature (also termed as surface temperature) were considered as responses. Further, multi-objective optimization was performed deploying GRA coupled with PCA for optimum cutting conditions.

25.2 Experimental Work

Experiments are designed using central composite rotatable design (CCD) of second order which is found to be the most efficient tool in response surface method (RSM). Five input parameters were varied at five levels throughout the experiment, and their influences on responses like surface roughness and workpiece surface temperature were examined. Levels and their factors are depicted in Table 25.1.

Experiments were conducted using variable speed and feed drive lathe in dry condition. Experimental set-up is depicted in Fig. 25.2. Round bars of 48 mm diameter and 500 mm length made of AISI 52100 steel were used for the experiment with a hardness of 57 HRC. Machining length and diameter of workpiece

Table 25.1 Factors and levels

S. No.	Factors	Units	Notation	Levels				
				−2	−1	0	1	2
1	Speed	rpm	v	200	400	600	800	1000
2	Feed	mm/rev	f	0.02	0.04	0.06	0.08	0.1
3	Depth of cut	mm	d	0.4	0.5	0.6	0.7	0.8
4	Nose radius	mm	r	0.4	0.6	0.8	1	1.2
5	Negative rake angle		α	−5	−15	−25	−35	−45

Fig. 25.2 Experimental set-up

were fixed at 30 and 48 mm respectively for every experimental run. In the present work, PCBN inserts were used with different nose radii such as r = 0.4, 0.6, 0.8, 1 and 1.2 mm with different rake angles (−5, −15, −25, −35 and −45). Inserts are mounted on a tool holder of ISO Designation PSBNR2525 M12.

Mitutoyo make surface roughness tester (SJ-210) was used to measure the surface roughness of the turned samples. Workpiece surface temperature was measured using an infrared thermometer manufactured by AMPROBE (range: −50 to 1550 °C, Model: IR 750). Experimental matrix with results is depicted in Table 25.2.

Table 25.2 Experimental results

Expt. No.	Speed (rpm)	Feed (mm/rev)	Depth of cut (mm)	Nose radius (mm)	Negative rake angle (°)	Surface roughness (μm)	Workpiece surface temperature (°C)
1	400	0.04	0.5	0.6	35	0.525	57.43
2	800	0.04	0.5	0.6	15	0.465	74.4
3	400	0.08	0.5	0.6	15	0.453	65.19
4	800	0.08	0.5	0.6	35	0.545	77.68
5	400	0.04	0.7	0.6	15	0.552	71.96
6	800	0.04	0.7	0.6	35	0.507	82.88
7	400	0.08	0.7	0.6	35	0.539	70.27
8	800	0.08	0.7	0.6	15	0.471	65.48
9	400	0.04	0.5	1	15	0.485	66.3
10	800	0.04	0.5	1	35	0.401	66.3
11	400	0.08	0.5	1	35	0.507	84.38

(continued)

Table 25.2 (continued)

Expt. No.	Speed (rpm)	Feed (mm/rev)	Depth of cut (mm)	Nose radius (mm)	Negative rake angle (°)	Surface roughness (μm)	Workpiece surface temperature (°C)
12	800	0.08	0.5	1	15	0.502	80.11
13	400	0.04	0.7	1	35	0.508	67.07
14	800	0.04	0.7	1	15	0.408	80.3
15	400	0.08	0.7	1	15	0.604	68.76
16	800	0.08	0.7	1	35	0.498	76.5
17	200	0.06	0.6	0.8	25	0.559	66.61
18	1000	0.06	0.6	0.8	25	0.456	82.73
19	600	0.02	0.6	0.8	25	0.468	70.85
20	600	0.1	0.6	0.8	25	0.53	74.88
21	600	0.06	0.4	0.8	25	0.45	71.39
22	600	0.06	0.8	0.8	25	0.48	74.2
23	600	0.06	0.6	0.4	25	0.514	67.18
24	600	0.06	0.6	1.2	25	0.485	76.05
25	600	0.06	0.6	0.8	5	0.484	70.32
26	600	0.06	0.6	0.8	45	0.509	73.68
27	600	0.06	0.6	0.8	25	0.507	68.6
28	600	0.06	0.6	0.8	25	0.518	74.94
29	600	0.06	0.6	0.8	25	0.52	71.41
30	600	0.06	0.6	0.8	25	0.512	66.36
31	600	0.06	0.6	0.8	25	0.488	76
32	600	0.06	0.6	0.8	25	0.522	69

25.3 Methodology Adopted

Hybrid GRA-PCA

Hybrid GRA-PCA is adopted to obtain optimum parametric combinations, and the methodology is shown in Fig. 25.3.

25.4 Results and Discussion

The experiments were conducted as per the CCD plan of response surface method. The experimental runs with rank, GRC and GRG are depicted in Table 25.3.

Better multiple performance characteristics are obtained at larger GRG. A larger delta value indicates the higher significance of the parameter in controlling the response.

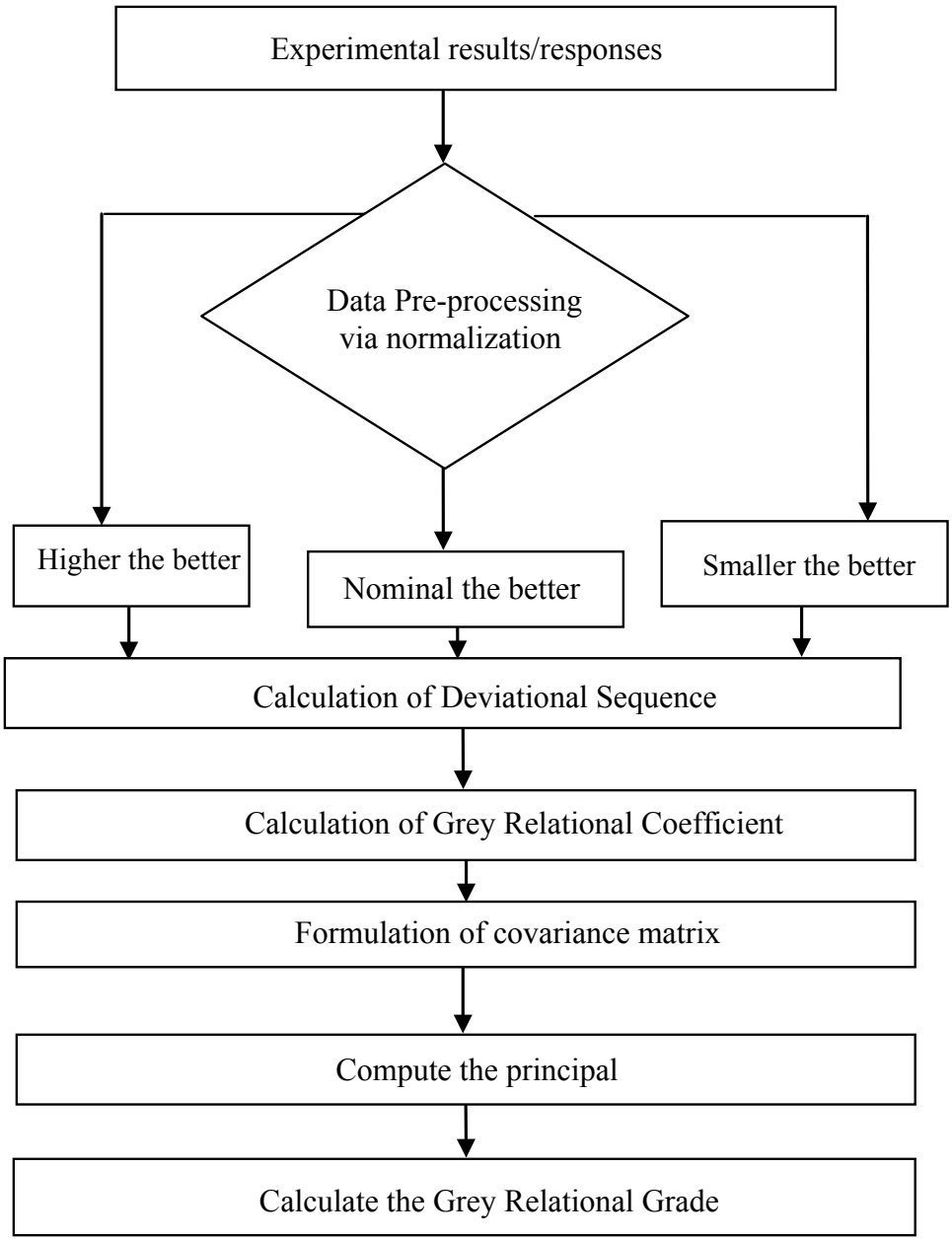


Fig. 25.3 Methodology for hybrid GRA-PCA

ANOVA is used to estimate each machining parameters percentage of contribution on multi-objective optimization. From the ANOVA analysis, it is clear that speed contribution is highest followed by nose radius, feed, depth of cut and negative rake angle as depicted in Table 25.4.

Quadratic equation for GRG after eliminating insignificant terms for 95% confidence level is given below

Table 25.3 Grey relational coefficient, grey relational grade and rank of the surface roughness and workpiece surface temperature

Expt. No.	Grey relational coefficient		Grey relational grade	Rank
	Surface roughness	Workpiece surface temperature		
1	0.45011	0.333333	0.391714	31
2	0.61329	0.498178	0.555725	7
3	0.66123	0.39276	0.52699	12
4	0.41344	0.550829	0.482126	19
5	0.40198	0.465106	0.433535	29
6	0.48915	0.661699	0.575416	5
7	0.42379	0.444661	0.434222	28
8	0.59183	0.395397	0.493607	16
9	0.54716	0.403041	0.475096	20
10	1	0.403041	0.701507	2
11	0.48915	0.702486	0.595809	4
12	0.50123	0.597622	0.549418	8
13	0.48681	0.410493	0.448643	24
14	0.93548	0.601618	0.768536	1
15	0.33333	0.427856	0.380587	32
16	0.51133	0.530653	0.520984	14
17	0.39113	0.406008	0.398565	30
18	0.64856	0.657879	0.653208	3
19	0.60237	0.451472	0.526913	13
20	0.44034	0.505246	0.472787	21
21	0.67441	0.458004	0.5662	6
22	0.56232	0.495291	0.528799	11
23	0.47319	0.411580	0.442378	26
24	0.54716	0.523343	0.535246	9
25	0.55013	0.445240	0.497678	15
26	0.48448	0.487940	0.486204	17
27	0.48915	0.426149	0.457644	23
28	0.46453	0.506143	0.485328	18
29	0.46031	0.458249	0.459274	22
30	0.47764	0.403612	0.440621	27
31	0.53846	0.522543	0.530492	10
32	0.45617	0.430441	0.443302	25

Table 25.4 ANOVA table for GRG

Source	SS	DF	MS	*F*	*P*	% C
Speed	0.09	1	0.09	140.82	<0.0001	39.63012
Feed	0.0094	1	0.0094	14.68	0.0028	4.139146
Depth of cut	0.0037	1	0.0037	5.77	0.0351	1.629238
Nose radius	0.0224	1	0.0224	35.01	0.0001	9.863496
Negative rake angle	0.0001	1	0.0001	0.2045	0.6599	0.044033
Speed * Feed	0.0346	1	0.0346	54.06	<0.0001	15.23558
Speed * Depth of cut	0.0082	1	0.0082	12.84	0.0043	3.610744
Speed * Nose radius	0.0064	1	0.0064	10	0.009	2.818142
Speed * Negative rake angle	0.0013	1	0.0013	1.96	0.1896	0.572435
Feed * Depth of cut	0.0114	1	0.0114	17.83	0.0014	5.019815
Feed * Nose radius	0.0067	1	0.0067	10.49	0.0079	2.950242
Feed * Negative rake angle	0.0025	1	0.0025	3.84	0.0759	1.100837
Depth of cut * Nose radius	0.0021	1	0.0021	3.28	0.0973	0.924703
Depth of cut * Negative rake angle	0.0016	1	0.0016	2.53	0.1399	0.704535
Nose radius * Negative rake angle	0.003	1	0.003	4.72	0.0526	1.321004
Speed * Speed	0.0058	1	0.0058	9.07	0.0118	2.553941
Feed * Feed	0.0017	1	0.0017	2.62	0.134	0.748569
Depth of cut * Depth of cut	0.0111	1	0.0111	17.38	0.0016	4.887715
Nose radius * Nose radius	0.0007	1	0.0007	1.05	0.3266	0.308234
Negative rake angle * Negative rake angle	0.0009	1	0.0009	1.43	0.2575	0.396301
Residual	0.007	11	0.0006			3.082343
Lack of fit	0.0013	6	0.0002	0.1885	0.967	0.572435
Pure error	0.0057	5	0.0011			2.509908
Total	0.2271	31				

$$\begin{aligned}\mathrm{GRG} = {} & 0.122501Xv - 0.039557Xf - 0.024805Xd + 0.061082Xr - 0.004668X\alpha \\ & + 0.056249XvXv + 0.077862XdXd - 0.185917XvXf + 0.090597XvXd \\ & + 0.079974XvXr - 0.106758XfXd - 0.081885XfXr + 0.469492\end{aligned} \tag{25.1}$$

$$S = 0.02529\ \mathrm{R{-}Sq} = 96.9\%\ \mathrm{R{-}Sq(adj)} = 91.3\%$$

From the main effects plot (Fig. 25.4) shows that grey relational grade increases significantly with an increase in speed which indexes the occurrence of optimum surface roughness and workpiece surface temperature. This is due to more thermal softening of material with an increase in speed [15, 16]. As feed increases, grey relational grade reduces due to the more feed marks which lead to higher surface

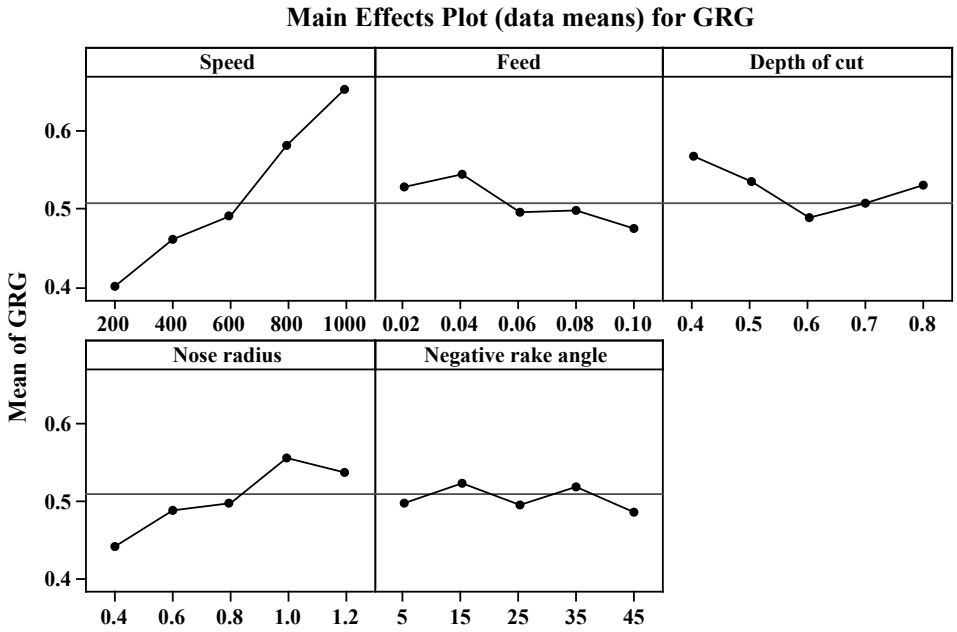


Fig. 25.4 Main effects plot for GRG

roughness and ploughing effect [17]. As the depth of cut increases, grey relational grade decreases and then increases. This is attributed to the increased cutting resistance and cutting force since chipping does not take place at nose radius which has a notable influence on the surface roughness as well as workpiece surface temperature [18]. Grey relational grade increases with an increase in nose radius. This is attributed to the increased nose radius leads to a reduction in surface roughness from general principles of metal cutting. Grey relational grade varies marginally with an increase in negative rake angle. This is due to the increased edge strength of the tool with an increase in negative rake angle causing ease in machining [18, 19]

From the main effect plot (Fig. 25.4), it is observed that the optimistic overall grey relational grade can be achieved with speed = 1000 rpm, feed = 0.04 mm/rev, depth of cut = 0.4 mm, nose radius = 1 mm and negative rake angle = 15°, respectively. GRG for the obtained optimum combination of parameters was 0.80887 estimated from Eq. 25.2 and was 5.24% more than the highest GRG in Table 25.3. Hence, the obtained values were optimum.

$$\gamma = \gamma_m + \sum_{i=1}^{q} (\overline{\gamma_j} - \gamma_m) \tag{25.2}$$

The interaction among the parameters is discussed with the help of response surface plots shown in Fig. 25.5. The optimum grey relational grade is noticed at the combinations of upper limits of speed due to thermal softening [15, 16], lower limits of feed due to less feed marks and upper limits of nose radius due to more

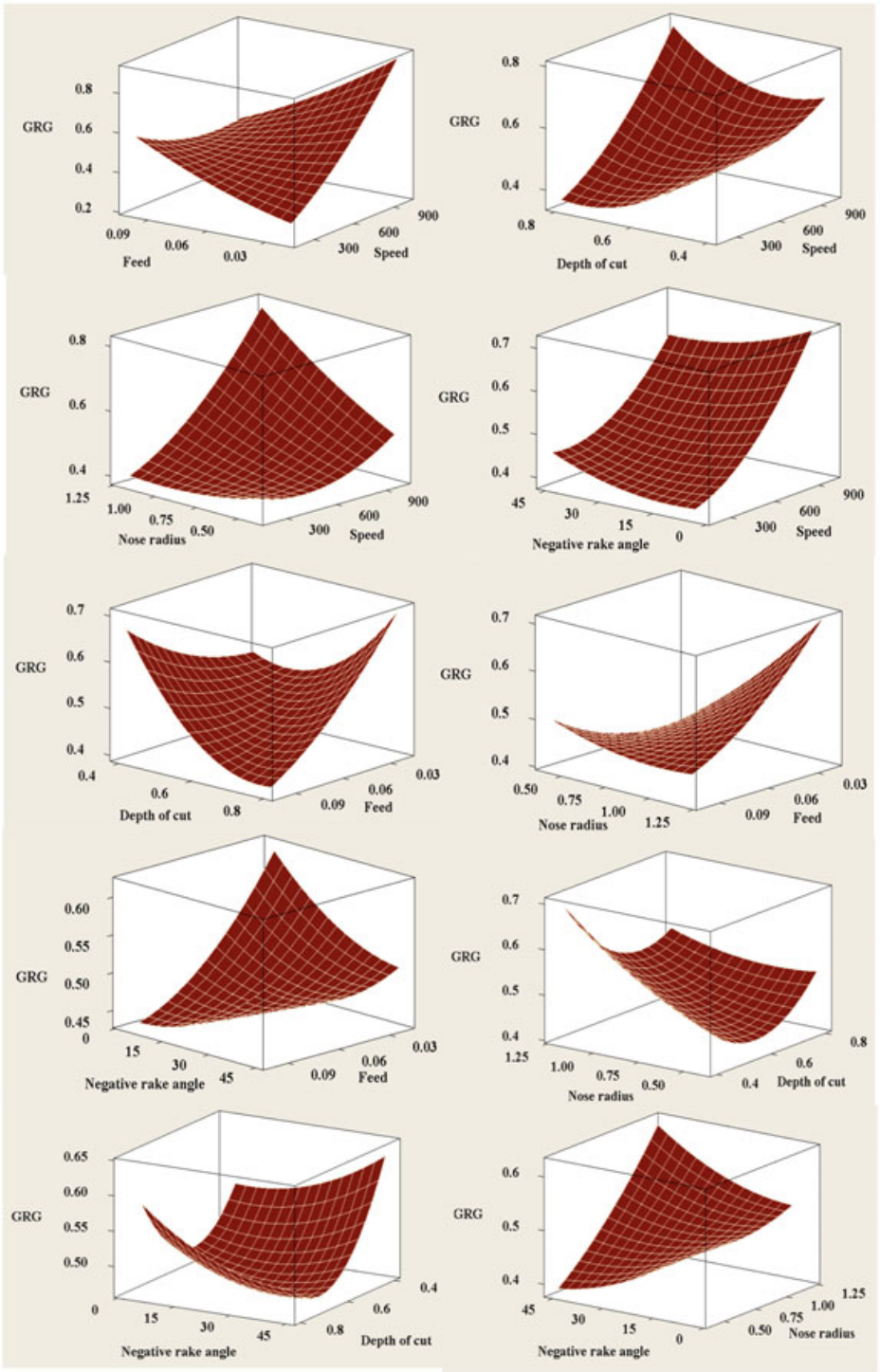


Fig. 25.5 Response surface plots [hold values were speed 600 rpm, feed 0.06, depth of cut 0.6, nose radius 0.8 and negative rake angle 25]

ploughing energy [20]. However, the interaction of depth of cut and negative rake angle with other parameters is observed to be significant, due to the variation of the developed cutting resistance and edge strength.

25.5 Conclusions

Multi-objective optimization for AISI 52100 hardened steel was performed using hybrid GRA-PCA

- The obtained optimum combination of parameters and their levels were A5B2C1D4E2, i.e. (speed = 1000 rpm, feed = 0.04 mm/rev, depth of cut = 0.4 mm, nose radius = 1 mm and negative rake angle = 15°).
- Responses were notably affected by speed followed by nose radius, feed, depth of cut and negative rake angle.
- Higher GRG is noticed during the interaction at upper limits of speed, lower limits of feed and higher limits of nose radius. However, the depth of cut and negative rake angle exhibit considerable interaction with other parameters.
- An increase in the value of predicted weighted GRG from 0.768536 to 0.80887 confirms the improvement in the performance of hard turning.

References

1. König, W., Hochschule, T., Komanduri, R., Schenectady, D., Tönshoff, H.K.: Machining of hard materials. Ann. CIRP **33**(2), 417–427 (1984)
2. Tönshoff, H.K., Arendt, C., Amor, R.B.: Cutting of hardened steel. Ann. CIRP **49**(2), 547–566 (2000)
3. Dogra, M., Sharma, V.S., Sachdeva, A., Suri, N.M., Dureja, J.: Tool wear, chip formation and work piece surface issues in CBN hard turning: a review. Int. J. Precis. Eng. Manuf. **11**(2), 341–358 (2010)
4. Huang, O.Y., Chou, Y.K., Liang, Y.S.: CBN tool wear in hard turning: a survey on research progresses. Int. J. Adv. Manuf. Technol. **35**, 443–453 (2007)
5. Bagawade, A.D., Ramdasi, P.G., Pawade, R.S., Bramhankar, P.K.: The effects of cutting conditions on chip area ratio and surface roughness in hard turning of AISI52100 steel. Int. J. Eng. Res. Technol. **1**(10) (2012)
6. Benga, G., Savu, D., Olei, A.: Influence of the cutting parameters on the surface roughness when machining hardened steel with ceramic and PCBN cutting tools. Adv. Eng. Forum **13**, 19–22 (2015)
7. Ravi Sankar, B., Umamaheswarrao, P., Nawaz Sharief, S.K., Suresh, T., Raju, R.: Parametric investigations on surface roughness of hard turned AISI 52100 steel. Appl. Mech. Mater. **867**, 171–176 (2017)
8. Suhail Adeel, H., Ismail, N., Wong, S.V., Abdul Jalil, N.A.: Optimization of cutting parameters based on surface roughness and assistance of work piece surface temperature in turning process. Am. J. Eng. Appl. Sci. **3**(1), 102–108 (2010)
9. Suhail Adeel, H., Ismail, N., Wong, S.V., Abdul Jalil, N.A.: Surface roughness identification using the grey relational analysis with multiple performance characteristics in turning operations. Arab. J. Sci. Eng. **37**(4), 1111–1117 (2012)
10. Das, S.R., Nayak, R.P., Dhupal, D.: Optimization of cutting parameters on tool wear and work piece surface temperature in turning of AISI D2 steel. Int. J. Lean Thinking **3**(2), 140–156 (2012)
11. Singh, H., et al.: Study of cutting parameters on turning using EN9. Int. J. Adv. Ind. Eng. **1**(2), 40–42 (2013)

12. Ravi Sankar, B., Umamaheswarrao, P.: Optimisation of hardness and tensile strength of friction stir welded AA6061 alloy using response surface methodology coupled with grey relational analysis and principle component analysis. Int. J. Eng. Sci. Technol. **7**(4), 21–29 (2015)
13. Ravi Sankar, B., Umamaheswarrao, P., Srinivasulu, V., Kishore Chowdari, G.: Optimization of milling process on jute polyester composite using Taguchi based grey relational analysis coupled with principle component analysis. Mater. Today Proc. **2**, 2522–2531 (2015)
14. Pradhan, M.K.: Estimating the effect of process parameters on MRR, TWR and radial overcut of EDMed AISI D2 tool steel by RSM and GRA coupled with PCA. Int. J. Adv. Manuf. Technol. **68**(1–4), 591–605 (2013)
15. Ebrahimi, A., Moshksar, M.M.: Evaluation of machinability in turning of microalloyed and quenched tempered steels: tool wear, statistical analysis, chip morphology. J. Mater. Process. Technol. **209**, 910–921 (2009)
16. Lin, H.M., Liao, Y.S., Wei, C.C.: Wear behavior in turning high hardness alloy steel by CBN tool. Wear **264**, 679–684 (2008)
17. Shihab, S.K., Khan, Z.A., Mohammad, A.A.S., Siddiquee, A.N.: Optimization of surface integrity in dry hard turning using RSM. Sadhana **39**(5), 1035–1053 (2014)
18. Suresh, R., Basavarajappa, S., Gaitonde, V.N., Samuel, G.L., Paulo Davim, J.: State-of-the-art research in machinability of hardened steels. Proc. Inst. Mech. Eng. Part B J. Eng. Manuf. **227** (2), 191–209 (2013)
19. Qian, L., Hossan, M.R.: Effect on cutting force in turning hardened tool steels with cubic boron nitride inserts. J. Mater. Process. Tech. **191**(1–3), 274–278 (2007)
20. Chou, Y.K., Song, H.: Tool nose radius effects on finish hard turning. J. Mater. Process. Technol. **148**, 259–268 (2004)

Chapter 26
Study of Cutting Edge Radius Effect on the Cutting Forces and Temperature During Machining of Ti6Al4V

Siddharam Mane, Shyamprasad Karagadde and Suhas S. Joshi

Abstract The cutting-edge radius of the tool has significant effects on the machining process, as it influences the cutting forces, stresses, and temperature at the tool–chip interface. These parameters ultimately affect the tool life and surface integrity of the finished workpiece. The presence of cutting-edge radius in the tools protects them from easily chipping off during the cutting process. A finite element based ABAQUS™ model is used to evaluate the effect of cutting-edge radius for 20, 40, and 60 μm on the cutting forces for orthogonal cutting of Ti6Al4V alloy at different cutting parameters. It was observed that the cutting-edge radius influences both the cutting and thrust forces. An increase of 4–8% and 12–14% in the cutting force and thrust force was observed when the cutting-edge radius changes from 20 to 60 μm. The temperature in the tool was increased with increasing cutting-edge radius.

Keywords Machining · Ti6Al4V · Cutting-edge radius · Finite element modeling

26.1 Introduction

In any machining process, a cutting tool with rounded cutting edges provides an enhanced tool life. The optimum cutting-edge radius for a machining process depends on the work material, tool material, and machining process parameters. Tool-edge geometry has significant effects on the cutting process, as it affects cutting forces, stresses, temperatures, the coefficient of friction, and tool–chip contact length. Nasr et al. [1] showed increasing cutting-edge radius increases both the residual stresses and temperature in the workpiece for AISI 316L steels. Liu and Melkote [2] have studied the tool-edge radius effect on the microstructure of the hard turned of AISI 52100 steel. They observed that large edge honed tools produce substantial subsurface

S. Mane · S. Karagadde · S. S. Joshi (✉)
Department of Mechanical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India
e-mail: ssjoshi@iitb.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_26

plastic flow. Wyen and Wegener [3] conducted the experimental study for orthogonal turning on Ti6Al4V with different cutting-edge radii. They found that the coefficient of friction is influenced by the cutting-edge radius of the tool.

Kim et al. [4] carried out numerical and experimental analysis considering the effect of the tool edge on the cutting process. They confirmed that a major cause for the size effect is the tool-edge radius and increase in the tool-edge radius that causes a change in the temperature distribution in the tool and changes the location of maximum temperature. Endres and Kountanya [5] showed that increasing the cutting-edge radius reduces the flank wear of the tool. Increasing cutting-edge radius increases mechanical stability, but the increasing contact length on the flank face due to the increase in cutting edge results in different thermo-mechanical loads on the cutting edge [6]. Zhao et al. [7] experimentally carried out hard turning of AISI 52100 steel under the cutting-edge radius of 20, 30, and 40. They found that the flank wear decreases with increasing radius.

Use of the cutting tool with some radius protects the edge from chipping off during machining, hence increases tool life. Considering the effect of cutting edge on the cutting forces, temperature profile, the coefficient of friction and contact length, an attempt is made to understand the effect of cutting-edge radius on the cutting forces and temperature in the machining of difficult to machine alloy like Ti6Al4V. This is done by developing a numerical model for two cutting-edge radii of 20 and 40, and 60 μm for different cutting speeds.

26.2 Numerical Modeling

A finite element model using ABAQUS™ is used for modeling and simulation of the tool and the workpiece. A 2D dynamic coupled temperature displacement model is used for simulating orthogonal cutting as shown in Fig. 26.1.

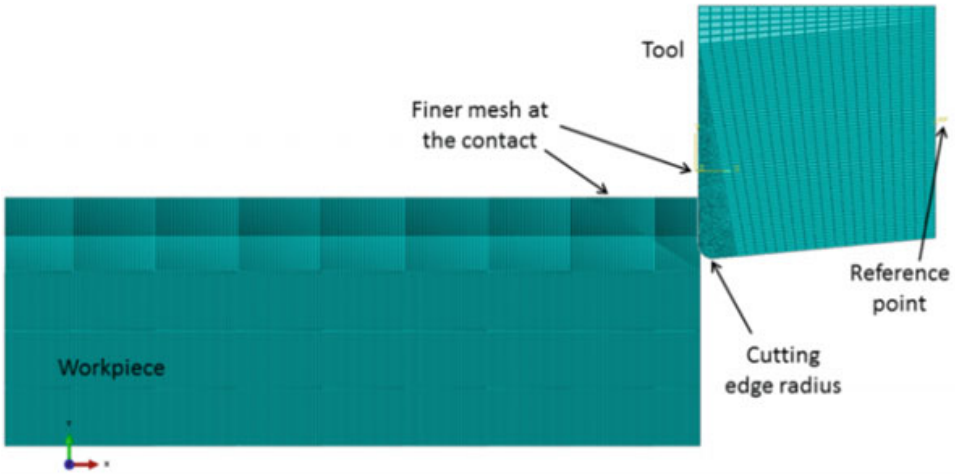


Fig. 26.1 Numerical modeling of orthogonal cutting

Table 26.1 Numerical model cutting parameters

Parameters	Value
Workpiece	Ti6Al4V
Tool	Tungsten carbide
Feed (mm/rev)	0.1
Cutting speed (m/min)	30, 50, 70, 90
Depth of cut (mm)	0.1
Rake angle	0
Cutting-edge radius (μm)	20, 40, 60

Table 26.2 Thermophysical properties of the workpiece and tool [8, 9]

Parameter	Workpiece	Tool
Specific heat (J/kg k)	670	178
Thermal conductivity (W/m k)	6.6	24
Density (kg/m^3)	4430	15700
Young's modulus (GPa)	110	705
Poisson's ratio	0.33	0.23

Table 26.3 Johnson–Cook parameter for Ti6Al4V alloy [10]

A (MPa)	B (MPa)	n	C	m	T_m	T_0
1098	1092	0.93	0.014	1.1	1630	20

The element type used for both the tool and workpiece refers to CPE4RT elements. The tool is assumed to be a rigid body since there will not be any deformation in the tool. A reference point is located at the tip of the tool for providing velocity to the tool. 20, 40, and 60 μm are the three different cutting-edge radii of the tool used in the simulation. The workpiece used is a rectangular block of 3 × 5 × 15 mm. Finer mesh size is used near the tool–workpiece contact region, as shown in Fig. 26.1. A total number of elements in the tool and the workpiece are 2392 and 39140, respectively. The cutting parameters for the modeling are listed in Table 26.1. The thermophysical properties of the tool and the workpiece are listed in Table 26.2.

Since machining is a dynamic process, there will be high strains, strain rate, and temperature involved. To accommodate this, a Johnson–Cook flow stress model was used for Ti alloy shown in Table 26.3. The Johnson–Cook fracture model is given in Table 26.4. The thermoplastic behavior of titanium alloy by Johnson–Cook model is given by,

Table 26.4 Johnson–Cook fracture model parameters for Ti6Al4V alloy [10]

D_1	D_2	D_3	D_4	D_5	ε_0
−0.09	0.25	−0.5	0.014	3.87	1

$$\sigma = [A + B\varepsilon^n]\left[1 + C\ln\frac{\dot{\varepsilon}}{\varepsilon_0}\right]\left[1 - \left(\frac{T - T_0}{T - T_m}\right)^m\right] \tag{26.1}$$

where σ is the equivalent stress, ε is the equivalent plastic strain, $\dot{\varepsilon}$ is the equivalent strain rate, and ε_0 is the reference strain rate. T, T_0, and T_m are the current temperature, room temperature, and melting temperature of the workpiece, respectively. The parameters A, B, n, C, and m are material constants, and they refer to initial yield stress, hardening modulus, strain hardening coefficient, strain rate dependency coefficient, and thermal softening coefficient, respectively.

26.3 Results and Discussion

Figure 26.2 shows an example of simulation at a cutting speed of 90 m/min. The stresses observed in the model are in the range of 1.7–1.9 GPa, which is generally observed in machining conditions.

26.3.1 Forces Evaluation

Forces generated in machining plays an important role in the heat flow and effectively the temperature distribution along the rake face of the tool. Hence, the estimation of cutting forces is important in order to understand the heat flow and temperature.

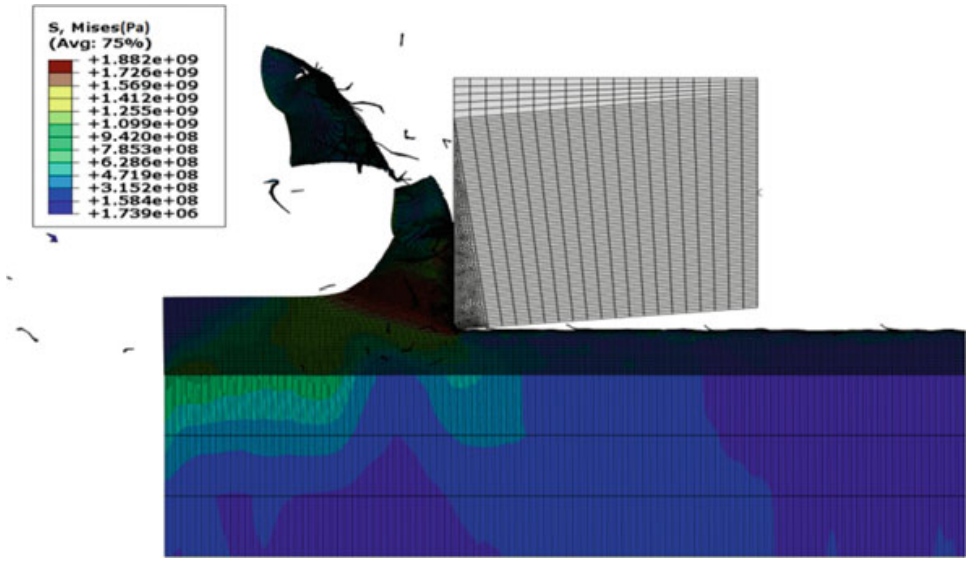


Fig. 26.2 Stresses (Pa) obtained in the finite element model

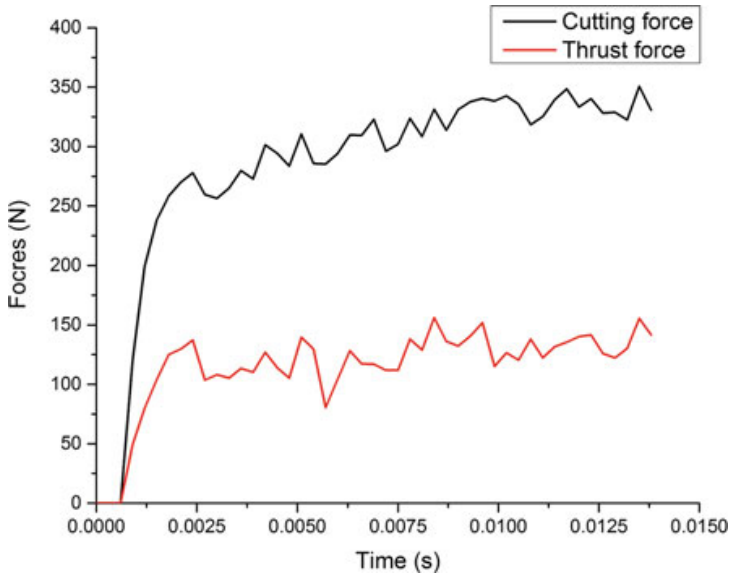


Fig. 26.3 Variation in the cutting and thrust forces with time

Figure 26.3 shows the variation in the cutting and thrust force with time. The cutting forces are large as compared to the thrust forces. These forces are evaluated for all four cutting speeds and three cutting-edge radii, as shown in Table 26.1.

The average cutting forces for cutting conditions are plotted in Fig. 26.4. As can be seen, with the change in cutting speed, change in the cutting force is around 4–5% only, whereas changing cutting edge from 20 to 60 µm changes cutting force and feed force on an average by 4–8 and 12–14%, respectively. For a radius of 20 µm, the cutting forces and thrust force are in the range of 250 and 80 N, respectively. However, for a larger radius of 40 µm, the cutting force shows some variation with cutting speeds, but the thrust force remains more ever in the range of 70–80 N. Overall, changing cutting edge influences the forces more as compared to the cutting force. The contact area at the tip of the tool is more in case of 60 µm as compared to the 20 µm. This increase in contact area increases the resistance the tool tip, and hence, the cutting forces observed in the case of 60 µm are higher.

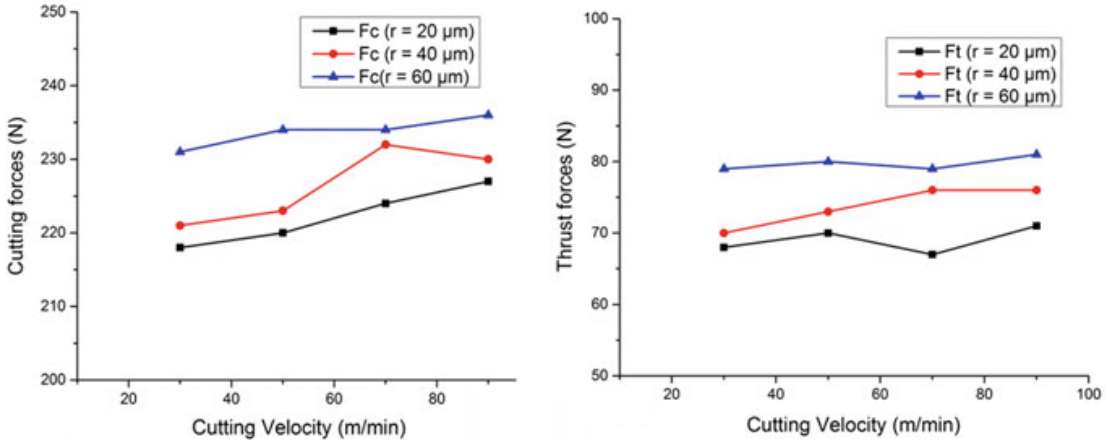


Fig. 26.4 Cutting and thrust forces for $f = 0.1$ mm/rev

26.3.2 Maximum Temperature in the Tool

Figure 26.5 shows the location of the maximum temperature developed around the tool tip. For a smaller cutting-edge radius of 20 µm, the maximum temperature location was away from the tool tip at all cutting speeds except 90 m/min where the maximum temperature was observed at the tool tip. In the case of 40 µm radius, two peaks of temperature are observed at all cutting speeds. These peaks are at the tool tip and around 1.2–1.8 mm from the tip of the tool, see Fig. 26.5b. The reason for observing the second peak on the rake face is due to the friction at the tool–chip interface. It is observed that for a larger radius of 60 µm, the maximum temperature location shifts away from the tool tip which is more toward the flank side of the tool, such a trend is observed because, in the case of larger cutting edge, the contact between the flank face of the tool and workpiece is more, can be seen in Fig. 26.5c. This causes more friction between the tool and workpiece away from the tool causing shifting of the maximum temperature location away from the tool tip. Whereas, in the case of 20 µm cutting-edge radius, the contact area was less, hence, less temperature around tool tip (Fig. 26.5a).

The temperature developed for the cutting process in the case of all the three cutting-edge radii is plotted in Fig. 26.6. The maximum temperature is observed away from the tool tip in the case of low edge radius (Fig. 26.6a), whereas, for an intermediate radius of the tool, i.e., 40 µm, the temperature obtained at the tool tip are higher as compared to 20 µm cutting radius. The maximum temperature obtained in the simulation is around 450 and 500°C for 20 and 40 µm at a cutting speed of 90 m/min, respectively. Also, the spread of the temperature is observed to be more in case of 40 and 60 µm as compared to the 20 µm cutting edge because of the large contact area available between the tool and workpiece for 40 and 60 µm edges. The maximum temperature of 550 °C was observed in the case of 60 µm

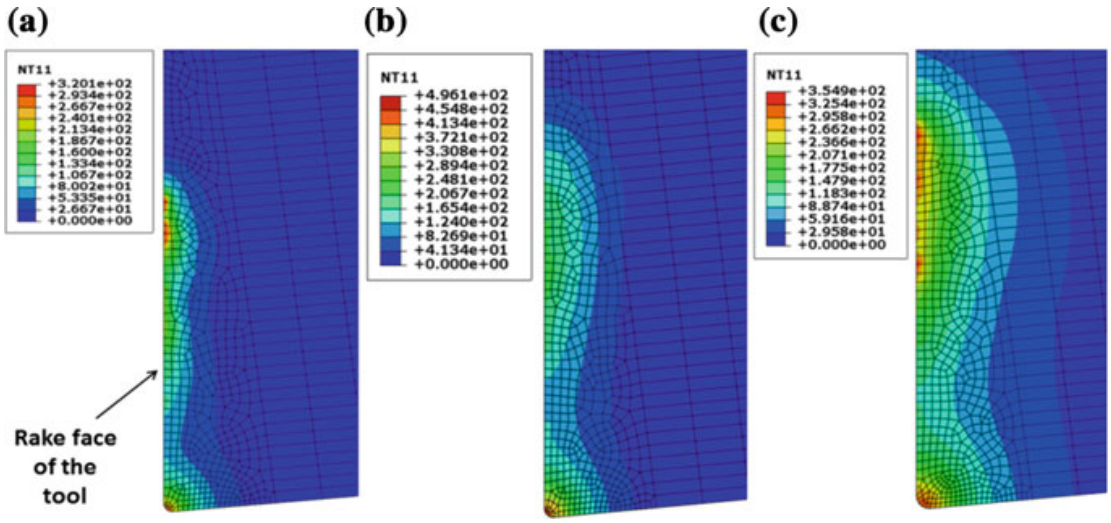


Fig. 26.5 Location of the maximum temperature generated along the rake face of the tool, **a** r = 20 µm, **b** r = 40 µm, **c** r = 60 µm

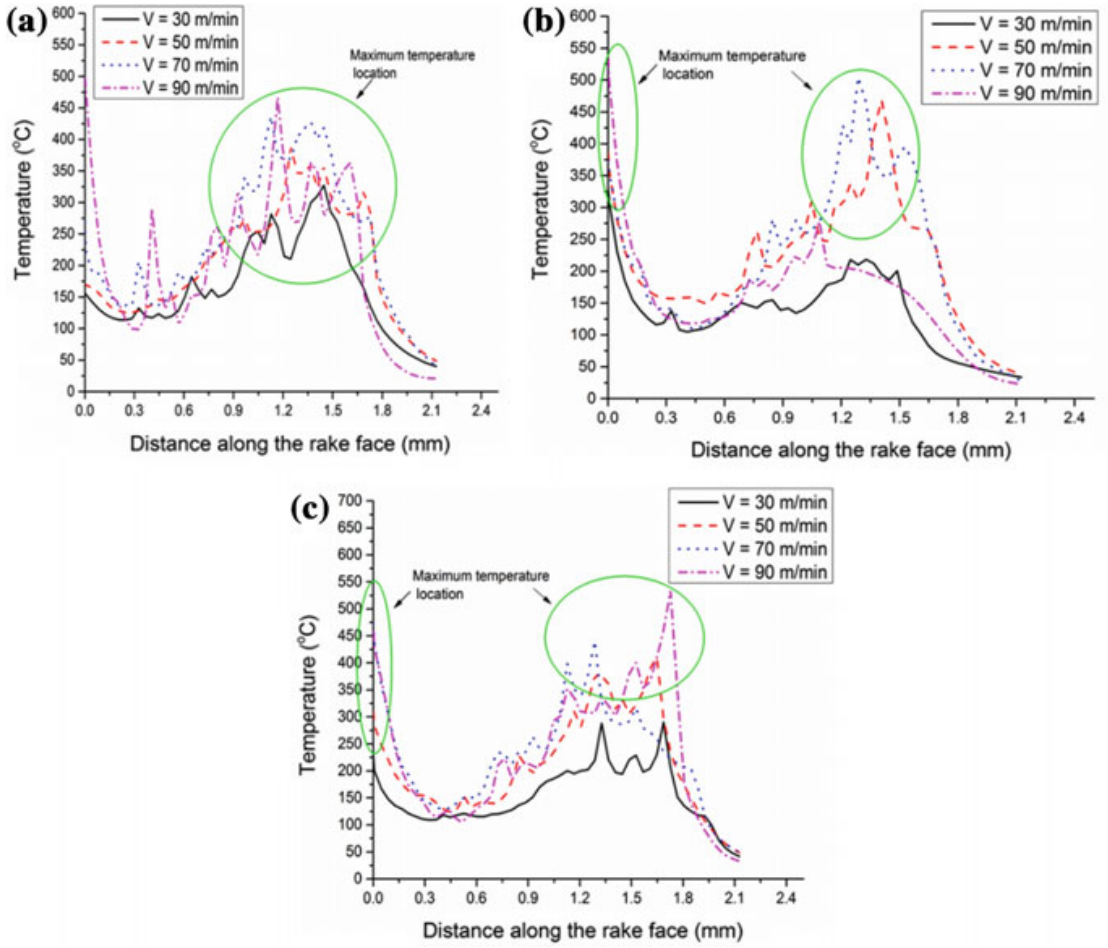


Fig. 26.6 Temperature along the rake face of the tool in the case of **a** r = 20 μm, **b** r = 40 μm, **c** r = 60 μm

cutting-edge radius at a speed of 90 m/min. Hence, the cutting-edge radius has an influence on not only the maximum induced temperature but also over the zone of temperature.

26.4 Conclusions

A FEM-based model is developed for predicting the cutting forces and temperature in the machining of Ti6Al4V alloy. It is seen that the cutting edge of the tool plays a major role in the induced temperature of machining and cutting forces. The cutting force and thrust force are seen to be increased by 4–8 and 12–14%, for a cutting-edge radius of 20 and 60 μm, respectively. The obtained temperature contours show that the maximum temperature location changes along the rake face of the tool as the cutting-edge radius changes. Although the effect of changing tool-edge radius on the forces is small, its effect on the tool temperature influences

significantly. Selecting a suitable cutting edge according to cutting conditions and material properties should be selected to improve the surface integrity and enhancing the cutting tool life.

References

1. Nasr, M., Ng, E., Elbestawi, M.: Modelling the effects of tool-edge radius on residual stresses when orthogonal cutting AISI 316L. Int. J. Mach. Tools. Manuf. **47**(2), 401–411 (2007)
2. Liu, K., Melkote, S.: Finite element analysis of the influence of tool edge radius on size effect in orthogonal micro-cutting process. Int. J. Mech. Sci. **49**(5), 650–660 (2007)
3. Wyen, C., Wegener, K.: Manufacturing technology influence of cutting edge radius on cutting forces in machining titanium. CIRP Ann. Manuf. Technol. **59**(1), 93–96 (2010)
4. Kim, K., Lee, W., Sin, H.: Finite-element analysis of machining with the tool edge considered. J. Mater. Process. Technol. **86**(1–3), 45–55 (1999)
5. Endres, W., Kountanya, R.: The effects of corner radius and edge radius on tool flank wear. J. Manuf. Process. **4**, 89–96 (2002)
6. Denkena, B., Biermann, D.: Manufacturing technology cutting edge geometries. CIRP Ann. Manuf. Technol. **63**(2), 631–653 (2014)
7. Zhao, T., Zhou, J., Bushlya, V., Stahl, J.: Effect of cutting edge radius on surface roughness and tool wear in hard turning of AISI 52100 steel. Int. J. Adv. Manuf. Technol. **91**(9–12), 3611–3618 (2017)
8. Welsch, G., Boyer, R., Collings, E.: Materials Properties Handbook: Titanium Alloys. ASM International, Ohio (1993)
9. Pierson, H.: Handbook of Refractory Carbides and Nitrides, pp. 100–117. William Andrew Publishing, Westwood, NJ (1996)
10. Leseur, D.: Experimental investigations of material models for Ti-6A1-4V and 2024-T3. Technical report (1999)

Chapter 27
Comparative Study on Effect of Inclusions and Machinability Aspects of Alloy Steels

V. Krishnaraj, K. Thillairajan and R. Rajeshshyam

Abstract In recent decades, the noticeable amount of research has been carried out in the role of non-metallic inclusions and their relationship to the machinability of various steels. This paper compares the performances of EN alloy steels for varied machinability aspects throughout the machining under dry cutting condition. Cutting speed, feed, and depth of cut were the major ruling parameters affecting the machinability of materials. The present works review the machining of different alloy steels and the role of non-metallic inclusions in them. More precisely, the effects of composition and morphology of inclusions on machinability factors such as cutting force and chip size and shape are discussed and summarized. EN 1A Pb&Te alloy steel provides lower cutting force, and good curl or breakdown chips thus provides the better machinability.

Keywords Alloy steel · Non-metallic inclusions · Machinability · Chip morphology

27.1 Introduction

Several alloy steels, which can be machined at higher cutting speeds and show longer tool life by adding additives, have been developed and used for making the screws, fasteners, and many automobile components. The performance of steels is becoming increasingly stringent. Consumer's expectation is high on the quality and performance of alloy steels. Microstructure, composition, segregation, and cleanliness are the dominant criteria in determining the properties of steel. At present, most of the steel manufacturers are unable to attain completely pure steel without any non-metallic inclusions. These are present in steel greater or lesser according to their mixture and conditions of production. Usually, these non-metallic inclusions

V. Krishnaraj (✉) · K. Thillairajan · R. Rajeshshyam
PSG College of Technology, Coimbatore 641004, Tamilnadu, India
e-mail: vkr.prod@psgtech.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_27

are the foreign substances; they affect the homogeneity of the structure and influence the machinability. Inclusions are compound materials, embedded inside the steel at the time of the manufacturing process, and it has a distinctive chemical composition and properties. Non-metallic inclusions influence the properties of steels like machinability, formability, and toughness and corrosion resistance.

In general, machinability can be represented as an optimal combination of the following factors that are low cutting forces, high metal removal rates, low tool wear rates, excellent surface finish, and good curl or breakdown of the chip. Tanaka et al. [1] investigated the machinability study of BN free-machined steel. They identified that BN free-machining steel exhibits slightly lower cutting temperature and smaller cutting forces. Jiang et al. [2] studied titanium diboride particles on machinability and machining criteria optimization during machining of TiB_2/Al MMCs. They revealed that TiB_2 particles influence on machining force, residual stress, and surface roughness. Lalbondre et al. [3] experimented the machinability of AISI 51100 and AISI 52100. They revealed that machinability of AISI 51100 is superior to AISI 52100. Anmark et al. [4] proposed that machinability of clean steels is significantly improved by resulfurization. The better machinability of the standard steel grade R is partly linked to its high content of sulphur (0.028 wt% S) which results in higher amount of MnS inclusion in the steel matrix. This MnS inclusion acts as stress risers in the primary shear region during machining. They concluded that MnS particles improved the chip formation process and prolong the cutting life. Alizadeh [5] examined the influence on the machinability of sintered steels and reported that addition of 0.3 MnS resulted in higher machinability. Joseph [6] revealed the effects of inclusions in carbon and alloy steels in their mechanical properties and cause of failure. Addition to that, inclusions, such as oxides and manganese sulphides, affect the anisotropy of these materials with respect to rolling condition. MnS particles particularly resist the anisotropy because of its elongated morphology after hot rolling.

Gupta et al. [7] quantified the micro-scale effects of inclusions on the steel sheets. They identified that the size, properties (hard and soft), and shape of the inclusions have a substantial effect on stress concentration and peak stress in a steel sheet in which inclusions are ingrained. Das et al. [8] described the comparative study on machinability of alloy steel using uncoated carbide and coated cermet inserts during hard turning. They experienced coated cermets attained a lower cutting force, minor flank wear, and reduced workpiece surface temperature. The revealed the depth of cut was the influential parameter for feed force and radial force.

The objective of the present work is on the comparative analysis of three different alloy steels. The machining operation was conducted in a dry environment based on Box–Behnken design using three factors and three levels. The machinability of the workpiece is investigated based on the output reacts like machining force and chip morphology.

27.2 Experimental Details

A complete testing procedure was carried out to examine the effect of inclusions on machinability of alloy steels when turning at several cutting conditions without coolant supply. The work materials were EN 8M free-machining steel, EN 1A alloy, and EN 1A Pb&Te alloy steel. Figure 27.1 represents the cylindrical bars of 28 mm diameter and 440 mm length made of different alloying elements which are used for the experiments.

Three parameters such as cutting speed, feed, and depth of cut are discrete at three levels during the experiments and their influences on output parameters such as cutting force and chip morphology are examined. The experimental runs are designed using Box–Behnken design without compromising the accuracy of results. Before getting into the actual machining, the rough layers were removed from the specimens.

27.2.1 Chemical Composition

A portion of the workpiece was cut using the parting tool and was polished using different grades of abrasive sheets and was subjected to the spectroscopy tests. The chemical composition obtained and major constitutions are summarized in Table 27.1.

27.2.2 Hardness and Microstructure

Hardness was determined by tests conducted using Vickers hardness testing machine wherein a square-based pyramid-shaped diamond indenter was used at 300 g load for 10 s. The diagonal of the indentation was measured and Vickers

Fig. 27.1 Workpiece material

Table 27.1 Chemical composition of alloy steels (in weight %)

Element (in mass%)	EN 8M steel	EN 1A steel	EN 1A Pb&Te
C	0.37	0.07	0.077
Si	0.21	0.057	0.057
Mn	1.27	1.28	1.25
S	0.14	0.23	0.24
Cr	0.06	0.038	0.04
Pb	–	–	>0.38
Te	–	–	0.001

Table 27.2 Vickers hardness of the alloy steels

Specimens	EN 8M steel	EN 1A steel	EN 1A Pb&Te
Vickers hardness	193	168	161

hardness was determined. The hardness values of the three steels are listed in Table 27.2.

The piece was then observed under an optical microscope after application of natal etchant. EN 8M has medium carbon steel 40% ferrite and 60% pearlite formation, EN 1A has low carbon steel with 70–75% ferrite and 25–30% pearlite, and EN 1A Pb&Te has almost ferrite structure with pearlite of 5–10%.

27.2.3 Inclusions

Inclusions are non-metallic compounds embedded in steel during the manufacturing process. It has a different chemical origin and gives different mechanical properties to steel. Inclusions have domination on several properties of steel, such as formability, toughness, and machinability and corrosion resistance. The workpieces were then observed under an optical microscope to study the distribution of inclusions on the surface as shown in Figs. 27.2, 27.3 and 27.4.

The above image indicates the Type B alumina inclusions. Literature has shown that the existence of non-metallic inclusions [particularly the hard ones such as alumina (Al_2O_3)] and complex oxides and voids has adverse effects on mechanical properties of steels. The inclusions are distributed in a particular area, and it looks like a clustered formation as shown in Fig. 27.2a. Clusters of inclusions are also unfavourable since they may result in a local drop of mechanical properties such as toughness and fatigue strength. These inclusions tend to increase the hardness of the workpiece while simultaneously increasing strength to a particular limit. This increase in hardness makes it difficult to machine.

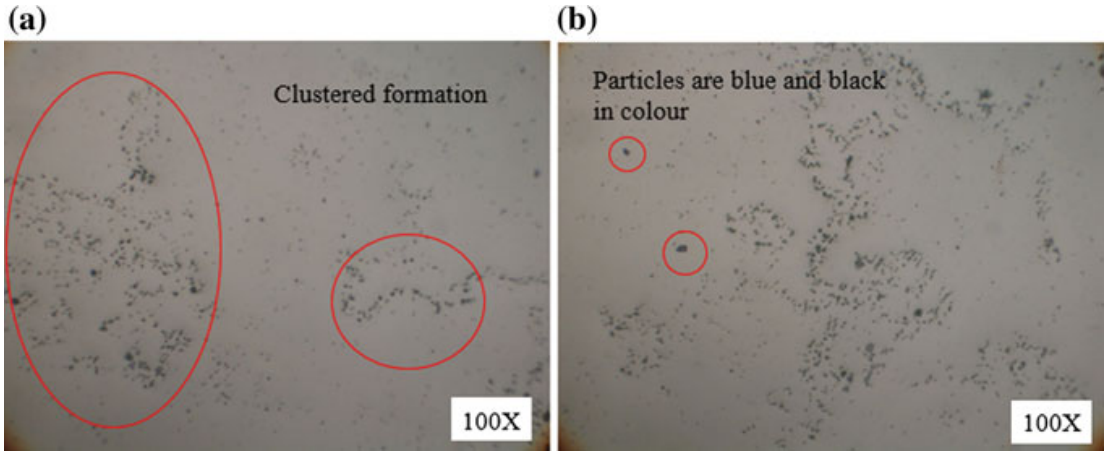


Fig. 27.2 **a** Type B–alumina type thin series 2.5 (fine size 2–9 μm). **b** Type B alumina thin series 1.5 inclusions in EN 8M free-machining steel

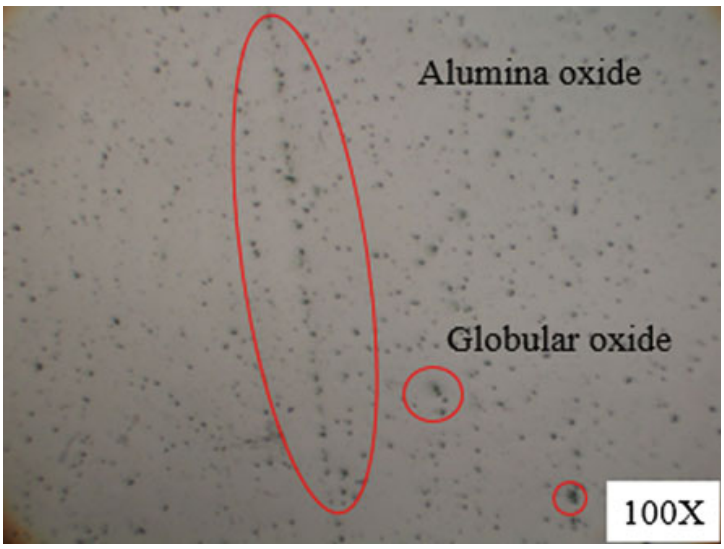


Fig. 27.3 Type D globular oxide thin series 3 (size 3–8 μm) inclusions in EN 1A steel

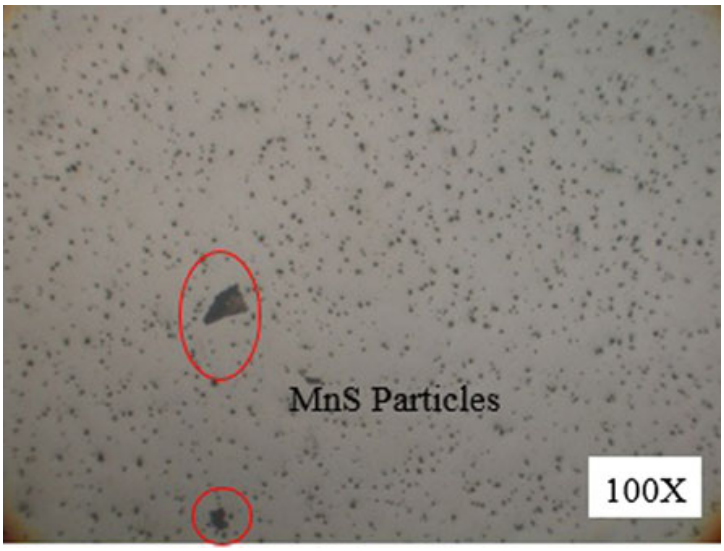


Fig. 27.4 Type D globular oxide thin series 3 inclusions in EN 1A Pb&Te steel

Figure 27.3 shows the type of inclusions formed in EN 1A Alloy steel. Globular shape of inclusions is desirable since their effect on the mechanical properties of steel is moderate. The spherical shape of globular inclusions has shown their formation in a liquid state at low content of aluminium, as indicated in Fig. 27.3. Examples of globular inclusions are manganese sulphide and oxysulphide formed during solidification in the spaces between the dendrite arms, iron aluminates, and silicates.

The MnS inclusions marked are shown in Fig. 27.4. MnS inclusions have a positive influence on the machinability, as they are soft in comparison with steel matrix and act as voids. As a result, the chips are smaller in length which leads to easier removal and declines the total power consumption during the machining.

27.2.4 Experimentation

In the present work, the multilayer CVD-coated inserts are used. The insert grade of CNMG 120408 MT TT5100 was procured from Taegu Tec. The ISO Range is (P20-P35) and shape of the insert is square (90° point angle) without any specific type of chip breaker geometry. This insert is mounted on a tool holder of ISO designation ECLNR-2020K12 as shown in Fig. 27.5. The inserts utilized in these experiments contain multilayer coatings (TiCN–Al_2O_3–TiN). Titanium nitride coatings were given at outer layer, alumina oxide, and titanium carbon nitride are as an intermediate and bottom layer of the insert.

The various control parameters adopted during the experiment are shown in Table 27.3.

Figure 27.6a, b shows the experimental setup. The turning operation was executed on a lathe of Pinacho SC 200x750 with spindle power of 4 kW and speed range of 40–2800 rpm. For measuring the cutting force, three-dimensional Kistler tool dynamometer was used.

Fig. 27.5 Cutting inserts and tool holder

Table 27.3 Control parameters and their levels for Box–Behnken design

S. No	Parameter	Unit	Symbol	Low level	Medium level	High Level
1	Cutting speed	m/min	*v*	125	162.5	200
2	Feed	mm/rev	*f*	0.1	0.15	0.2
3	Depth of cut	mm	*d*	0.5	0.75	1

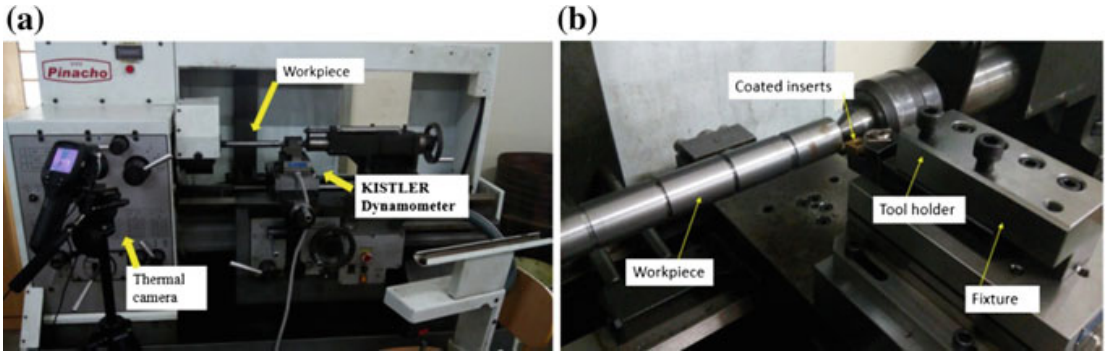


Fig. 27.6 Experimental setup

27.3 Result and Discussion

From Fig. 27.7, it was observed that the resultant cutting force of EN 8M steel was higher when compared to other two materials. When comparing to these three steels, EN 1A Pb&Te steel, accomplished the lower cutting forces at all cutting conditions owing to their low hardness and effects of lead and tellurium content in the steel.

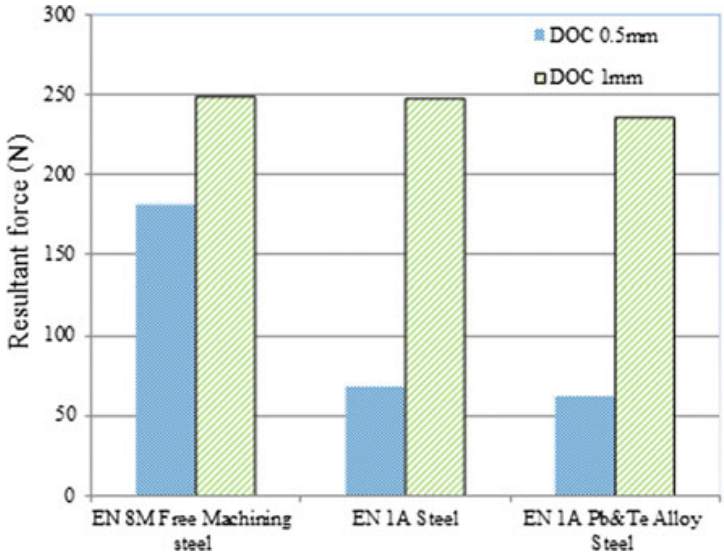


Fig. 27.7 Comparison of resultant forces for three different steels (when cutting speed 162.5 m/min, feed–0.1 mm/rev)

27.3.1 Chip Morphology

All cases which occur in a material during machining are depicted in the chip character, for example, in shape and density of chip morphology. Machining of EN 8M free machining steel obtained helical and continuous chips at higher cutting speed. On the other hand, chips were obtained blue in colour because of carbon content, which is 0.37%. During machining, an adequate amount of heat is expelled from the workpiece. Figure 27.8 shows the chips generated during machining of EN 8M steel. EN 1A alloy steel has produced the chips with light golden in colour and helical continuous one at higher cutting speed as shown in Fig. 27.9. The Lead and Tellurium alloy steel obtained small and discontinuous chips in all cutting conditions. The high content of sulphur (0.24%) results in an increased amount of MnS inclusions in the steel matrix. It is due to the effect of MnS inclusions and leads which predominantly improves the machinability. MnS inclusions act as a stress raiser in the primary shear region during the machining. As a result, they improve the chip formation process as shown in Fig. 27.10.

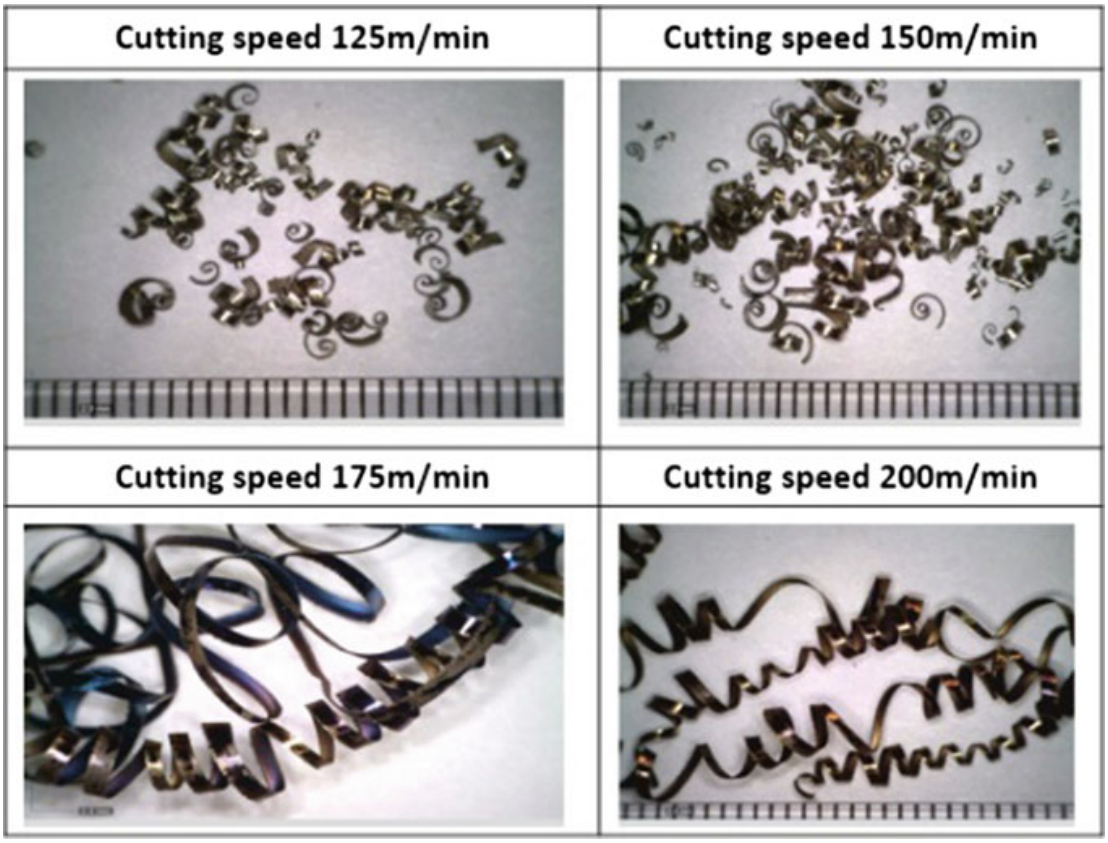


Fig. 27.8 Chips formation of EN 8M free-machining steel

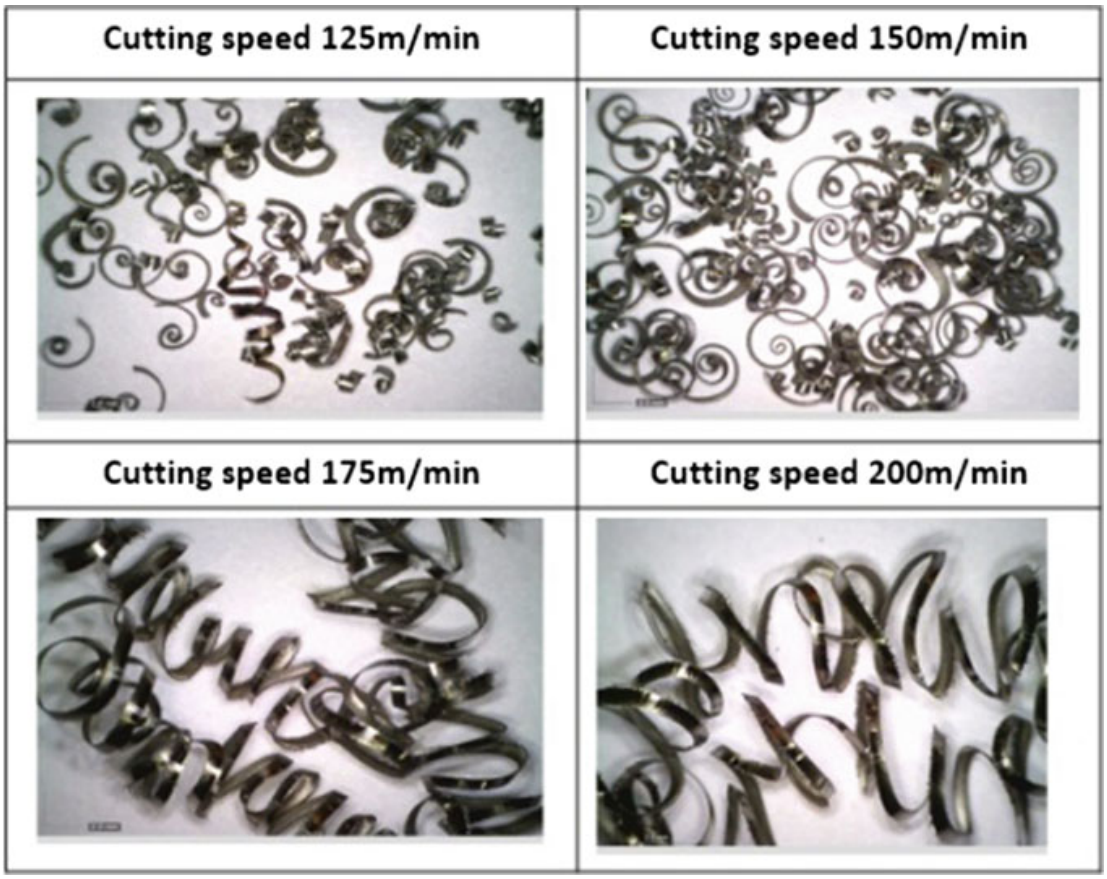


Fig. 27.9 Chip formation of EN 1A alloy steel

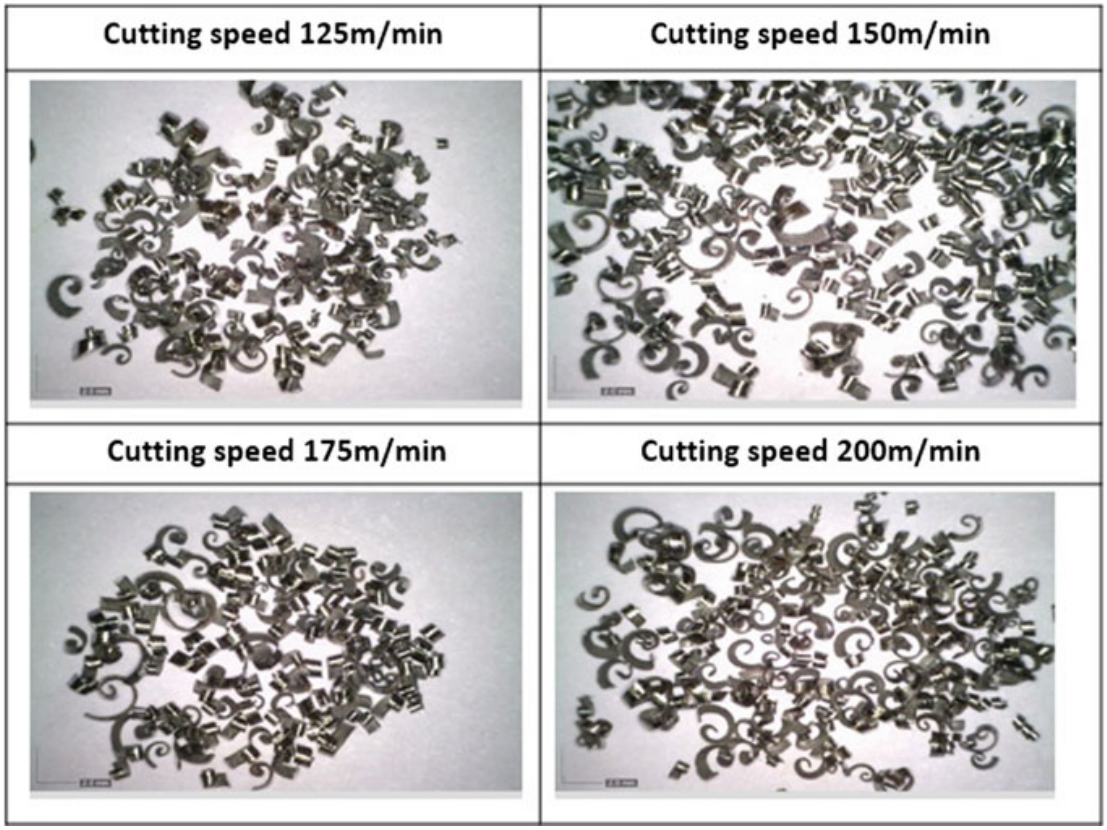


Fig. 27.10 Chip formation of EN 1A Pb&Te alloy steel

27.3.2 Discussion

In this overall experimentation, dynamic cutting forces under different cutting speeds and chip morphology were measured during dry machining of different alloy steels. Due to the chemical composition and effect of inclusions, the results were significantly varied. EN 8M free-machining steel is the hardest one because it has 40% ferrite and 60% pearlite, so it generated the highest cutting forces. The chips were segmented associated with low cutting speed, and it is lengthy and continuous at higher cutting speed.

EN 1A has a significant amount of manganese and sulphur, due to this effect, it has a lower level of cutting forces were observed. Chips were discontinuous at lower cutting speed, and it was longer and continuous at higher cutting speed.

EN 1A Pb&Te alloy has an almost ferrite structure and only 5–10% pearlite structure with large amount of inclusions (MnS Particles). Hence, it has generated the least amount of cutting forces; moreover, it has Pb content of >0.38% and Mn 1.25 with S 0.24. These three elements improve machinability. It has produced discontinuous chips in all cutting conditions when compared to other alloy steels.

27.4 Conclusion

The following outcomes are drawn out from the results gained from the current work.

- An influence of carbon content in EN 8M free-machining steel has experienced higher cutting force and obtained long and continues chips at higher cutting speeds.
- Manganese combined with sulphur (MnS) provides better machinability. The tendency of MnS inclusions in EN 1A alloy steel has generated cutting forces moderately and chips were seen curled and continues at higher cutting speed.
- As manganese, sulphur, and lead contents are increased, cutting forces were significantly decreased. Particularly in EN 1A Pb&Te alloy, lower cutting force was seen and discontinuous chips were found. This provides better machinability.

Acknowledgements We would like to thank JSW Steels Pvt. ltd. Salem, for funding this project. We also thank the faculty members of Metallurgical department and Centre for advanced tooling and precision dies, PSG Institute of Advanced Studies, Coimbatore.

References

1. Tanaka, R., Yamane, Y., Sekiya, K., Narutaki, N., Shiraga, T.: Machinability of BN free machining steel in turning. Int. J. Mach. Tools Manuf. **47**(12–13), 1971–1977 (2007)
2. Jiang, R., Chen, X., Ge, R., Wang, W.: Influence on TiB_2 particles on machinability parameter optimization of TiB_2/Al MMCs. Chin. J. Aeronaut. **31**(1), 187–196 (2017)
3. Lalbondre, R., Krishna, P., Mohankumar, G.C.: An experimental investigation on machinability studies of steels by face turning. Procedia Mater. Sci. **6**, 1386–1395 (2014)
4. Anmark, N., Karasev, A., Jönsson, P.G.: The Influence of Microstructure and Non-Metallic Inclusions on the Machinability of Clean Steels. KTH Royal Institute of Technology, WILEY-VCH, Weinheim (2015)
5. Alizadeh, E.: Factors influencing the machinability of sintered steels. Powder Metall. Met. Ceram. **47**(5–6), 304–315 (2008)
6. Joseph, M.: The effeects of sulfide inclusions on mechanical properties and failures of steel components. J. Fail. Anal. Prev. **15**(2), 169–178 (2015)
7. Gupta, A., Goyal, S., Padmanabhan, K.A., Singh, A.K.: Inclusions in steel: micro-macro modelling approach to analyse the effects of inclusions on the properties of steel. Int. J. Adv. Manuf. Technol. **77**(1–4), 565–572 (2014)
8. Das, A., Mukhopadhyay, A., Patel, S.K., Biswal, B.B.: Comparative assessment on machinability aspects of AISI 4340 alloy steel using uncoated carbide and coated cermets insets during hard turning. Arab. J. Sci. Eng. **41**, 4531–4552 (2016)

Chapter 28
Process Parameters Optimization of Electrical Discharge Machining of Al7075/SiC/WS2 by Using MCDM

Rakesh Kumar Patel and M. K. Pradhan

Abstract This paper determines the optimum process parameters of the non-conventional electrical discharge machining. The performance of the EDM machine depends upon the process parameter used. In this analysis, weight percentage, pulse current (I_p), discharge voltage (V), and pulse duration (T_{on}) are used as process parameters. The optimized output parameters are MRR, TWR, surface roughness, and radial overcut. By using face-centered composite design, nine trials were conducted on the workpiece which is made up of Al7075/SiC/WS2 hybrid composite. The trial results obtained were used in decision-making method correlation coefficient and standard deviation (CCSD) integrated approach. These results give useful information on how to control the machining parameters and accuracy of the components produced from EDM. Decision-making method used is simple, and results obtained are confirmed by conducting confirmation experiments.

Keywords Electrical discharge machining (EDM) · Material removal rate (MRR) · Tool wear rate (TWR) · Correlation coefficient and standard deviation (CCSD)

28.1 Introduction

In this article, investigation has been done on the effect of EDM parameters on machining characteristics of EDM machine for Al7075/SiC/WS_2 hybrid composite and the Multi-criteria decision-making (MCDM) method, Correlation coefficient and standard deviation (CCSD) integrated approach used for determining the weights of attributes. CCSD can be used in a multi-objective problem in various fields. EDM is a non-conventional manufacturing process, which is widely used to

R. K. Patel (✉) · M. K. Pradhan
Maulana Azad National Institute of Technology, Bhopal 462003, India
e-mail: rakesh1993patel@gmail.com

M. K. Pradhan
e-mail: mohanrkl@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_28

make more precise, accurate, and attractive components in modern manufacturing industries. In this study, the main focus was to increase material removal rate and minimize surface roughness, tool wear rate, and radial overcut for composite. The Al7075/SiC/WS_2 hybrid composite has been selected to analyze this optimization method. The various EDM input parameters such as discharge voltage (V), discharge current (I_p), pulse-on time (T_{on}), and wt% of reinforcement materials have applied to machining of a composite. The machining of particulate metal matrix composites presents a significant challenge since a number of reinforcement materials are significantly harder than the commonly used high-speed steel (HSS) and carbide tools. The reinforcement phase causes rapid abrasive tool wear, and therefore, the widespread usage of particulate metal matrix composites is significantly impeded by their poor machinability and high machining costs. Although often categorized as difficult to machine, MMCs are actually readily machinable. They form short cutting chips and require moderate cutting forces, and the range of machining parameters at which they can be machined is quite wide. However, MMCs are highly abrasive, and tools can wear rapidly.

To avoid these problems, EDM was used due to its non-contact machining characteristic. More optimum machining parameters in EDM machine are optimized to get a desirable magnitude of EDM responses on particulate metal matrix composites.

Even many attempts have been made by the experimenters for the enhancement of the MRR and improvement in accuracy, but association of CCSD method for obtaining optimal setting on EDM of Al7075/SiC/WS_2 hybrid composite has never been attempted; thus, in this research, the aforesaid unique combination has been attempted.

28.2 Literature Review

A significant number of research papers are reviewed to know the approach of yielding optimal EDM performance measures of high MRR, low TWR, and acceptable ROC. This section provides a study of each of the performance measures and the scheme for their enhancement. Roy et al. [1] used dielectric fluid mixed with Al powder and conducted experiment in EDM machine, and to analyze the experimental results, response surface methodology (RSM) is used. The input process parameters used are pulse current, pulse-on time, and concentration of Al powder on kerosene dielectric and found that due to increase in Al powder on dielectric fluid reduced the MRR and increases surface finish. Pradhan and Biswas [2] have implemented successfully by employing RSM and investigated the individual effect of the input parameters on the responses like MRR and surface roughness (R_a). Further, Pradhan and Biswas [3] have explored the effects of I_p, T_{on}, duty cycle (T_{au}), and V on various response parameters using two neuro-fuzzy and one neural network model. Mandal et al. [4] used an artificial neural network model and multi-objective optimization approach for obtaining the best process parametric

combinations for a better productivity and fine surface finish. Dvivedi et al. [5] used Al6063 SiCp metal matrix composite to check machineability in EDM. They investigated the optimum value of MRR and TWR by considering the input variables T_{on}, T_{off}, and I_p. Kanagarajan et al. [6] had investigated the EDM performance of tungsten carbide/cobalt cemented carbide by considering the input parameters such as electrode rotation, I_p, T_{on}, and flushing pressure. Jaharah et al. [7] found that on machining of tool steel on EDM, MRR increases and TWR decreases with increase in the I_p and T_{on}. Dhar et al. [8, 9] used Al–4Cu–6Si alloy—10 wt% SiCp composites on EDM machine—to find the effect of I_p, T_{on}, and V on MRR, TWR, and gap on EDM. To develop the relationship between the machining parameters used a second-order nonlinear mathematical model and found that MRR, TWR, and gap on EDM increase with increase in I_p and T_{on}.

In past, substantial efforts are carried out to increase output parameters such as MRR, TWR, and accuracy of EDM process. The important issue is to pick the process parameters such as I_p, T_{au}, T_{on}, V, flushing pressure, dielectric fluid, and polarity in such a way that increases MRR and accuracy, and simultaneously ROC, TWR, and surface roughness should diminish.

28.3 Description of the Experiments

The experiments were conducted under various parameter settings of discharge current, pulse-on time, and discharge voltage. The Minitab 18 software was used for Taguchi's method to modeling alternatives and attributes of EDM parameters. Finally, the soft computing techniques were employed for modeling of MRR, TWR, radial overcut, and R_a.

Calculate the responses of EDM machine

$$\text{MRR}\left(\frac{\text{mm}^3}{\text{min}}\right) = \frac{\text{Weight Loss (gm)} \times 60}{\text{Density of Sample}\left(\frac{\text{gm}}{\text{mm}^3}\right) \times \text{Machining Time (s)}}$$

$$\text{TWR}\left(\text{mm}^3/\text{min}\right) = \frac{\text{Weight Loss(gm)} \times 60}{\text{Density of Tool(gm/mm}^3) \times \text{Machining Time (s)}}$$

$$\text{Radial Over Cut (ROC)} = \frac{D_1 - D_2}{2}$$

where D_1 = right side reading on digital micrometer and D = left side reading on digital micrometer. The Toolmaker microscope was used to calculate ROC. The digital surface roughness meter is used to calculate surface roughness of surface of cavity on composite surface generated by EDM machine (Table 28.1), where 0.75, 1, and 1.5 represent the wt% of WS_2 in composite with constant 10 wt% of SiC (Fig. 28.1).

Table 28.1 Experimental results

Run	Wt. (%)	I_p (A)	T_{on} (μs)	V (V)	MRR (mm^3/min)	TWR (mm^3/min)	ROC (μm)	R_a (μm)
1	0.75	8	75	50	38.144	0.146	5.646	7.100
2	0.75	10	100	60	17.764	0.091	5.706	11.328
3	0.75	12	150	70	56.070	0.119	5.678	12.510
4	1	8	100	70	18.332	0.292	5.693	8.041
5	1	10	150	50	19.132	0.378	5.626	10.381
6	1	12	75	60	39.220	0.363	5.637	8.844
7	1.5	8	150	60	18.033	0.231	5.625	10.323
8	1.5	10	75	70	18.287	0.047	5.678	7.120
9	1.5	12	100	50	39.716	0.102	5.581	9.723

Fig. 28.1 Samples machined on EDM machine

28.4 Analysis Method

Wang et al. [10] used CCSD multi-criteria decision making method. Suppose there are n decision alternatives A_1, …, A_n to be evaluated in terms of m attributes $O_1, O_2, \ldots, O_m$, it forms a decision matrix which can be represented as $X = (x_{ij})$ $n \times m$ where x_{ij} is the performance value of A_i with respect to O_j. The m attributes

can be categorized into two groups: beneficial and non-beneficial. The beneficial attributes are those whose values are always larger the better. The non-beneficial attributes are those whose values are smaller the better. Due to the attributes in different units, the decision matrix $X = (x_{ij})n \times m$ needs to be normalized to eliminate its dimensional units. The commonly used normalization method is:

for beneficial attributes

$$z_{ij} = \frac{x_{ij} - x_j^{\min}}{x_j^{\max} - x_j^{\min}} \tag{28.1}$$

for non-beneficial attributes

$$z_{ij} = \frac{x_j^{\max} - x_{ij}}{x_j^{\max} - x_j^{\min}} \tag{28.2}$$

z_{ij} Normalized value of output parameter.

Now remove attribute O_j one by one from the set of attributes one at a time, and see its impact on decision making. When O_j is removed, the overall assessment value of each alternative can be calculated as

$$d_{ij} = \sum_{k=1,k\neq j}^{m} z_{ik} w_k \quad i = 1, 2, \ldots, n$$

The correlation coefficient between the attributes and the score values can be expressed as

$$R_j = \frac{\sum_{i=1}^{n} \left(z_{ij} - p_j\right)\left(d_{ij} - q_j\right)}{\sqrt{\sum_{i=1}^{n} \left(z_{ij} - p_j\right)^2 \left(d_{ij} - q_j\right)^2}}$$

where

$$p_j = \frac{1}{n}\sum_{i=1}^{n} z_{ij}, \quad q_j = \frac{1}{n}\sum_{i=1}^{n} d_{ij}$$

The attribute given more weightage which have bigger standard deviation value and smaller weightage which have lower standard deviation value in comparison.

From the above calculations, the weights of attributes can be determined as

$$w_j = \frac{\sigma_j \sqrt{1 - R_j}}{\sum_{j=1}^{m} \sigma_j \sqrt{1 - R_j}} \quad j = 1, 2, \ldots, n \tag{28.3}$$

R_j Correlation coefficient between the values of O_j.
σ_j Standard deviation of the values of O_j.

The score of each decision alternative can be determined by

$$d_{ij} = \sum_{j=1}^{m} z_{ij} w_j \quad i = 1, 2, \ldots, n \tag{28.4}$$

28.5 Results and Discussion

To get the optimum result of the machining parameters and to reduce the manufacturing cost of composite product, the CCSD method was used to optimization of EDM parameters for machining of hybrid composite (Al7075 + SiC + WS2) and find the rank of alternatives. According to the CCSD integrated approach, the weight of the attributes is 0.2528, 0.2518, 0.2118, and 0.2737 for the attribute MRR, TWR, ROC, and R_a, respectively, which was calculated by using Eq. (28.3).

The overall assessment value of each decision alternative can be computed by

$$d_i = \sum_{j=1}^{m} z_{ij} w_j, \quad i = 1, 2, \ldots, n$$

From Table 28.2, values shown are normalized value which was calculated by using Eq. (28.1) for beneficial and Eq. (28.2) for non-beneficial attribute, and score was calculated by using above Eq. (28.4) (Fig. 28.2).

From the above graph, alternative 9 has highest score, so it is best suitable alternative according to MCDM. EDM of hybrid composite Al7075/SiC/WS2 with input value of alternative 9 gives optimum value of output parameters.

Table 28.2 Overall score assessment and ranking

Alternatives	MRR (mm^3/min)	TWR (mm^3/min)	ROC (mm)	R_a (μm)	Score	Rank
1	0.5320	0.7013	0.4784	1	0.6913	2
2	0	0.8680	0	0.2184	0.287	8
3	1	0.7819	0.2192	0	0.5009	4
4	0.0148	0.2593	0.1	0.8260	0.3171	7
5	0.0357	0	0.6384	0.3935	0.2553	9
6	0.5601	0.0469	0.5528	0.6776	0.4555	5
7	0.0070	0.4454	0.6464	0.4042	0.3698	6
8	0.0136	1	0.2192	0.9963	0.5835	3
9	0.5730	0.8346	1	0.515	0.7197	1

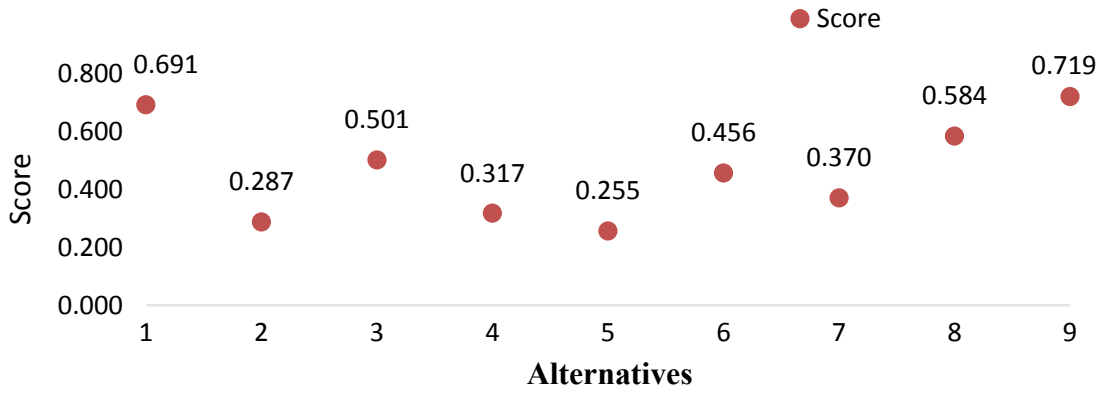


Fig. 28.2 Score of the alternatives

28.6 Conclusion

This study proposes optimization of the EDM input process parameters. CCSD method shows that the alternative 9 is the best alternative. The optimal condition out of the given alternatives determined by the CCSD methods is as follows: pulse current 12 A, pulse-on time 100 µs, weight percentage 1.5, and discharge voltage 100 V. These results will help in reducing the machining cost, error, and time consumption in the machining process and also enhance the surface quality and efficiency. This study gives the optimum input process parameter of EDM to achieve desired MRR, TWR, surface roughness, and radial overcut, in the die sinking of Al7075/SiC/WS2 hybrid composite.

References

1. Roy, C., Syed, K.H., Kuppan, P.: Machinability of Al/10% SiC/2.5% TiB2 metal matrix composite with powder-mixed electrical discharge machning. Procedia Technol. **25**, 1056–1063 (2016)
2. Pradhan, M.K., Biswas, C.K.: Neuro-fuzzy and neural network-based prediction of various responses in electrical discharge machining of AISI D2 steel. Int. J. Adv. Manuf. Technol. **50** (5–8), 591–610 (2010)
3. Pradhan, M.K., Biswas, C.K.: Investigating the effect of machining parameters on EDMed components a RSM approach. J. Mech. Eng. **7**(1), 47–64 (2010)
4. Mandal, D., Pal, S.K., Saha, P.: Modeling of electrical discharge machining process using back propagation neural network and multi-objective optimization using non-dominating sorting genetic algorithm-II. J. Mater. Process. Technol. **186**(1–3), 154–162 (2007)
5. Dvivedi, A., Kumar, P., Singh, I.: Experimental investigation and optimization in EDM of al 6063 SiCp metal matrix composite. Int. J. Mach. Mach. Mater. **3**(3–4), 293–308 (2008)
6. Kanagarajan, D., Karthikeyan, R., Palani kumar, K., Sivaraj, P.: Influence of process parameters on electric-discharge machining of WC/30%Co composites. Proc. Inst. Mech. Eng. Part B J. Eng. Manuf. **222**(7), 807–815 (2008)

7. Jaharah, A.G., Liang, C.G., Wahid, S.Z., Ab, R.M., Che, H.C.: Performance of copper electrode in electrical discharge machining (EDM) of AISI H13 harden steel. Int. J. Mech. Mater. Eng. **3**(1), 25–29 (2008)
8. Dhar, S., Purohit, R., Saini, N., Sharma, A., Kumar, G.H.: Mathematical modeling of electric discharge machining of cast Al–4Cu–6Si alloy–10 wt.% SiCP composites. J. Mater. Process. Technol. **194**(1–3), 24–29 (2007)
9. Purohit, R., Krishna, C.M., Rana, R.S., Kumar, V., Patel, K.K.: Optimization of electric discharge machining parameters for Al–3 weight% nano SiCp composites using copper electrode. Trends Ind. Mech. Eng. **105** (2016)
10. Wang, Y.M., Luo, Y.: Integration of correlations with standard deviations for determining attribute weights in multiple attribute decision making. Math. Comput. Model. **51**(1–2), 1–2 (2010)

Chapter 29
CNC Turning and Simulation of Residual Stress Measurement on H13 Tool Steel

Ravi Butola, Qasim Murtaza and Ranganath M. Singari

Abstract This paper deals with the measurement of residual stress produced during turning operation. CNC turning is a machining process to remove unwanted material from the workpiece to produce desired rotational parts. Turning is a widely used cutting process in the manufacturing industry, and thus, continuous research is going on to optimize it because of cutting high-performance dynamics, workpiece materials, and cutting tools. Turning operation was performed on H13 tool steel workpiece dimensions 110 × Ø22 mm on CNC Lathe machine tools. H13 tool steel is used for manufacturing of friction stir welding and processing tool. The main aim of the current study is to determine the residual stress using PulsetecμX-360n portable stress analyzer setup, and Abaqus 6.14 software was used to simulate the turning process. Experimental tests were performed on H13 tool steel materials which are in form of FSW/FSP tool, which consists of different pin profiles. The experimental outcomes are then compared with the outcomes obtained from the simulation. Experimental results validate the simulation results for both the tool pin profiles which are lying within the acceptable range.

Keywords Turning · H13 tool steel · Simulation

R. Butola · Q. Murtaza (✉) · R. M. Singari
Department of Mechanical, Production and Automobile Engineering,
Delhi Technological University, Delhi 110042, India
e-mail: qasimmurtaza@dce.ac.in

R. Butola
e-mail: ravibutola33855@gmail.com

R. M. Singari
e-mail: ranganath@dce.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_29

29.1 Introduction

Turning operation was performed on H13 tool steel workpiece on CNC Turning Center Lathe. The dimension of workpiece was 110-mm-long cylinder with 22 mm diameter. The cutting tool insert used was uncoated tungsten carbide CNMG 120404–THM for turning operation, and it was a dry turning since no cutting fluid was used. In turning processes, large plastic deformation takes place; the prediction and control of the material fracture is a critical issue. In order to surface finish and integrity of the new component, the damage and fracture of the materials should be predicted. H13 tool steel is used for manufacturing of friction stir welding and processing tool which are used for making welding joints and surface composite. Friction stir processing, a development based on friction stir welding (FSW), is a solid-state processing technique for microstructure modification [1]. FSW/FSP tool consists of shoulder and different pin profiles, i.e., cylindrical and taper cylindrical pin profiles have been used to process the aluminum-based surface composite at different tool rotation and tool traveling speeds. Cylinder pin profile is simplest in construction considering the machining operations performed in making the cylindrical pin profile. The parameters considered for the tool design are shoulder diameter, pin diameter, and length of the pin. Huang et al. [2] studied the impact of residual stress at the initial stage and turning-induced stresses after deformation of aluminum and found that turning-induced residual stresses are the primary cause of distortion. Plate deformation is increased by coupling of compressive residual stress and machining-induced residual stress while deformation is decreased by the coupling action of tensile initial residual stress and machining-induced residual stress. Wang et al. [3] carried out a study on residual stress in machined surface of monocrystalline silicon based on molecular dynamics simulation. They found that on the increasing depth of cut, the maximum tensile residual stress decreases while compressive residual stress increases underneath the surface and same happened when the tool rake angle changed from positive to negative. Fergani et al. [4] predicted the residual stress in multi-pass milling machining using thermomechanical stress model using Neumann–Duhamel law. A regeneration algorithm was proposed that considers the combination of initial stresses and material cyclic hardening in residual stress calculations. Sharma and Butola [5] reviewed Taguchi's technique, response surface methodology, and gray fuzzy algorithm for optimal values of machining parameters in the CNC process of turning. They found that material removal rate increased with increase in feed rate and depth of cut, while the increase in cutting speed decreased it. Huang et al. [6] investigated the influence of residual stress on the monolithic component deformation using FEM as well as experimental work and found that initial residual stress in the blank was the main factor of deformation for three-frame monolithic beam and deformation is increased by coupling of initial and turning residual stresses. Johnson and Cook [7] have been developed that the flow stress is defined as a function of strain, strain rate, and temperature. Pres et al. [8] have been reported a measurement, simulation, and the comparison between the cutting force and geometrical feature of the chip. Wang

et al. [9] have been presented the frequency and cutting force for Ti6Al4V after that they compare their simulation and experimental result, and they found good agreement between them. Umbrello et al. [10] develop a hardness flow stress simulation model of turning processes based on Johnson–Cook equation. Chaudhary et al. [11] have reviewed that different FSP tool pin is used at different parameter. FSP can be widely used in metals such as Al alloy that is lightweight and can be used extensively in various applications. The significance of this research will have a major and robust impact on the manufacturing industry. The prediction of residual stresses by using Abaqus software could become an opportunity to undertake by the tool manufacturers by changing or improving the cutting tool design prior to manufacturing and same supported by experimental data using PulsetecμX-360n portable stress analyzer setup.

The objective of this research is to the measurement of residual stress using PulsetecμX-360n portable stress analyzer setup. After CNC turning, the process on H13 tool steel materials which are in form of FSW/FSP tool has manufactured two different tool profiles, i.e., cylindrical and taper cylindrical pin and shoulder. To simulate the turning of H13 tool steel workpiece and determine stresses, Abaqus 6.14 software was used.

29.2 Experimental Method

29.2.1 Selection of Tool Materials

Each material is characterized by a unique set of physical, mechanical, and chemical properties which can be treated as attributes of a specific material. The selection of material is primarily dictated by the specific attributes that are required for manufacturing friction stir processing tool. H13 tool steel is used for manufacturing of friction stir welding and processing tool which are used for making welding joints and surface composite. The tool requisite for friction stir processing must resist thermal fatigue cracking under hot working operations. H13 tool steel is versatile chromium–molybdenum hot work steel that is widely used in manufacturing of friction stir welding and processing tool. The workpiece in form H13 tool steel rod of diameter 22 mm and length 110 mm, processed by Mahindra and Mahindra Steel, was used in this investigation. The chemical composition of these materials in weight percentage as supplied by the producer is given in Table 29.1.

Table 29.1 Chemical composition of tool material-H13 tool steel

C	Mn	Si	Cr	Mo	V
0.40	0.40	1.00	5.25	1.35	1.00

29.2.2 Machining Methodology

Machining methodology consists of machining techniques and equipment used for machining or shaping the tools. In this research, two workpieces were made with different pin profiles like cylindrical and taper cylindrical. For machining the tool material into the desired tool pin profile, CNC Lathe machine tool was used. Primarily, turning was used for making the shoulder of the tool. For a cylindrical pin, the further turning operation was performed. In the first step, turning reduced the diameter from 22 to 19.95 mm for the entire length of the tool because 19.95 mm diameter is the maximum diameter that can accommodate the FSW machine in our laboratory. In the second step, turning reduced the diameter from 19.65 to 7.22 mm for pin length of 3.50 mm. The taper cylindrical pin has generated a taper of length 3.50 mm with diameter linearly varying from 7.22 to 3.50 mm, and tool pin length depends upon the thickness of the workpiece to be processed (Fig. 29.1; Table 29.2).

29.2.3 Measurement Condition of H13 Tool Steel

Measurement of residual stress in cylindrical pin and taper cylindrical pin was performed on the following conditions: pitch 50 µm, X-ray irradiation time (setup) 30 s, X-ray irradiation time (Meas.) 30 s, X-ray irradiation time (Max) 101 s, X-ray

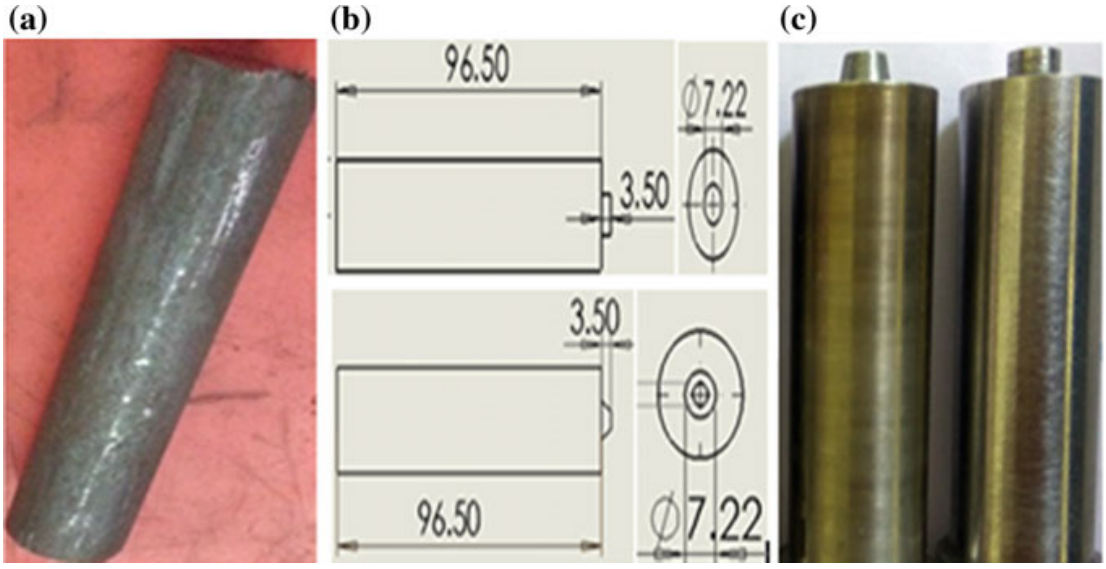


Fig. 29.1 **a** H13 tool steel rod, **b** dimension of tool, and **c** after turning

Table 29.2 Input machining parameter

Cutting speed (S)	Feed rate (F)	Depth of cut
1500 (Rpm)	0.10 (mm/rev)	1 (mm)

tube current 1.00 mA, X-ray tube voltage 30.00 kV, X-ray incidence angle 35.0°, offset of alpha angle 0°, X-ray wavelength (K-Alpha) 2.29093 A (Cr), X-ray wavelength (K-Beta) 2.08480 A (Cr), total measurement count 6631, oscillation count 7, detection sensitivity 26.1% (224,879), peak strength (Ave) 209 k, and temperature 31.56 °C.

29.2.3.1 Sample Information

The basic idea was to measure residual stresses on the machined workpieces of H13 tool steel as follows: wavelength K-Alpha, diffraction angle (2θ) = 156.396°, diffraction lattice angle (2η) = 23.604°, interplanar spacing (d) = 1.170, diffraction plane (h, k, l) = 2, 1, 1, crystal structure BCC, Young's modulus (E) = 224.000 GPa, Poisson's ratio (v) = 0.280, sigma (x) stress constant (K) = −465.097 GPa, tau (xy) stress constant (K) 380.985 GPa, and sigma (y) stress constant (K) = −2091.661 GPa.

29.2.3.2 Measurement of Residual Stresses

Residual stress measurement was performed in Precision Manufacturing Lab at Delhi Technological University. After turning operation of H13 tool steel, the next step is to the measurement of residual stress has been done on the specimen using PulsetecμX-360n portable stress analyzer setup. Two specimens were performed for measurement of residual stress, i.e., cylindrical and taper cylindrical profiles. Mechanical residual stresses are generally produced because of the plastic deformation in the non-uniform manner caused during manufacturing operations. Residual stresses can be made naturally during the operation or may be produced in order to generate the desired stress profile in the component [12] (Fig. 29.2).

29.3 Modeling of Cylindrical and Taper Cylindrical Tool

The following steps will illustrate in different stages involved in the FEA of the turning process to create a cylindrical and taper cylindrical tool profile discussed below

- Standard and explicit model was selected upon opening the Abaqus 6.14 package. This allows for explicit integration scheme to solve highly nonlinear systems with many complex contacts under transient loads.
- In the Part module, a 3-D deformable part with solid extrusion feature is chosen. This part is the H13 tool steel workpiece.
- Figure 29.3a shows a circular sketch of 22 mm radius which is drawn. This was the raw dimension of the workpiece. Figure 29.3b shows the part is extruded to

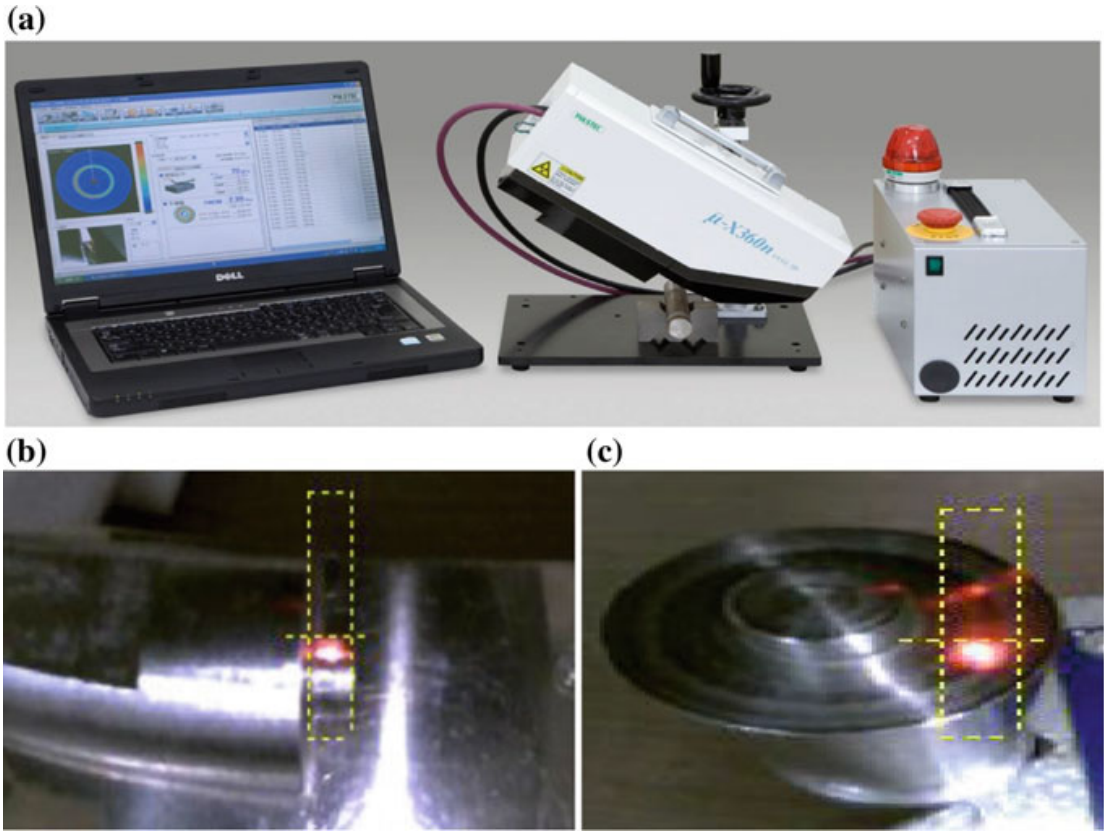


Fig. 29.2 **a** PulsetecμX-360n portable stress analyzer set up, **b**, **c** image captured by stress analyzer

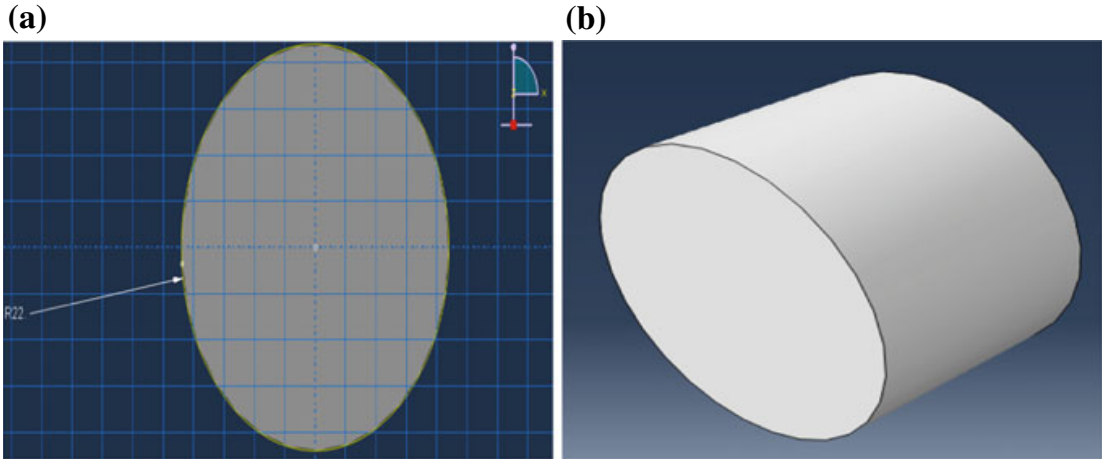


Fig. 29.3 **a** Geometry for workpiece with R 22 mm and **b** shown 30-mm-extruded cylindrical workpiece

30 mm depth. Note that it is shorter than the actual length of the workpiece. This is because we are only interested in the region most affected by the turning process.

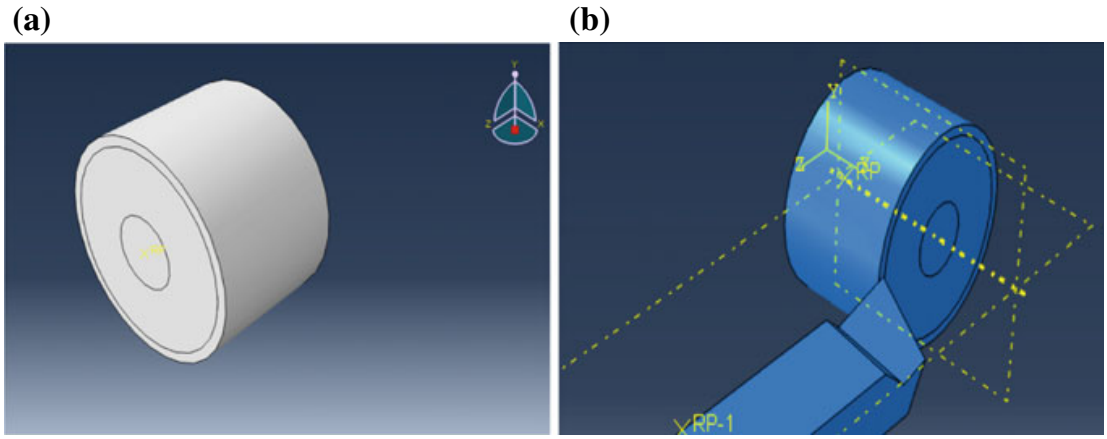


Fig. 29.4 **a** Partitioned regions in the workpiece and **b** assembly

- Figure 29.4a shows the part is partitioned into three different regions. It will help in improving the mesh generation. Since there is step cutting involved, different regions can be meshed to be of finer/coarser quality.
- In the Materials module, the material is defined by entering the properties of H13 tool steel under the *Material Behaviors* section. Properties like conductivity, ductile damage, shear damage with damage evolution, density, elastic, expansion, and plastic behavior are defined. Johnson-Cook damage criterion was followed.
- Figure 29.5a shows the Mesh module, 3D stress family of explicit, and linear geometric order, and Hex element type is assigned to the workpiece. Element deletion is chosen since we are not interested in chip flow and only the stresses on the workpiece surface. In the interaction module, friction contact with the coefficient of 0.4 is used. In the load module, boundary conditions are added. Here, the angular velocity to the workpiece, the depth of cut, and feed rate are given to the tool.

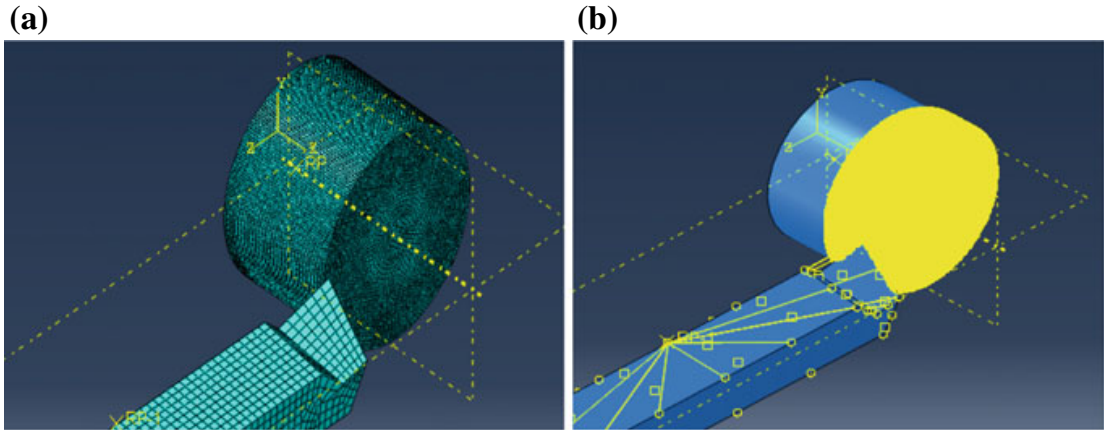


Fig. 29.5 **a** Meshing and **b** constraints added to the workpiece and tool

The procedure followed for the stress analysis of the cylindrical tool pin was similar to the procedure described in the previous section. The only exception lies in the load module, where the boundary conditions are modified so that the tool travels in a straight path parallel to the workpiece axis.

29.4 Results and Discussion

As these stresses play a very important role in the working performance of the component, it becomes quite necessary to examine the magnitude and nature of the residual stresses contained in the component. Stresses present can be tensile or compressive in nature but generally tensile nature of residual stresses is undesired as it can because quench cracking, stress-induced corrosion cracking, and fatigue failure. Compressive nature is desired one as it gives opposite effect of the tensile stresses on the component.

The developed FE analysis model was validated by comparing the simulation of the turning process with the experimental data. Figures 29.5 and 29.6 show the simulation of turning process by using Abaqus 6.14 software. Due to advancement in technology and new methodology, now residual stresses due to turning operation can be visualized using computer simulation. Experiments have done by using PulsetecμX-360n portable stress analyzer machine. Residual stress name is given to the stress which is left in the component after removal of the load. It remains almost in every rigid structure generated due to every operation performed on it such as metallurgical, thermal, and mechanical during manufacturing. Simulation of residual stress is an important process when H13 tool steel is used for CNC turning operation. Measurement of residual stresses due to CNC turning can now be visualized via computer simulations due to advancements in technology.

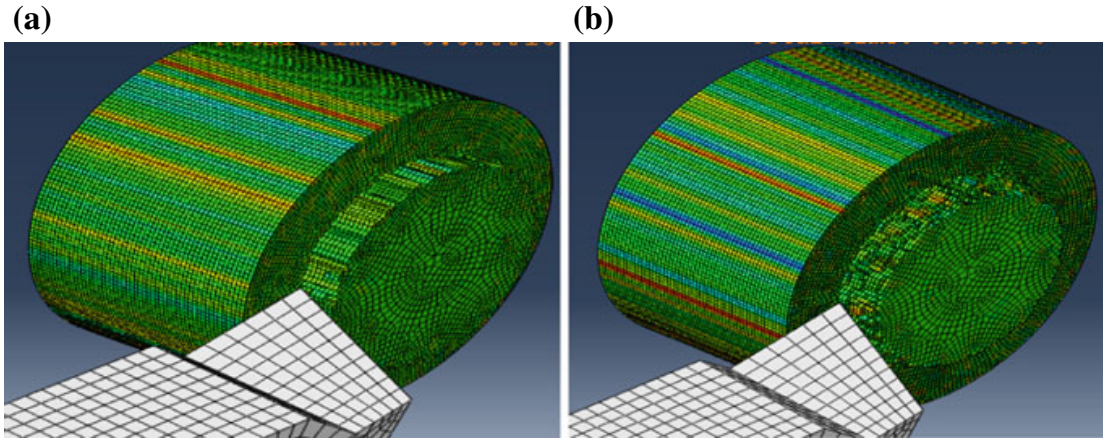


Fig. 29.6 **a** Stress contours of cylindrical tool pin and **b** stress contours of taper cylindrical tool pin

Unfortunately, turning processes have not yet been simulated to perfection in software due to a large number of deformations and variables involved. To overcome this problem, FE analysis is carried out in Abaqus/CAE 6.14. Experimental data is obtained from by using a PulsetecμX-360n portable stress analyzer setup. Experimental and simulation results show the percentage error lying within the acceptable range for both the cylindrical and taper cylindrical tool of H13 tool steel (Figs. 29.7, 29.8, 29.9 and 29.10; Tables 29.3 and 29.4).

The outcome of the residual stress on the periphery of a cylindrical pin and shoulder was observed to be compressive 166 and 280 MPa, respectively. The standard deviation was 16, 19 MPa, respectively. The outcome of the residual stress on the periphery of the taper cylindrical pin and shoulder was observed to be compressive 301 and 444 MPa, respectively. The standard deviation was 17 and 18 MPa, respectively.

After the measurement of residual stress, an experimental data and simulation data for cylindrical tool pin were found to be compressive 223 MPa and compressive 240 MPa, respectively, so the percentage error was observed that 7.62%. For taper cylindrical tool pin, an experimental data and simulation data were found to be compressive 372 MPa and compressive 398 MPa, respectively, so the percentage error observed was 6.98%.

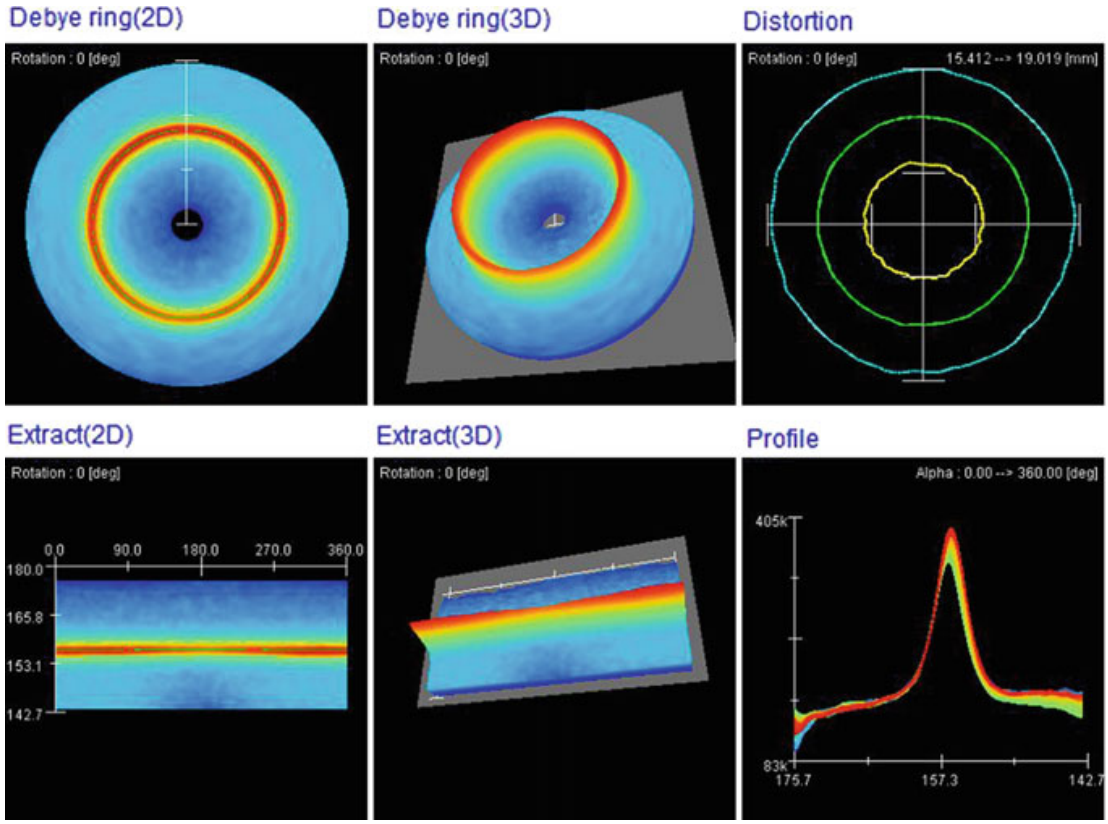


Fig. 29.7 Cylindrical tool pin. **a** Debye ring 2D and 3D, **b** distortion, **c** extract 2D and 3D, and **d** profile

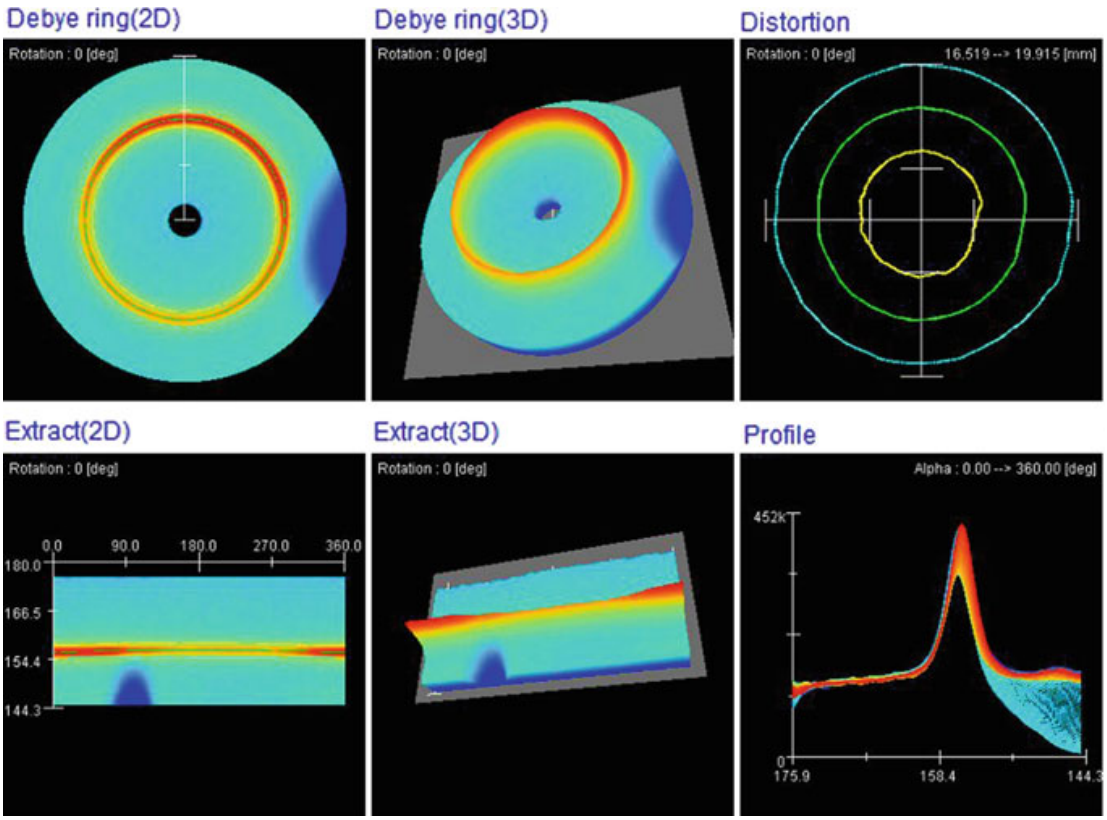


Fig. 29.8 Cylindrical tool shoulder. **a** Debye ring 2D and 3D, **b** distortion, **c** extract 2D and 3D, and **d** profile

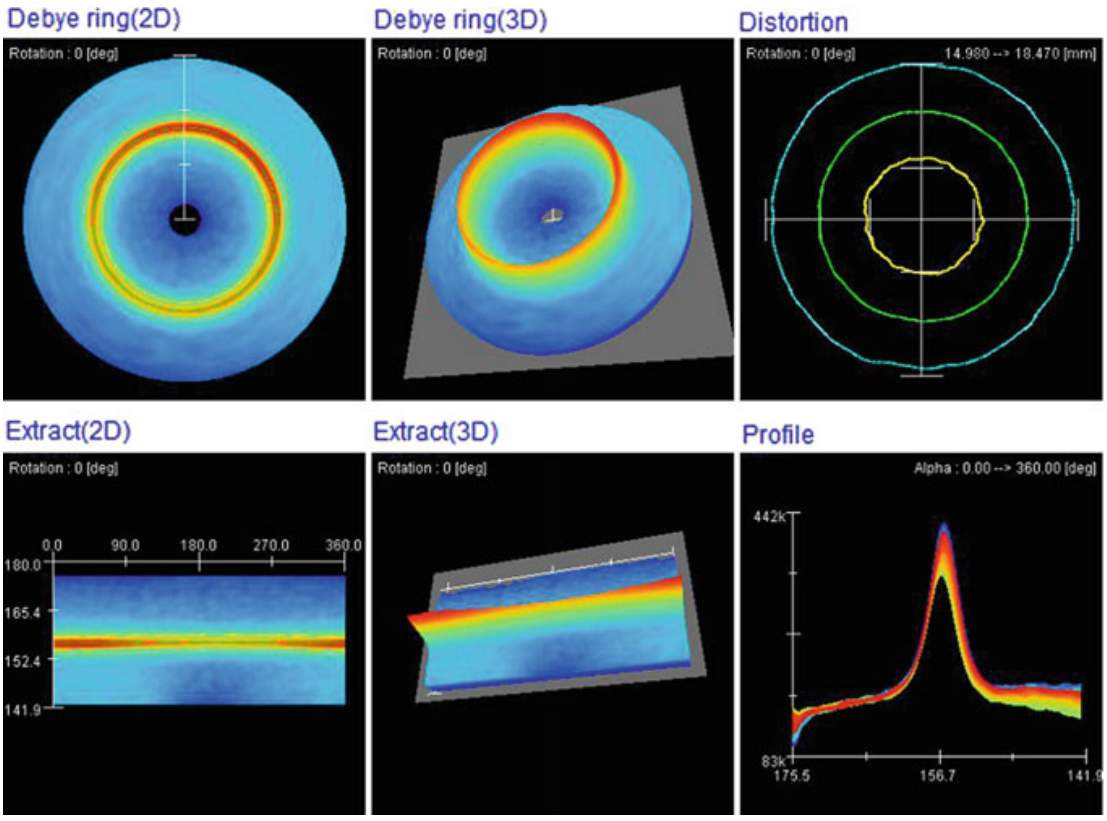


Fig. 29.9 Taper cylindrical tool pin. **a** Debye ring 2D and 3D, **b** distortion, **c** extract 2D and 3D, and **d** profile

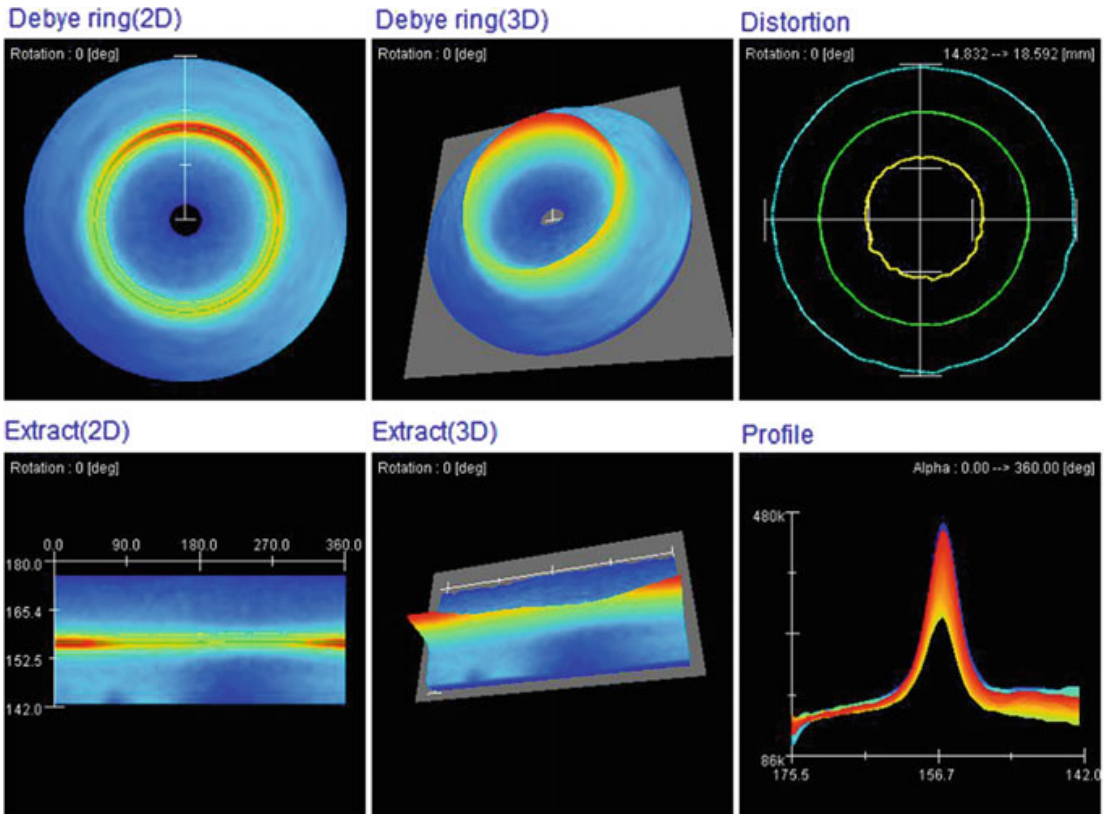


Fig. 29.10 Taper cylindrical tool pin. **a** Debye ring 2D and 3D, **b** distortion, **c** extract 2D and 3D, and **d** profile

Table 29.3 Measurement of residual stress in different tool pin and shoulder

S. No.	Pin profile	Residual stress [MPa] (Absolute)	Average stress [MPa] (Absolute)
1	Cylindrical pin	166	223
2	Cylindrical shoulder	280	
3	Tapered cylindrical pin	301	372
4	Tapered cylindrical shoulder	444	

Table 29.4 Summary of results and error %

S. No.	Tool pin profile	Residual stress [MPa] (Absolute) from simulation	Average stress [MPa] (Absolute) from experiment	Error %
1	Cylindrical pin	240	223	7.62
2	Tapered cylindrical pin	398	372	6.98

29.5 Conclusions

In the current research, the study was carried out on simulation of residual stresses stimulated by turning of H13 tool steel. Two different tool pin profiles, i.e., cylindrical and taper cylindrical is used.

After analyzing the result of an experimental data and simulation data, the percentage error was observed that 7.62% for cylindrical tool profile and 6.98% for taper cylindrical tool profile and the percentage error can be rectified with the help by improving the mesh size quality that could be helpful for production of new material. Experimental results validate the simulation results for both the tool pin profiles which are lying within the acceptable range.

References

1. Mishra, R.S., Mahoney, M.W., McFadden, S.X., Mara, N.A., Mukherjee, A.K.: High strain rate super plasticity in a friction stir processed 7075 Al alloy. Scr. Mater. **42**, 163–168 (2000)
2. Huang, X., Sun, J., Li, J.: Effect of Initial residual stress and machining-induced residual stress on the deformation of aluminium alloy plate. J. Mech. Eng. **61**(2), 131–139 (2015)
3. Wang, Y., Shi, J., Ji, C.: A numerical study of residual stress induced in machined silicon surfaces by molecular dynamics simulation. Appl. Phys. A **115**, 1263–1279 (2014)
4. Fergani, O., Jiang, X., Shao, Y., Welo, T., Yang, J., Liang, S.: Prediction of residual stress regeneration in multi-pass milling. Int. J. Adv. Manuf. Technol. **83**, 1153–1160 (2016)
5. Sharma, V., Butola, R.: Optimization of machining parameters in CNC turning of hybrid metal matrix composites using different techniques: a review. IJARI **5**, 78–82 (2017)
6. Huang, X., Sun, J., Li, J.: Finite element simulation and experimental investigation on the residual stress-related monolithic component deformation. Int. J. Adv. Manuf. Technol. **77**, 1035 (2015)
7. Johnson, G.R., Cook, W.H.: A constitutive model and data for metals subjected to large strains, high strain-rate, and temperatures. In: Proceedings of the 7th International Symposium on Ballistics, The Hague, The Netherlands, pp. 1–7 (1983)
8. Preś, P., Skoczyński, W., Jaśkiewicz, K.: Research and modeling workpiece edge formation process during orthogonal cutting. Arch. Civ. Mech. Eng. **14**(5), 622–635 (2014)
9. Wang, B., Liu, Z.: Investigations on the chip formation mechanism and shear localization sensitivity of high-speed machining Ti6Al4V. Int. J. Adv. Manuf. Technol. **75**(5), 1065–1076 (2014)
10. Umbrello, D., Rizzuti, S., Outeiro, J., Shivpuri, R., M'Saoubi, R.: Hardness-based flow stress for numerical simulation of hard machining AISI H13 tool steel. J. Mater. Process. Technol. **199**, 64–73 (2008)
11. Chaudhary, A., Kumar, D.A., Goel, A., Butola, R., Ranganath, M.S.: The mechanical properties of different alloys in friction stir processing: a review. Mater. Today Proc. **5**, 5553–5562 (2018)
12. Brien, E.O.: Residual Stress—Friend or Foe A Civil Aircraft Perspective, Vol. 1, pp. 13−21. ICRS-6, Oxford, UK, 10–12 July (2000)

Chapter 30
Small-Scale Machining Simulations

Anuj Bisht, Anish Roy, Uday S. Dixit, Satyam Suwas and Vadim V. Silberschmidt

Abstract Molecular dynamics (MD) and single-crystal plasticity finite-element method (CP FEM) are approaches used to simulate the micro-machining process. At such small-length scales, anisotropic behaviour of material becomes important; the two methods can essentially capture it. Therefore, it is important to understand the fundamental principle behind these methods as well as their capabilities and limitations in order to select the scheme to simulate the micro-machining process. In this paper, the fundamentals of MD and CP FEM are introduced in brief. The applicability of the respective method is further illustrated with the help of examples from the literature. Thereafter, the two methods are compared and discussed in terms of their various aspects and capabilities. This discussion should enrich the reader and assist their choice of an appropriate simulation method.

Keywords Micromachining · CP FEM simulation · Molecular dynamics

30.1 Introduction

There is a push for miniaturization of devices with improved performance in electronic and sensor industry. Precision micromachining in such components is unavoidable often to facilitate assembly of parts. The behaviour of such devices under the process of micro-machining is not completely understood. This is primarily due to the change in behaviour of material at lower length scales, which in turn depends on intrinsic and extrinsic material parameters. Firstly, grain size is the most important intrinsic material parameter, which influences the materials

A. Bisht · A. Roy (✉) · V. V. Silberschmidt
Loughborough University, Loughborough LE11 3TU, UK
e-mail: a.roy3@lboro.ac.uk

U. S. Dixit
Indian Institute of Technology Guwahati, Guwahati 781038, India

S. Suwas
Indian Institute of Science, Bangalore 560012, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_30

behaviour. The strength of the material increases with the decrease in grain size from micron to submicron and nanometre regime, known as "Hall–Petch (HP) effect" [1, 2]. The increase in strength upon grain size reduction is mainly attributed to the increase in grain boundary area, which acts as barrier for dislocation motion [3]. Secondly, the size of the sample is an extrinsic material factor, and the strength of the material increases in decreasing the sample dimension to micron and submicron scale [4–6], an effect similar to the HP effect. This phenomenon is mainly attributed to the starvation of dislocation (plasticity carriers) on reducing the sample dimension. Apart from the mentioned phenomena, material is observed to drift from the conventional behaviour when it is machined at micron and nanoscale. Similar behaviour is also observed under indentation at various length scales, known as "indentation size effect (ISP)" [7–9]. Due to the unconventional behaviour of material showing scale effects, there is a need to study the process to better understand the phenomenon occurring during micro-machining process.

Experimental studies are instrumental in reflecting the actual events. However, due to the limitations of method involved in experiments, the phenomena can be studied only after its completion and the evolution of defects in the material is speculated based on the findings. In this regard, simulation methods are quite instrumental to study the evolution of material. Finite-element methods are used conventionally to simulate various material processes [10–12], which includes deformation behaviour, using bulk phenomenological-based models. They are good in simulating phenomenon where the scales are larger compared to the inherent dislocation slip phenomena. Johnson–Cook model is a popular model used and is for isotropic material. However, materials are crystalline and, hence, inherently anisotropic in nature [13]. The anisotropic nature becomes prominent under the event of limited small-scale deformation process, where the deformation length scale becomes comparable to deformation features like slip bands due to dislocation slip activity. To account for the crystal anisotropy, there are two simulation methods adopted for simulating the micro-machining processes, namely single-crystal plasticity (CP) finite-element method (FEM) [14–16] and molecular dynamics (MD) [17–19]. The two mentioned simulation methods are based on different principles.

The CP FEM modelling takes into account the anisotropy in material in a FEM framework. On the other hand, MD is an atomic-level simulation and is primarily based on Newton's second law of motion to describe the trajectory of atoms with time. Atomic potential function is used to describe the interatomic interaction between atoms. This essentially means the scale of the two simulations is very different.

The aim of the present paper is to introduce to the reader the two methods used for simulation of micro-machining process. They are described using two examples from the literature. The principle behind the two methods is introduced in brief in their respective section. The example also focuses on the information, which can be extracted from the respective simulation.

30.2 Simulation Techniques

Here, we discuss the basics of the two simulation methods, namely molecular dynamics (MD) simulation and finite-element (FE)-based crystal plasticity (CP) numerical models. In this section, we briefly discuss the two techniques, which are commonly used to predict machining outcomes in the small-length scales.

30.2.1 Molecular Dynamics Simulation

Molecular dynamics (MD) is a versatile simulation technique, which finds applications in the simulation of material behaviour at very small-length scales (scale of molecules). In MD, the behaviour of an ensemble of molecules of the material is studied when subjected to an externally imposed thermo-mechanical load. It proceeds by solving the many body problem involving atoms at each time step. The principle relation in MD is based on Newton's second law of motion. Newton's equation of motion for an atom i in the system is given by

$$f_i = m_i a_i, \tag{30.1}$$

where f_i is the net force exerted on the atom i, m_i is its mass and a_i is the acceleration of atom i. The net force f_i on atom i is the summation of individual forces (f_{ij}) exerted by atoms j on atom i.

The net force f_i is given by the gradient of the potential energy (V):

$$f_i = -\nabla_i V. \tag{30.2}$$

Potential functions are used to describe the interacting forces (f_{ij}) among the atoms i and j. Lennard-Jones (LJ) potential [20, 21], analytical bond-order potential (ABOP) [22], embedded atom method (EAM) [23] and Tersoff potential [24] are important potential functions to name a few. Specific potential functions were developed for specific bond types including ionic bond, metallic bond, covalent bond, valence bond and others. The choice of an appropriate potential function depends on the material and problem at hand [25]. For example, the EAM potential is useful for describing the interaction among interacting metallic atoms, while ABOP is trustable for describing covalent materials.

Once the force is obtained at the current time step t, the position (r_i) and velocity (v_i) of atom i can be obtained for the next time step $t + \delta t$ by time integration step. Various time integration schemes are available to march ahead in time. Velocity Verlet integrator [26] and rRESPA multi-timescale integrator [25] are popular time integration schemes. The position (r_i) and velocity (v_i) of atom i at time step $t + \delta t$ as given be Velocity Verlet algorithm are

$$x_i(t+\delta t) = x_i(t) + v_i(t)\Delta t + \frac{1}{2}a_i(t)\delta t^2, \tag{30.3}$$

$$v_i(t+\delta t) = v_i(t) + \frac{a_i(t) + a_i(t+\delta t)}{2}\delta t. \tag{30.4}$$

30.2.2 Crystal Plasticity Finite-Element Simulation

The classical crystal plasticity theory is discussed in the present section. When a material is deformed plastically, it experiences both elastic and plastic strain. A deformation gradient ($\boldsymbol{F}$), in a continuum sense, can be decomposed into an elastic part ($\boldsymbol{F}_e$) and plastic part ($\boldsymbol{F}_p$):

$$\boldsymbol{F} = \boldsymbol{F}_e\boldsymbol{F}_p. \tag{30.5}$$

On differentiating the deformation gradient ($\boldsymbol{F}$), we get:

$$\dot{\boldsymbol{F}} = \dot{\boldsymbol{F}}_e\boldsymbol{F}_p + \boldsymbol{F}_e\dot{\boldsymbol{F}}_p. \tag{30.6}$$

Now, the velocity gradient ($\boldsymbol{L}$) is defined as

$$\boldsymbol{L} = \dot{\boldsymbol{F}}\boldsymbol{F}^{-1} \tag{30.7}$$

and, thus, $\boldsymbol{L}$ can be decomposed into

$$\boldsymbol{L} = \dot{\boldsymbol{F}}_e\boldsymbol{F}_e^{-1} + \boldsymbol{F}_e\left(\dot{\boldsymbol{F}}_p\boldsymbol{F}_p^{-1}\right)\boldsymbol{F}_e^{-1} = \boldsymbol{L}_e + \boldsymbol{L}_p. \tag{30.8}$$

Plastic deformation in metals is attributed to the cumulative motion of individual dislocations on their respective slip planes, which leads to shear deformation along the slip planes in a crystal. Following this, the plastic velocity gradient ($\boldsymbol{L}_p$) is defined as the sum of shear rates due on all slip systems:

$$\boldsymbol{L}_p = \sum_{i=1}^{N} \dot{\gamma}^{\alpha}\boldsymbol{s}^{\alpha} \otimes \boldsymbol{m}^{\alpha}, \tag{30.9}$$

where $\dot{\gamma}^{\alpha}$ is the shear slip rate on slip system α; unit vectors $\boldsymbol{s}^{\alpha}$ and $\boldsymbol{m}^{\alpha}$ define the slip direction and slip plane; and N is the total number of slip systems. The velocity gradient can further be decomposed in terms of rate of stretching $\boldsymbol{D}$ and rate of spin $\boldsymbol{W}$, for both elastic and plastic component. From Eqs. (30.8) and (30.9), the elastic and plastic part of velocity gradient $\boldsymbol{L}$ is expressed as

$$\boldsymbol{L}_e = \boldsymbol{D}_e + \boldsymbol{W}_e = \dot{\boldsymbol{F}}_e \boldsymbol{F}_e^{-1}, \quad \boldsymbol{L}_p = \boldsymbol{D}_p + \boldsymbol{W}_p = \sum_{i=1}^{N} \dot{\gamma}^{\alpha} \boldsymbol{s}^{\boldsymbol{\alpha}} \otimes \boldsymbol{m}^{\boldsymbol{\alpha}}. \tag{30.10}$$

The relationship between the elastic part of the rate of stretching ($\boldsymbol{D}_e$) and the Jaumann rate of Cauchy stress ($\boldsymbol{\Sigma}$) as a constitutive law is given by Huang [27] as:

$$\overset{\nabla}{\boldsymbol{\Sigma}} + \boldsymbol{\Sigma}(\boldsymbol{I} : \boldsymbol{D}_e) = \boldsymbol{C} : (\boldsymbol{D} - \boldsymbol{D}_p) \tag{30.11}$$

where $\boldsymbol{C}$ is the fourth-order elastic stiffness tensor and $\boldsymbol{I}$ is the second-order unit tensor. The Jaumann rate of stress evolution is

$$\overset{\nabla}{\boldsymbol{\Sigma}} = \dot{\boldsymbol{\Sigma}} - \boldsymbol{W}_e \boldsymbol{\Sigma} + \boldsymbol{\Sigma} \boldsymbol{W}_e. \tag{30.12}$$

The calculation of plastic velocity gradient $\boldsymbol{L}_p$ is based on the shear strain rate $\dot{\gamma}^{\alpha}$ experienced by individual slip system α and is typically expressed as a power law [28] in terms of resolved shear stress τ^{α} as

$$\dot{\gamma}^{\alpha} = \dot{\gamma}_0 \left| \frac{\tau^{\alpha}}{g^{\alpha}} \right|^{n} \operatorname{sgn}(\tau^{\alpha}), \tag{30.13}$$

where $\dot{\gamma}_0$ is the reference shear strain rate, τ^{α} is the strain rate sensitivity and g^{α} is the slip resistance of the slip system α. The resolved shear stress τ^{α} on a slip system α is obtained using Schmid's law

$$\tau^{\alpha} = \operatorname{sym}(\boldsymbol{s}^{\alpha} \otimes \boldsymbol{m}^{\alpha}) : \boldsymbol{\Sigma}. \tag{30.14}$$

and the slip resistance $\dot{g}^{\alpha}$ is given by

$$\dot{g}^{\alpha} = \sum_{i=1}^{N} h_{\alpha\beta} \left| \dot{\gamma}^{\beta} \right|, \tag{30.15}$$

where $h_{\alpha\beta}$ is the hardening modulus. The hardening in a slip system α can be due to dislocation present on the same slip system α (given by $h_{\alpha\alpha}$) or due to the presence of dislocation on other slip system β (given by $h_{\alpha\beta}$). Modified version of the hardening relation proposed by Asaro [29] is used in the case studied [14] and is given by,

$$h_{\alpha\alpha} = (h_0 - h_s) \operatorname{sech}^2 \left(\frac{(h_0 - h_s)\gamma}{\tau_0 - \tau_s} \right) + h_s,$$

$$h_{\alpha\beta} = q h_{\alpha\alpha} (\alpha\beta), \qquad \gamma = \sum_{\alpha} \int_0^t |\dot{\gamma}^{\alpha}| dt, \tag{30.16}$$

where τ is the shear stress, h is the hardening modulus and q is the latent hardening ratio. The subscripts "0" and "s" represent the initial and saturated condition. γ is the accumulative shear strain over all the slip systems.

The CP formulation is implemented in a general purpose finite-element software ABAQUS/Explicit via a material subroutine (VUMAT). Details of the implementation are available elsewhere [14]. Material removal in a single crystal is non-trivial. To this end, a new approach was proposed, which has capability in accounting for shear deformation in individual slip systems. This is given by [14].

$$\max\left(\gamma - \gamma_{cr},\, \gamma_{sl,min} - \gamma_{sl,cr}\right) \geq 0, \tag{30.17}$$

$$\gamma = \min(\gamma^{\alpha}),\ \alpha = 1, 2, \ldots N, \tag{30.18}$$

where subscripts "sl, cr" and "cr" correspond to a critical value for a single slip system and for all slip systems, respectively. The criterion accounts for the total slip (from all slip systems) at a spatial location as well as the induced slip from individual slip systems. Once the mentioned criteria were satisfied, the element was either deleted or converted into smoothed-particle hydrodynamics (SPH) particles.

30.3 Case Studies

In this section, we illustrate two case studies for simulation of machining processes at lower length scales, each for MD and CP FEM. For MD, we refer to the recent work of Goel et al. [19] on the cutting behaviour of silicon (Si). For CP FEM, the work of Liu et al. [14] on the micromachining of single-crystal metal is referred.

30.3.1 MD: Influence of Microstructure on Cutting Behaviour of Silicon [19]

Goel et al. [19] carried out the MD simulation of nanometric cutting of silicon in order to study the difference in the cutting behaviour of single crystalline (SC) and polycrystalline (PC) Si. The MD simulation was performed using a large-scale atomic/molecular massively parallel simulator (LAMMPS) [30]. The simulation set-up at the start of cutting simulation after equilibration is shown in Fig. 30.1. The atoms in the workpiece (Si) and diamond cutting tools (C) were categorized into three zones according to their functions. The Newton atoms are atoms in the region, which were directly involved in the tool–workpiece interaction. The thermostatic atoms are a thin boundary layer of atoms, which were subjected to a thermostat in order to mimic the heat dissipation in the cutting region due to coolant during the machining operation. The boundary atoms were subjected to rigid boundary condition. A sharp cutting edge was used to ensure the brittle-regime machining in

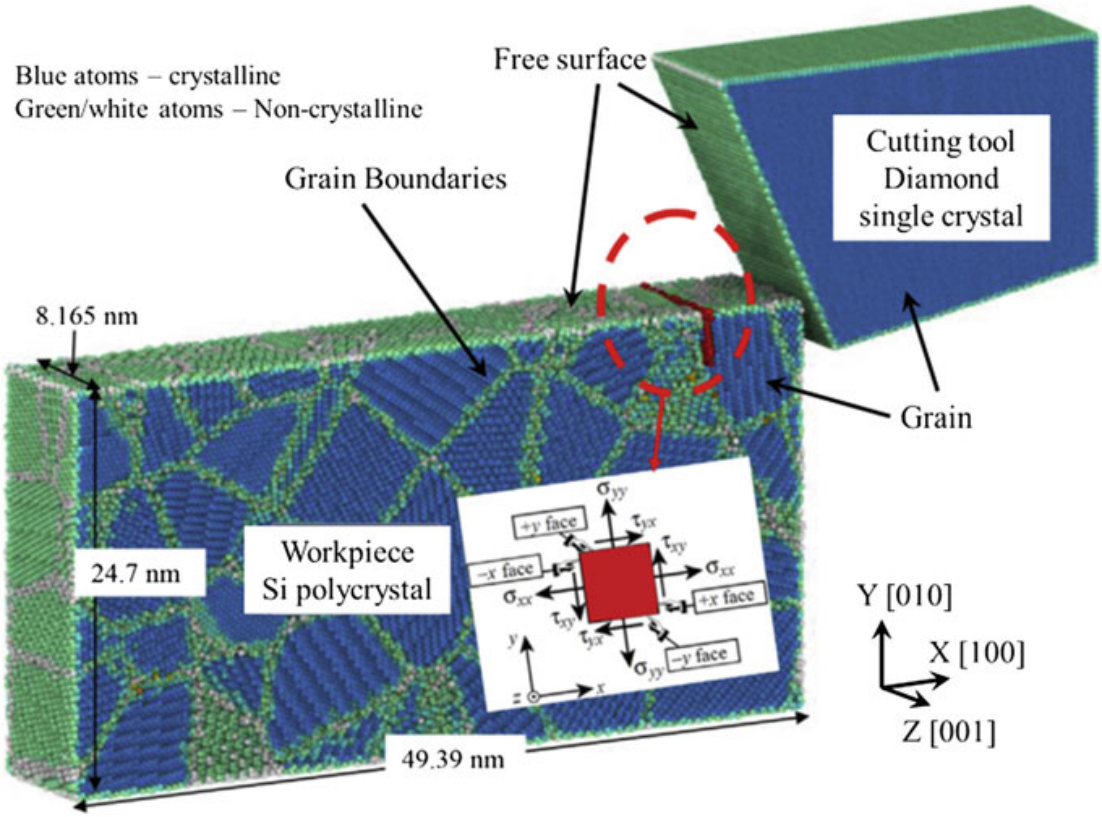


Fig. 30.1 A schematic diagram of the MD simulation model of nanometric cutting of polycrystalline silicon. The blue atoms are diamond atoms while white and green atoms are atoms at grain boundaries and free surface [19]

order to study such a process. To simulate the atomic interactions, ABOP potential was found to be suitable and advantageous as it describes the interaction between carbon and silicon.

The SC silicon required a larger cutting force during machining when compared to PC silicon. The chip for machining SC silicon was more curled (Fig. 30.2a) than the chip obtained from machining PC silicon (Fig. 30.2b). The chip formation in PC silicon takes place with plastic deformation concentrated along the grain boundaries via amorphization of the grain boundaries (Fig. 30.2d). Chunks of grains are pulled off as chips with no defect activity (Fig. 30.2f), indicating minimal plastic activity within the grains. The grain boundaries had large residual stresses (Fig. 30.2b) which help to reduce the critical stress to activate plastic deformation mediated along grain boundary, and, thus, resulting in lower cutting force for PC silicon.

An interesting observation in machining of SC silicon is the periodic occurrence of nano-grooves on the machined surface (Fig. 30.2c). The nano-grooves are regions of amorphous silicon indicating brittle cracking. Such nano-grooves were also observed in experimental studies of nanomachined silicon [25] (inset in Fig. 30.2c). These grooves form along (010) orientation and were inclined 45°–55° to the cutting direction.

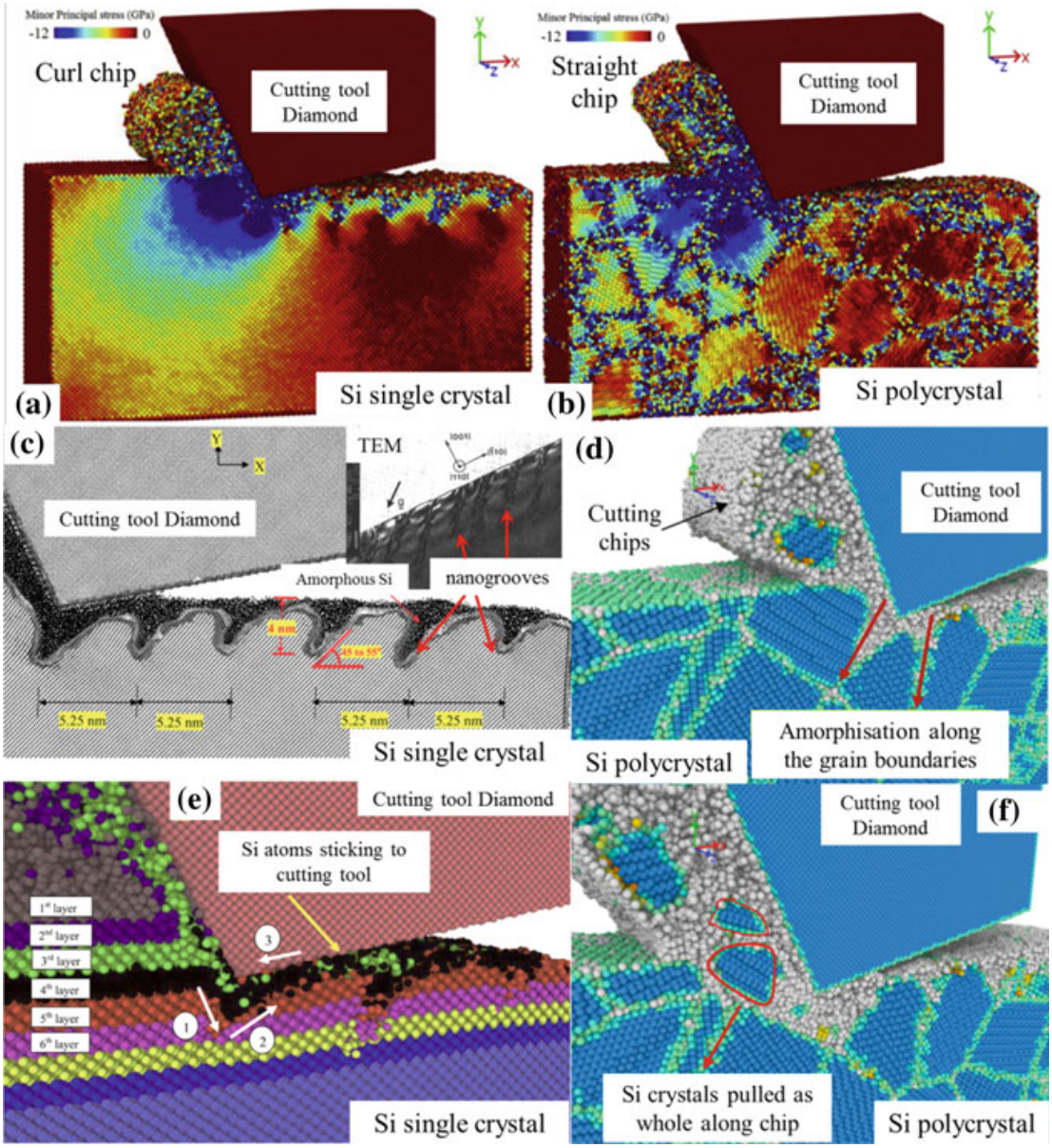


Fig. 30.2 Variation in the minor principal stress (compressive) during cutting of single-crystal (**a**) and polycrystalline (**b**) silicon [19]. (**c**) Formation of nano-grooves in the cutting zone. The inset in (**c**) is TEM image of diamond-turned specimen along (010) orientation [25]. (**e**) The path of vorticity as observed from displacement of atoms from different coloured layers [19]. Images (**d**) and (**f**) are the snapshot at various stages of cutting showing amorphization and grain pulled-outs along the chip [19]

The authors further investigated the grooving mechanism by visualization. For this, the layers of atoms of 0.54 nm thickness before machining are coloured differently. A snapshot of the machined SC silicon shows the grooving process (Fig. 30.2e). On machining, the material layer above the point of stagnation (layer 4 black Fig. 30.2e) was sheared and eventually separated from the workpiece as chips, while the layer beneath was under pure compression and slipped under the bottom the tool face. This process of flow of silicon under the tool creates a vortex.

This vorticity phenomenon leads to the flow of material (black layer) below the cutting tool and downwards in the wake of the cutting edge [19]. Sites formed by such process have abundant amorphous silicon.

One surprising finding in this study as reported by the authors [19] was the more or less absence of dislocation formation upon cutting. A burst of dislocations are normally observed under nanoindentation simulations [31]. Only a few ¼ <111> partial dislocations were observed in a selective case in the sub-surface after machining. Plasticity was mainly via amorphization and flow of silicon and was responsible for the ductile behaviour [19].

30.4 Crystal Plasticity Simulation of Micromachining of Single-Crystal Metal: Methodology and Analysis [14]

A finite-element model for single-crystal copper (FCC structure) was used to study micro-scratching (Fig. 30.3a) [14]. The micro-scratching was performed in two stages using a rigid frictionless cutting tool with 60° wedge angle and 6.25° clearance angle. The two-stage micro-scratching consisted of a linearly increasing cutting depth up to 18 µm (engagement stage) followed by a constant cutting depth of 18 µm (full-engagement stage), as shown in Fig. 30.3b. A cutting velocity of 10 mm/min was used. The model was meshed using eight-node brick elements with reduced integration (C3D8R), and a finer mesh was employed near the cutting zone for better accuracy. The copper single-crystal model is oriented such that (110) plane is parallel to the cutting plane. In the study, the cut was performed along [1–10] direction (0°) (symmetric) and at 45° to [1–10] direction (asymmetric).

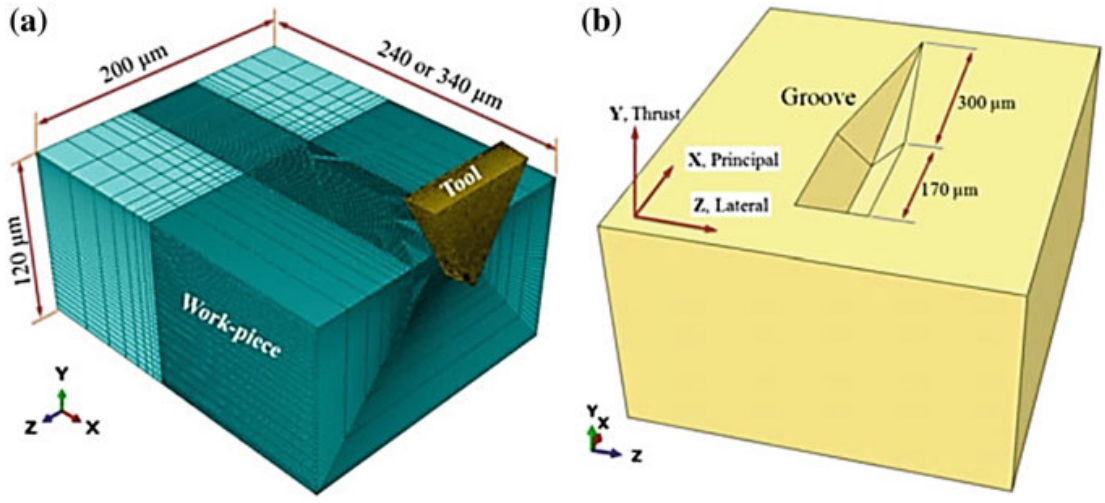


Fig. 30.3 **a** Finite-element model for micro-scratching simulation. **b** Schematic of the two-step micro-scratching experiment [14]

Different material removal models (element deletion, ALE and SPH) adopted resulted in similar cutting and reaction force in the full-engagement stage. However, the time for simulation was significantly different for the three methods. The element deletion was the most computationally efficient. ALE and SPH had higher computational costs, around 7 and 10 times that of element deletion, respectively. Thus, in the study, all further simulations were performed using element deletion.

The cutting forces during the two-step micro-scratching process for 0° and 45° direction are shown in Fig. 30.4a and b, respectively. The principal force matches well with the experimental results. The principal force for the 45° direction was 40% higher than that for 0° direction and shows the anisotropy in the material. In comparison, the predictions for thrust forces are relatively poor. It was observed to be not significantly dependent on the cutting direction.

The chip morphology for 0° direction and 45° direction from simulations (Fig. 30.5a and c) and experiments (Fig. 30.5b and d) is shown in Fig. 30.5. The simulations are capable of capturing the chip morphology and compare well with experiments. For 0° direction, the chip is symmetric, while for 45° direction, it is asymmetric. This is due to the crystal symmetry/asymmetry along the cutting direction. The crystal possesses a twofold symmetry along [1–10] direction (0° direction); hence, activated slip systems were symmetrical with respect to the cutting direction 0°. This results in a symmetrical chip formation (Fig. 30.5a) and stress field. However, no crystal symmetry exists along the cutting 45° direction, which explains the asymmetrical chip (Fig. 30.5c).

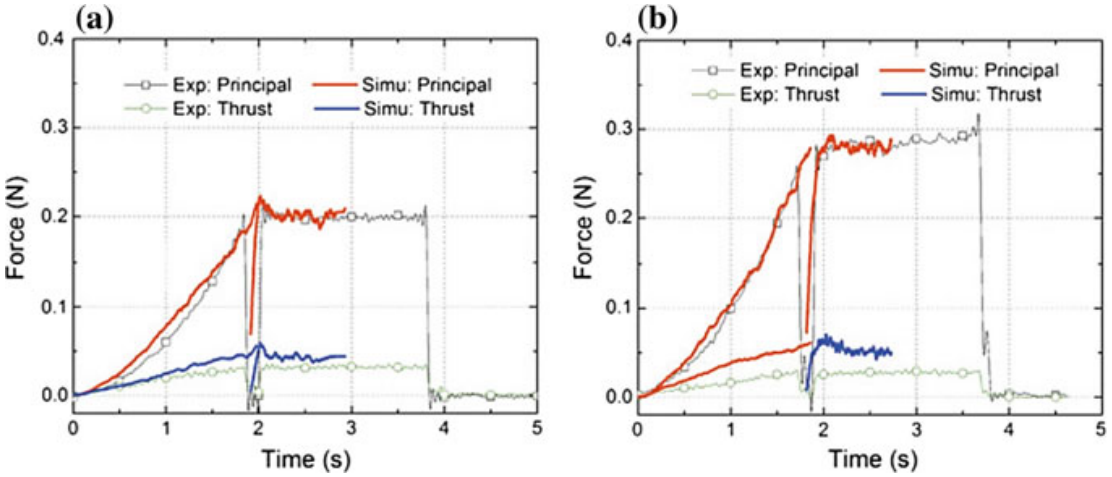


Fig. 30.4 Comparison of cutting force for FE simulations and experimental data in micro-scratching of single-crystal copper along 0° (**a**) and 45° (**b**) direction [14]

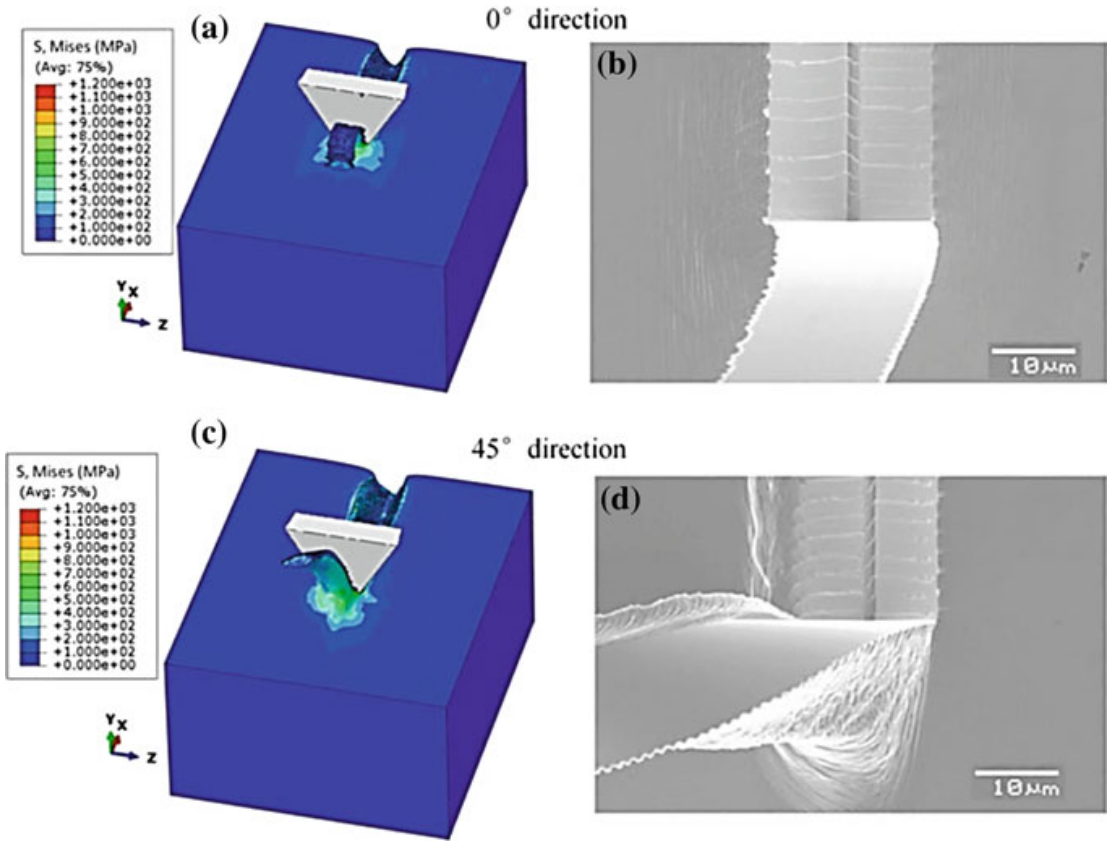


Fig. 30.5 Comparison of chip morphology on micro-scratching of single-crystal copper simulation (**a**) and experiment along 0° (**b**) direction; (**c**) simulation along 45° direction; (**d**) experiment along 45° direction [14]

30.5 Comparison of the Two Methods

The two methods for simulating micromachining processes, namely MD and CP FEM, are versatile techniques with several advantages and disadvantages. A right choice of simulation technique depends on several factors including the workpiece material, processing condition, size of machining cut to name a few.

Molecular dynamics (MD) is very helpful in situations when the underlying deformation mechanism is not well understood. Since potential function is the only necessary information needed to run a simulation. The results obtained provide fundamental insight into the nuances of the deformation processes. Depending on the boundary conditions and interaction among atoms, the system evolves in time. No prior knowledge of how the system will evolve is required. However, we note that the results obtained are dependent on an accurate potential function being defined for the material of interest. This is because a wide variety of interactions, such as ionic, metallic and covalent, are possible among atoms. This particularly becomes important as the number of interacting atom types increases in the system, which leads to an increase in the complexity of the potential and simulation computational cost. Nevertheless, a great depth of information ranging from

vacancy cluster, dislocation type, surface defects to phase transformation can be extracted from the simulation. Timescale and domain size are other factors, which often limit the applicability of MD simulations. With an increase in size, the computational cost increases exponentially. Even with the advances in computational facility available today, the simulation length scale for an MD simulation is limited to a few hundreds of nanometre. MD usually needs a time step of a few fem to seconds for numerical convergence. This again limits the time available for the simulation to a few nanoseconds. Due to this limitation, MD simulations generally involve a very high strain rate, typically 10^8 s^{-1} and above. These strain rates are unrealistic for real-world applications and merely provide an insight into the nuances of deformation processes.

Single-crystal plasticity-based FEM (CP FEM) uses phenomenological constitutive material laws to account for the anisotropic material deformation mechanisms. From a practical standpoint, CP FEM is computationally efficient and possesses the capability of modelling reasonable spatial domains. Various deformation modes including slip systems [14, 32] and twin systems [33] can be incorporated in the CP FEM model. The choice of deformation modes to be considered for CP FEM simulation is principally based on the active dominant mechanisms. Slip-based gradient plasticity is dominant in FCC materials, especially those having high stacking fault energy like Al and Ni. The twin-induced plasticity is dominant primarily in FCC materials with low stacking fault energy (brass and silver) and in HCP materials (Mg, Ti, Zr). By incorporating multiple modes in simulation, the activity of each system during the deformation can be studied via CP FEM simulations. Moreover, the deformation modes changes with sample size and temperature. Such dependency of deformation modes on size [32] and temperature [33] can be incorporated with the help of experimental data, which widens the applicability of CP FEM modelling to a large-scale scenario. Furthermore, the effect of individual deformation mode can be investigated via CP FEM by only considering a particular mode during the simulation, which helps in better understanding the role of individual deformation system. Nevertheless, it should be noted that CP FEM simulation is based on phenomenological constitutive material laws and, thus, for new materials initial experimental investigation for identification of deformation modes and calibration of material model is required.

30.6 Conclusion

In this paper, two methods, namely molecular dynamics (MD) and single-crystal plasticity (CP) FEM, used for simulation of micro-machining process are described in brief with the help of two examples from the literature. Both methods capture the anisotropic behaviour of the material. The basics of the two methods along with their capability are illustrated. Various aspects of the two methods are compared to give the reader a better understanding of respective advantages and deficiencies. MD is more suitable when the underlying deformation mechanisms are not well

understood. On the other hand, CP FEM is apt for realistic simulations and can be executed for a wide range of loading conditions. It is clear that even the type of data obtained from the two simulation methods is very different. Thus, the appropriate choice of simulation method depends primarily on the investigation at hand.

References

1. Chokshi, A.H., Rosen, A., Karch, J., Gleiter, H.: On the validity of the hall-petch relationship in nanocrystalline materials. Scr. Metall. **23**, 1679–1683 (1989). https://doi.org/10.1016/0036-9748(89)90342-6
2. Petch, J.N.: The cleavage strength of polycrystals. J. Iron Steel Inst. **174**, 25–28 (1953)
3. Siegel, R.W., Fougere, G.E.: Mechanical properties of nanophase metals. Nanostruct. Mater. **6**, 205–216 (1995). https://doi.org/10.1016/0965-9773(95)00044-5
4. Suzuki, H., Ikeda, S., Takeuchi, S.: Deformation of thin copper crystals. J. Phys. Soc. Japan. **11**, 382–393 (1956). https://doi.org/10.1143/JPSJ.11.382
5. Bei, H., Shim, S., George, E.P., Miller, M.K., Herbert, E.G., Pharr, G.M.: Compressive strengths of molybdenum alloy micro-pillars prepared using a new technique. Scr. Mater. **57**, 397–400 (2007). https://doi.org/10.1016/J.SCRIPTAMAT.2007.05.010
6. Uchic, M.D., Dimiduk, D.M., Florando, J.N., Nix, W.D.: Sample dimensions influence strength and crystal plasticity. Science **305**, 986–989 (2004). https://doi.org/10.1126/science.1098993
7. Nix, W.D., Gao, H.: Indentation size effects in crystalline materials: a law for strain gradient plasticity. J. Mech. Phys. Solids **46**, 411–425 (1998). https://doi.org/10.1016/S0022-5096(97)00086-0
8. Ma, Q., Clarke, D.R.: Size dependent hardness of silver single crystals. J. Mater. Res. **10**, 853–863 (1995). https://doi.org/10.1557/JMR.1995.0853
9. Fleck, N.A., Muller, G.M., Ashby, M.F., Hutchinson, J.W.: Strain gradient plasticity: theory and experiment. Acta Metall. **42**, 475–487 (1994). https://doi.org/10.1016/0956-7151(94)90502-9
10. Misra, A., M. Pandey, P., Dixit, U.S.: Modeling of material removal in ultrasonic assisted magnetic abrasive finishing process. Int. J. Mech. Sci. 131–132, 853–867 (2017). https://doi.org/10.1016/j.ijmecsci.2017.07.023
11. Dutta, P.P., Kalita, K., Dixit, U.S.: Electromagnetic-force-assisted bending and straightening of AH36 steel strip by laser irradiation. Lasers Manuf. Mater. Process. **5**, 201–221 (2018). https://doi.org/10.1007/s40516-018-0062-6
12. Misra, A., Pandey, P.M., Dixit, U.S.: Modeling and simulation of surface roughness in ultrasonic assisted magnetic abrasive finishing process. Int. J. Mech. Sci. **133**, 344–356 (2017). https://doi.org/10.1016/j.ijmecsci.2017.08.056
13. Hall, R.C.: Single crystal anisotropy and magnetostriction constants of several ferromagnetic materials including alloys of NiFe, SiFe, AlFe, CoNi, and CoFe. J. Appl. Phys. **30**, 816–819 (1959). https://doi.org/10.1063/1.1735247
14. Liu, Q., Roy, A., Tamura, S., Matsumura, T., Silberschmidt, V.V.: Micro-cutting of single-crystal metal: Finite-element analysis of deformation and material removal. Int. J. Mech. Sci. **118**, 135–143 (2016). https://doi.org/10.1016/j.ijmecsci.2016.09.021
15. Demiral, M., Roy, A., Sayed, T. El, Silberschmidt, V.V.: Numerical modelling of micro-machining of f.c.c. single crystal: Influence of strain gradients. Comput. Mater. Sci. **94**, 273–278 (2014). https://doi.org/10.1016/j.commatsci.2014.05.058
16. Liu, Q., Dodla, S., Roy, A., Silberschmidt, V.V.: Crystal-plasticity simulation of micromachining of single-crystal metal: methodology and analysis, 1st edn. Springer, Singapore (2016)

17. Goel, S., Luo, X., Reuben, R.L., Rashid, W.: Bin: atomistic aspects of ductile responses of cubic silicon carbide during nanometric cutting. Nanoscale Res. Lett. **6**, 1–9 (2011). https://doi.org/10.1186/1556-276X-6-589
18. Goel, S., Luo, X.C., Reuben, R.L., Bin Rashid, W., Sun, J.N.: Single point diamond turning of single crystal silicon carbide: molecular dynamic simulation study. Key Eng. Mater. **496**, 150–155 (2011). https://doi.org/10.4028/www.scientific.net/KEM.496.150
19. Goel, S., Kovalchenko, A., Stukowski, A., Cross, G.: Influence of microstructure on the cutting behaviour of silicon. Acta Mater. **105**, 464–478 (2016). https://doi.org/10.1016/j.actamat.2015.11.046
20. Jones, J.E., Ingham, A.E.: On the calculation of certain crystal potential constants, and on the cubic crystal of least potential energy. Proc. R. Soc. A Math. Phys. Eng. Sci. **107**, 636–653 (1925). https://doi.org/10.1098/rspa.1925.0047
21. Jones, J.E.: On the determination of molecular fields. II. From the equation of state of a gas. Proc. R. Soc. A Math. Phys. Eng. Sci. **106**, 463–477 (1924). https://doi.org/10.1098/rspa.1924.0082
22. Erhart, P., Albe, K.: Analytical potential for atomistic simulations of silicon, carbon, and silicon carbide. Phys. Rev. B. **71**, 035211 (2005). https://doi.org/10.1103/PhysRevB.71.035211
23. Daw, M.S., Foiles, S.M., Baskes, M.I.: The embedded-atom method: a review of theory and applications. Mater. Sci. Reports. **9**, 251–310 (1993). https://doi.org/10.1016/0920-2307(93)90001-U
24. Tersoff, J.: Modeling solid-state chemistry: interatomic potentials for multicomponent systems. Phys. Rev. B. **39**, 5566–5568 (1989). https://doi.org/10.1103/PhysRevB.39.5566
25. Becker, C.A., Tavazza, F., Trautt, Z.T., Buarque de Macedo, R.A.: Considerations for choosing and using force fields and interatomic potentials in materials science and engineering. Curr. Opin. Solid State Mater. Sci. **17**, 277–283 (2013). https://doi.org/10.1016/j.cossms.2013.10.001
26. Verlet, L.: Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. Phys. Rev. **159**, 98–103 (1967). https://doi.org/10.1103/physrev.159.98
27. Huang, Y.: A user-material subroutine incorporating single crystal plasticity in the ABAQUS finite element program, Harvard (1991)
28. Hutchinson, J.W.: Bounds and self-consistent estimates for creep of polycrystalline materials. Proc. R. Soc. A Math. Phys. Eng. Sci. **348**, 101–127 (1976). https://doi.org/10.1098/rspa.1976.0027
29. Asaro, R.J.: Crystal Plasticity. J. Appl. Mech. **50**, 921 (1983). https://doi.org/10.1115/1.3167205
30. Plimpton, S.: Fast parallel algorithms for short-range molecular dynamics. J. Comput. Phys. **117**, 1–19 (1995). https://doi.org/10.1006/JCPH.1995.1039
31. Goel, S., Cross, G., Stukowski, A., Gamsjäger, E., Beake, B., Agrawal, A.: Designing nanoindentation simulation studies by appropriate indenter choices: case study on single crystal tungsten. Comput. Mater. Sci. **152**, 196–210 (2018). https://doi.org/10.1016/j.commatsci.2018.04.044
32. Liu, Q., Roy, A., Silberschmidt, V.V.: Size-dependent crystal plasticity: from micro-pillar compression to bending. Mech. Mater. **100**, 31–40 (2016). https://doi.org/10.1016/j.mechmat.2016.06.002
33. Liu, Q., Roy, A., Silberschmidt, V.V.: Temperature-dependent crystal-plasticity model for magnesium: a bottom-up approach. Mech. Mater. **113**, 44–56 (2017). https://doi.org/10.1016/j.mechmat.2017.07.008

Chapter 31
Process Parameters Optimization of Cylindrical Grinding Machining Using Taguchi Method

T. Nancharaiah and M. Subramanyam

Abstract The main objective of this work is to study the effect of material hardness, workpiece speed and depth of cut on surface finish and material removal rate in cylindrical grinding process. For three process parameters at three different levels, L9 orthogonal array (OA) is selected. Experiments were conducted, and roughness values and material removal rates are calculated. S/N ratio is performed to get the optimal process parameters for surface roughness and material removal rate. ANOVA is carried out, and the percentage of contribution of each selected parameter was found.

Keywords Cylindrical grinding · Surface roughness · Material removal rate

31.1 Introduction

Machining involves removing materials using cutting tools for getting rid of unwanted materials from workpiece and converting into the desired shape. There are four movements in cylindrical grinding. They are the work must revolve, the wheel must revolve, the work must pass the wheel and wheel must pass work. Surface finish is the important output parameter in cylindrical grinding. Many researchers worked in this area, a few are cited here for references.

Athreya and Venkatesh [1] illustrated the parameter design of the Taguchi method and optimizing the process parameters. Rupesh [2] applied Taguchi method to decide the optimal process parameter for surface roughness. Prasad [3] measured the vibration and temperature rise. Habrat [4] revealed that grinding with the use of resin bond grinding wheel provides significantly lower grinding force components

T. Nancharaiah (✉)
Department of Mechanical Engineering, Bapatla Engineering College, 522102 Bapatla, India
e-mail: tata_nr@rediffmail.com

M. Subramanyam
Department of Mechanical Engineering, Bapatla Engineering College, Bapatla 522102 AP, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_31

during the process. Mohite et al. [5] focused on the ability of a grinding wheel measured in terms of surface of workpiece during the subsequent grinding operation.

31.2 Experimentation

31.2.1 Grinding Machine Specifications

See (Table 31.1).

31.2.2 En Alloy Steels

Mild steel of different compositions is considered for analysis. EN 19, EN 24 and EN 31 of different grades of the mild steels are turned on a lathe and then cylindrical grinding operation is performed as per L9 orthogonal array. The chemical composition of EN 19, EN 24 and EN3 1 is as shown in Tables 31.2, 31.3 and 31.4, respectively.

Table 31.1 Specifications of grinding machine

Model	DEVCO UC-150
Center distance	160 mm
Center height	102 mm
Angle of swiveling	±9 °C
Speed of transverse	140-260-370-700 mm/min
In feed of hand wheel	0.01 Division
Grinding wheel size	250*25*76.2 mm
Grinding wheel speed	2300 rpm
Work head single spindle speed	90-210-360 rpm
Spindle in taper	MT-3
Total power required	3.5HP (MAX)

Table 31.2 Chemical composition of EN 19 steel

Carbon	0.36–0.44%
Silicon	0.10–0.35%
Manganese	0.45–0.70%
Sulfur	0.040 Max
Phosphorus	0.035 Max
Chromium	1.00–1.40%
Molybdenum	0.20–0.35%

Table 31.3 Chemical composition of EN 24 steel

Carbon	0.36–0.44%
Silicon	0.10–0.40%
Manganese	0.70–1.00%
Sulfur	0.040 Max
Phosphorus	0.035 Max
Chromium	0.90–1.20%
Molybdenum	0.25–0.35%

Table 31.4 Chemical composition of EN 31 steel

Carbon	0.90–1.20%
Silicon	0.10–0.35%
Manganese	0.30–0.75%
Sulfur	0.040%
Phosphorus	0.040%
Chromium	1.00–1.60%

31.2.3 Input Process Parameters and Output Variables

The aim of grinding process is to get a better surface finish and higher rate of production. Currently there is a need for precise and accurate components in every field, so grinding operation is used to achieve this goal.

31.2.3.1 Process Parameters

In this study, the process parameters at their levels are given in Table 31.5.

31.2.3.2 Output Parameters

The material removal rate (MRR) and surface roughness of the workpiece have been measured from machined surfaces. To have minimum MRR and minimum roughness value, the optimal process parameters should be chosen.

Table 31.5 Selection of different levels of input process parameters

Parameter	Level 1	Level 2	Level 3
Hardness	40	47	55
Speed (rpm)	100	214	340
Depth of cut (microns)	10	20	30

31.2.4 Experimental Setup

The experiments were performed on DEVCO UC-150 cylindrical grinding machine. Depth of cut (D), speed (N) and hardness (H) is set at different levels, and the effects of these parameters on material removal rate and surface roughness are studied using Taguchi method. Surface roughness values are measured using SURFTEST SJ301.

For better surface roughness, smaller is better S/N ratio was used in this study and is calculated using the formula given below.

$$\mathbf{S/N\ ratio} = -\mathbf{10log}\ (\mathbf{Ra}^2)$$

In the Taguchi method of the smaller the better method, the smallest S/N value which is calculated is the best optimized running in the system. Smallest S/N ratio has highest surface finishing (Tables 31.6, 31.7, 31.8, 31.9, 31.10 and 31.11).

From the above Ra values, the signal to noise ratio is calculated using the below formula

Table 31.6 Surface roughness values as per L9 orthogonal array

S. No. (microns) (Ra)	Hardness roughness	Speed (rpm)	Depth of cut	Roughness (Ra)
1	40	100	10	0.81
2	40	214	20	0.78
3	40	340	30	1.25
4	47	100	20	1.06
5	47	214	30	1.08
6	47	340	10	1.20
7	55	100	30	1.60
8	55	214	10	1.04
9	55	340	20	1.54

Table 31.7 S/N values for machining the EN 19, 24, 31 workpieces

S. No.	Hardness	Speed (rpm)	Depth of cut (microns)	Ra	S/N ratio
1	40	100	10	0.81	−18.17
2	40	214	20	0.78	−14.83
3	40	340	30	1.25	−17.20
4	47	100	20	1.06	−17.50
5	47	214	30	1.08	−15.98
6	47	340	10	1.20	−21.65
7	55	100	30	1.60	−19.32
8	55	214	10	1.04	−20.34
9	55	340	20	1.54	−20.75

Table 31.8 Experimental data for MRR

S. No.	Depth (microns)	Speed (rpm)	Hardness	Metal removal rate (MRR) (mm^3/min)
1	10	100	40	152
2	10	214	47	163
3	10	340	55	158
4	20	214	55	169
5	20	340	40	176
6	20	100	47	171
7	30	340	47	246
8	30	100	55	224
9	30	214	40	282

Table 31.9 S/N values for machining the EN 19, 24, 31 workpieces

S. No.	Depth (microns)	Speed (rpm)	Hardness	Metal removal rate (MRR) (mm^3/min)	S/N ratio
1	10	100	40	152	−20.81
2	10	214	47	163	−19.89
3	10	340	55	158	−20.52
4	20	214	55	169	−19.42
5	20	340	40	176	−18.53
6	20	100	47	171	−18.86
7	30	340	47	246	−16.18
8	30	100	55	224	−17.27
9	30	214	40	282	−15.36

Table 31.10 ANOVA table for surface roughness

Parameter	Dof	SS	SS_m	% of contribution	F-test
Hardness	2	0.3066	0.1533	47.52	37.85
Speed	2	0.1983	0.0945	30.74	22.59
Depth of cut	2	0.1322	0.0661	20.49	16.32
Error	2	0.0081	0.00405	1.25	–
Total	8	0.6452	–	–	–

$$\mathbf{S/N\ ratio} = -\mathbf{10log}\ (\mathbf{Ra^2})$$

Lower S/N ratio indicates better surface finish.

From the above MRR values, the signal to noise ratio is calculated using the below formula

Table 31.11 ANOVA table for MRR

Parameter	Dof	SS	SS_m	% of contribution	F-test
Depth of cut	2	15042.8	7521.43	89.30	31.82
Speed	2	748.223	391.11	4.44	1.65
Hardness	2	580.223	280.11	3.44	1.22
Error	2	472.69	236.349	2.806	–
Total	8	16844	–	–	–

$$\mathrm{SN} = -10x\log\left[\frac{1}{n}\sum_{i}^{n} y_i^2\right]$$

Higher S/N ratio indicates that maximum MRR value.

31.3 Results and Discussions

31.3.1 Discussions

From Fig. 31.1, it was observed that surface roughness value increases if the depth of cut and speed increases and decreases if hardness increases.

It was observed from Fig. 31.2 that material removal rate increases if depth of cut increases and decreases if hardness and speed increases.

From the ANOVA, it was found that hardness effects on surface roughness are 47.52%, speed are 30.74% and depth of cut are 20.49%, and hardness effects on MRR are 3.44%, speed are 4.44% and depth of cut are 89.30%.

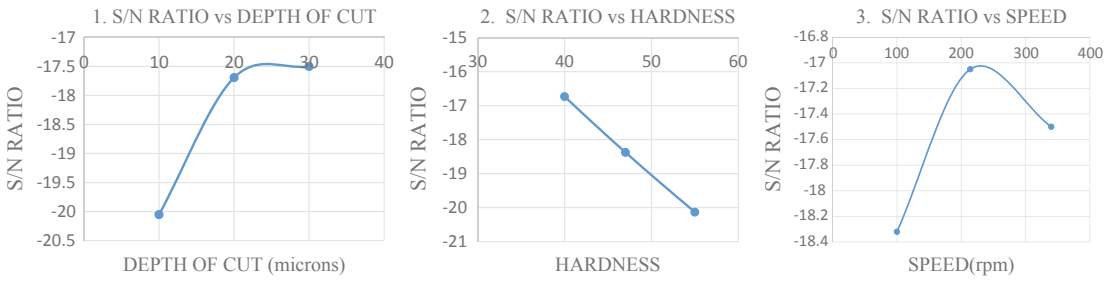


Fig. 31.1 Mean effective plots for surface roughness

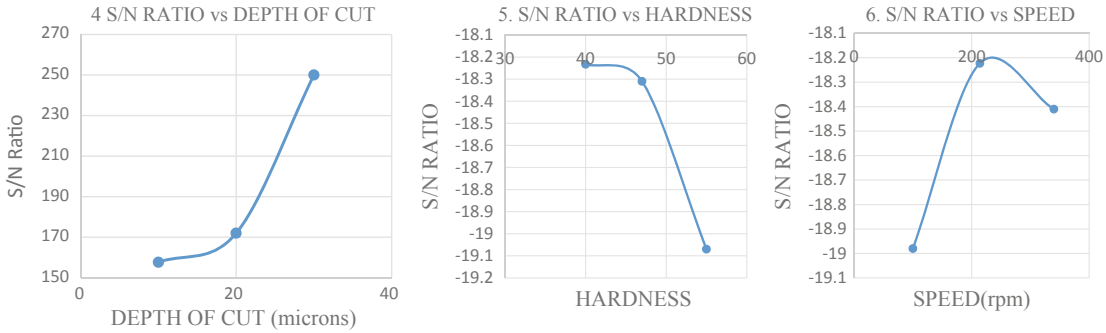


Fig. 31.2 Mean effective plots for material removal rate

31.4 Conclusions

Based on Taguchi method, experiments were conducted using EN alloy steels by varying depth of cut, hardness and speed parameters on cylindrical grinding machine, and the following conclusions were drawn.

(a) Higher depth of cut results MRR increases and surface finish decreases
(b) Increase in hardness, MRR decreases and surface finish increases
(c) Increase in speed, MRR increases and then decreases, surface finish decreases and then increases

From ANOVA, it concludes that hardness is more effective on surface roughness, and depth of cut is more effective on MRR.

References

1. Athreya, S., Venkatesh, Y.D.: Application of Taguchi method for optimization of process parameters in improving the surface roughness of lathe facing operation. Int. Refereed J. Eng. Sci. pp. 13–19 (2012)
2. Karande, R.J.: Optimization of cylindrical grinding machine parameters for minimum surface roughness and maximum MRR. GRD J. Global Res. Develop. J. Eng. **2**(5) (2017, April). ISSN: 2455–5703
3. Prasad, P.P: Design and analysis of grinding wheel of fly ash grinding machine. Int. J. Sci. Eng. Technol. Res. (IJSETR) **5**(8) (2016, August)
4. Habrat, W.F.: Effect of bond type and process parameters on grinding force components in grinding of cemented carbide. In: International Conference on Manufacturing Engineering and Materials, ICMEM 2016, 6–10 June 2016, Nový Smokovec, Slovakia (2016)
5. Mohite, D.D., Patil, K.R., Karande, R.J., Jadhav, V.S. (2017) Modeling and optimization of cylindrical grinding parameters for MRR and surface roughness. Int. J. Eng. Sci. Res. Technol. **6**(4) (2017, April)

Chapter 32
Effect of Sequential Drilling Process on In-situ Bone Temperature

Varatharajan Prasannavenkadesan and Ponnusamy Pandithevan

Abstract In orthopedic surgery, sequential drilling is the process of making multiple holes to facilitate the implant fixation. During the bone drilling process, the generated heat will cause thermal damage that affects the implant fixation strength because of "osteonecrosis," a permanent death of the tissues around the drilling site. In this work, an attempt was made to identify the effects of sequential drilling of cadaveric human femur bone on the heat accumulation. To perform the sequential drilling process, the rotational speed of 1500 rpm and 80 mm/min feed rate was considered with the 4 mm deep holes. Thermal images were captured using Sonel® infrared thermography during the drilling process. It was identified that the accumulated heat increased as the number of sequential holes increased. It was also observed that the distance between the drilled holes influenced the amount of temperature rise. The number of sequential holes drilled and the distance between the holes greatly influenced the temperature rise. This study showed the accumulation of heat during sequential drilling and its consequences in the temperature rise. It is recommended to the surgeons to increase the field of a drilling site in such a way to increase the distance between the subsequent holes to avoid thermal damage to the bone to prevent from further complications.

Keywords Orthopedic surgery · Sequential drilling · Femur bone · In-situ temperature · Osteonecrosis

32.1 Introduction

Bone drilling procedure is one of the most commonly employed surgical procedures in orthopedics, during joint replacement and implant fixations. During the drilling of bone, the interaction between tool and bone; and the contact between the chip

V. Prasannavenkadesan · P. Pandithevan (✉)
Department of Mechanical Engineering, Indian Institute of Information Technology Design and Manufacturing Kancheepuram, Chennai 600127, India
e-mail: ppthevan@iiitdm.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_32

and the drill bore, creates friction which in turn induces heat [1]. When the generated temperature exceeded the threshold will lead to a condition called "thermal necrosis," a permanent death of the bone cells and because of that, the strength of the bone will reduce [2]. The phenomenon of necrosis will lead the bone to loss its healing capability and this will become more common in case of drilling, especially in multi-pass or sequential drilling [3]. This kind of thermal damage increases with increase in temperature and also with respect to the exposure time [4]. Apart from the threshold value that causes the thermal necrosis, the time being exposed also decides the severity of the damage. It has been reported in the literature that the necrosis condition will occur if the temperature is at 47 °C for 1 min and if the temperature increased to 53 °C, then the time to cause the failure will be less than a second [5, 6]. Based on the literature, it was identified that the temperature which is greater than or at 70 °C will result in a sudden state of thermal injury, and also even a lesser temperature with prolonged exposure time will provide the same adverse effects. Other than time, the significant parameters that increase the temperature rise include the thickness of the bone considered and also depth of the holes drilled [7]. Even though the temperature generated during drilling can be reduced with the help of irrigation, it is not recommended in some actual clinical cases [4]. Because of the low thermal conductivity of bone as reported in the literature, the dissipation of heat will be difficult and the build-up heat due to the multi-pass drilling will be more [8]. The main objective of this work was to identify the effect of heat accumulation in the sequential drilling with respect to the distance between the drilling sites, as the severity due to the temperature is large when compared with the single-hole drilling of femoral cortical bone. Based on this work, necessary recommendations can be provided to the surgeons to aid the drilling process in clinical trials.

32.2 Materials and Methods

Sequential drilling was performed in a vertical CNC milling machine (MTAB Maxmill®). Sonel® KT-160A infrared thermal imaging camera was used to capture the temperature data and processed with the Sonel ThermoAnalyze® software. Figure 32.1 shows the experimental setup considered. The surgical drill bit of 3.2 mm diameter was used in this study. Sequential drilling was performed at a rotational speed of 1500 rpm and feed rate of 80 mm/min with 4 mm deep holes to study the accumulation of heat.

The most commonly used parameter range for orthopedic surgery was identified through the literature and from surgeon guidance. The combination of rotational speed and feed rate that provides the higher temperature rise was identified through trial experiments. Based on that, the parameters were selected for this work. To facilitate the process of drilling, the femur bone was cut into pieces and mount into polymethylmethacrylate (PMMA). In order to replicate the clinical case, the dwell time of 1 s was provided between each pass. The accumulation of heat during the sequential drilling is due to the adjacent holes. In general, the thermocouple is

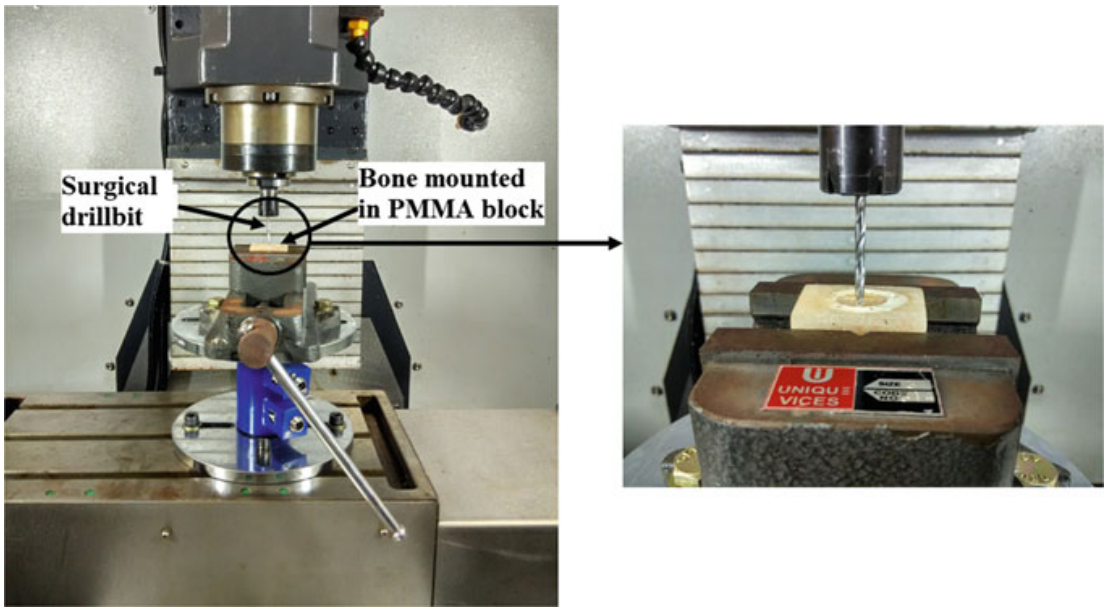


Fig. 32.1 Experimental setup used to conduct the sequential drilling in bone

employed to capture the data during single-hole drilling. But, in the case of sequential hole drilling, the multiple placements of thermocouple require several holes to house the thermocouples, which affects the integrity of the bone [9]. Because of this, the thermocouple is not employed in this study. A 2 × 3-hole array was considered and its schematic representation is shown in Fig. 32.2. The scheme considered for the sequential drilling is based on the distance from margin to margin (4 mm) of the drilled hole. Two points A and B were selected to study the accumulation of heat with respect to the drilling (Fig. 32.2). Six holes were drilled and two reference points "A" and "B" were considered as an analyzing region to identify the heat accumulation phenomenon.

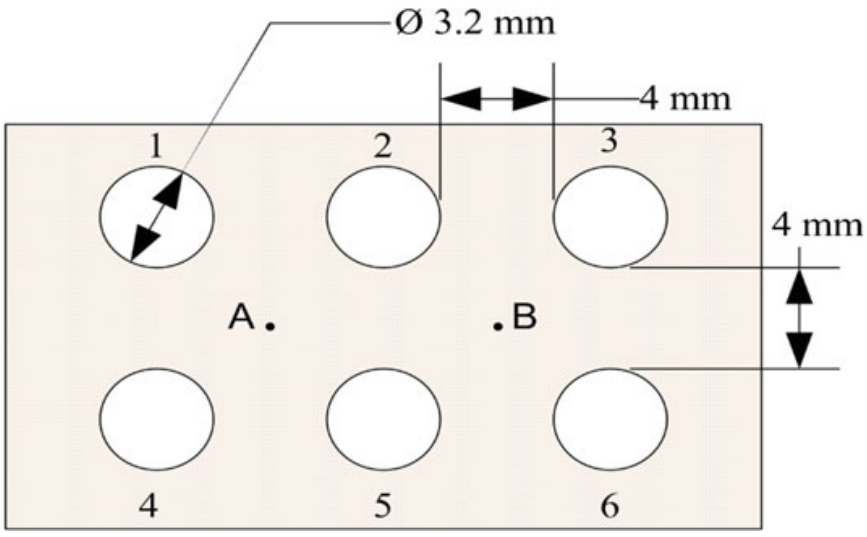


Fig. 32.2 Drilling sequence considered in this work with the reference points A and B

32.3 Results and Discussion

The variation in temperature (T) with respect to the time (t) for each drilling action is shown in Fig. 32.3. A representative thermal image and the corresponding thermal profile (Temperature, T vs. Distance, d) captured are shown in Fig. 32.4 for illustrative purpose. Based on the considered feed rate (80 mm/min), to achieve the full depth (4 mm), 3 s was taken. The dwell period provided was 1 s and thus a total time period of 4 s was taken to complete a single hole. The upright line, shown in Fig. 32.3 represents the hole number and the corresponding abscissa provides the time (t).

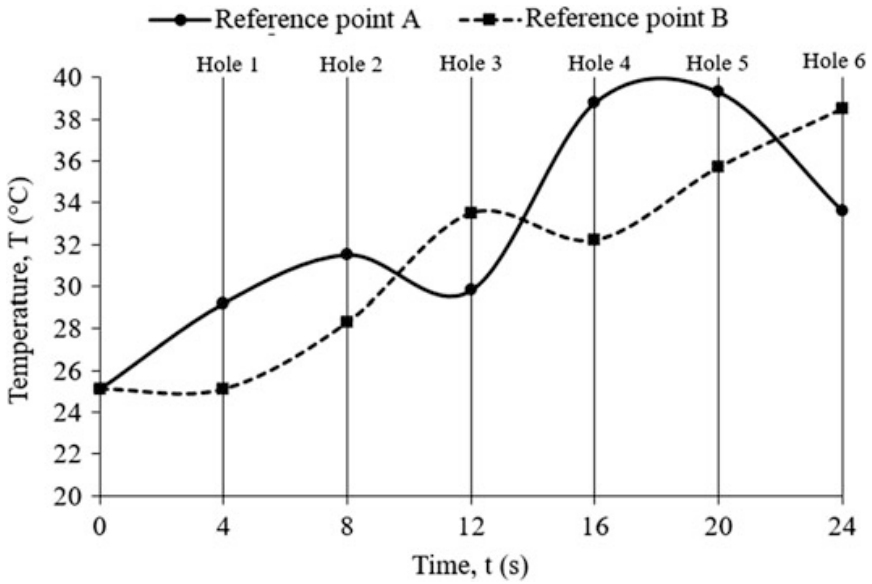


Fig. 32.3 Captured temperature data with respect to the time during sequential drilling

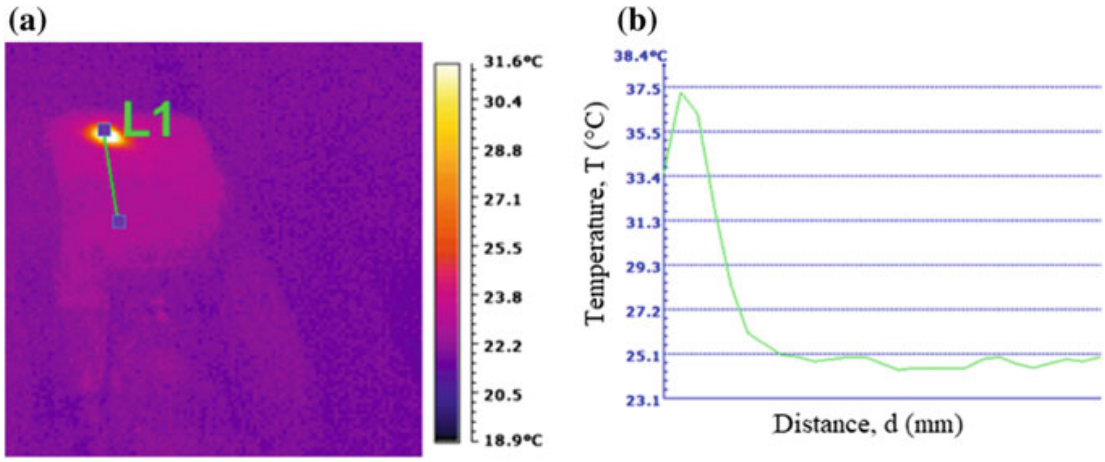


Fig. 32.4 Representative temperature data, **a** thermal image captured during the drilling with its temperature contour and **b** temperature profile of the processed thermal image

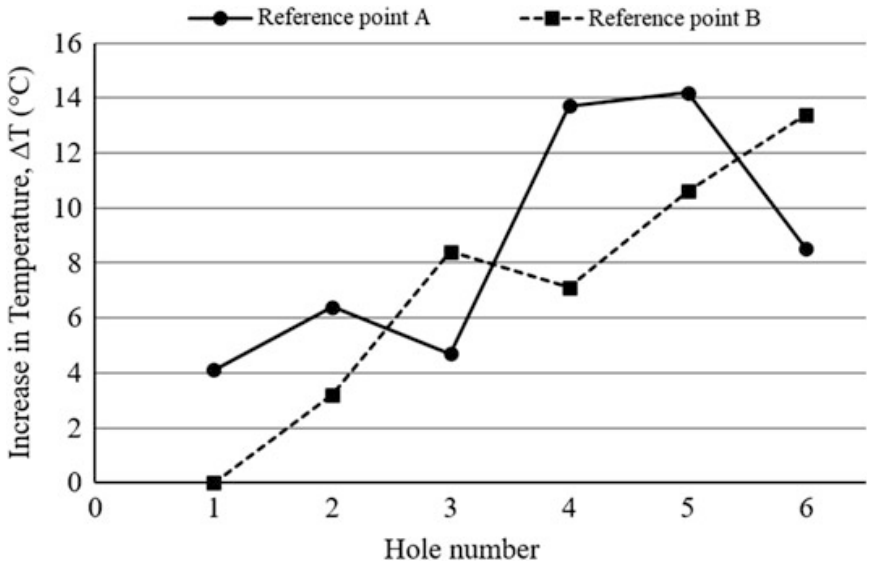


Fig. 32.5 Temperature rise during drilling with respect to the hole number

As the sequential drilling procedure starts, the reference point A (RP_A) responds first and then the reference point B (RP_B). This can be identified through Fig. 32.5, which shows the relation between the increase in temperature rise (ΔT) during the drilling and the hole number.

When drilling was made at the position of the hole number 1, the temperature information was obtained through RP_A and at this particular instance, the absence of heat was observed at RP_B. As the sequential drilling process proceeded through the subsequent passes, the accumulation of heat was clearly visible. The temperature rise was decreased substantially as the position of RP_A was far from the hole number 3. When the drilling site is in proximal (hole number 4) to the RP_A, again the temperature was increased drastically, which may be due to the accumulation of heat. Hole number 6 also showed the same trend as that of hole number 3. In the case of RP_A, it may be reported as the heat accumulation was increased for the drilling sites which are proximal to RP and gets decreased when the drilling site was farther. In the case of RP_B, it is noticed from the obtained data (Fig. 32.5) that there was not an initial rise in temperature when drilling was performed at hole number 1. A similar trend was obtained in the case of RP_B, in such a way that the heat accumulation was greater when the RP_B is at nearer to the drill site. Also, it can be noticed from Fig. 32.5 that the temperature rise at RP_A and RP_B follows the opposite trend. It may be due to the reason that the considered reference points A and B are at opposite to each other. Due to this, at a particular instance, for a considered drilled hole higher temperature rise or heat accumulation was observed at the reference point which is nearer to the drilled hole. Thus, the RP_A responds in a greater way to the sites 1, 2, 4, and 5; the RP_B provides the temperature information for the sites 2, 3, 5, and 6. The farther distance might be the reason for the RP_B for not providing any information about hole number 1 at the initial time. As the sequential drilling is involved in the drilling of more than a hole, the heat gets

accumulated and gets transferred to the next drilling site. Due to this, more chance for the osteonecrosis condition when the temperature reaches 47 °C. So to avoid that in sequential drilling, the surgeons are advised to select the drilling site farther from the previously drilled hole. In this present work, even though, the sequence of the holes to be drilled is changed since the distance between the drilling sites were maintained at the same distance, the temperature profile will not vary. As a preliminary work, those effects were not included in this study. Based on the conducted experiments, it may be concluded that the temperature rise during multi-pass drilling can be reasonably reduced by increasing the dwell time, and this discussion is in good agreement with the numerical study reported earlier [10]. By increasing the dwell period or the time interval between the passes, more time to disintegrate the heat generated can be obtained so that the accumulation can be reduced.

32.4 Conclusion

In this work, an attempt was made to investigate the accumulation of temperature in the cadaveric human femur bone during sequential drilling with a 3.2 mm surgical drill bit for orthopedic applications. Temperature is higher at the margin of the drilled holes, but the accumulation of heat was analyzed at the considered reference points A and B. Accumulation of heat was found to be increased during the subsequent passes of the drilling. From the investigation made, it may be concluded that the distance which is considered as a datum for comparison has a significant impact over the captured temperature information. Because of this kind of heat accumulation due to the sequential drilling, the thermal damage occurred during the orthopedic surgery will be more. By this preliminary study, it was identified that by increasing the margin to margin distance between the drilled holes, the thermal damage can be decreased through the decrease in the heat accumulation or temperature rise. The main limitation of this work is that the observations provided were based on the very limited experimental study. So more follow-up experiments are need to be conducted to have better repeatability and understanding about the effects of sequential drilling. The postoperative success of the drilling process in orthopedic surgery not only depends on the reduction of temperature rise through the identification of optimal distance or optimal parameters but also the geometry of drill bit design, which reduces the efforts in the process of surgery. The authors tend to travel in further directions of this work in such a way to explore the effects of exposure time and the distance between the holes to the thermal damage caused, to implement this phenomenon in the virtual orthopedic surgery protocol. As the temperature generation is more at the proximity of the drill site, and it is also necessary to optimize the way in which the sequential drills can be made to avoid heat accumulation. So the association between the personnel from engineering and physicians is required to develop the surgical protocol for the automated surgeries.

Acknowledgements This research work was funded by Science and Engineering Research Board (File No. YSS/2015/001053), Department of Science and Technology, New Delhi, Government of India.

Ethical Approval This study was approved by the local ethics committee of the Institute and the participants' next of kin provided informed consent before commencing the present study.

References

1. Hillery, M.T., Shuaib, I.: Temperature effects in the drilling of human and bovine bone. J. Mater. Process. Technol. **92**, 302–308 (1999)
2. Bachus, K.N., Rondina, M.T., Hutchinson, D.T.: The effects of drilling force on cortical temperatures and their duration: an in vitro study. Med. Eng. Phys. **22**(10), 685–691 (2000)
3. Baron, R., Horne, W.C.: Bone resorption: regulation of osteoclast activity. In: Bronner, F., Farach-Carson, M.C. (eds.) Topics in Bone Biology. Springer, London (2005)
4. Eriksson, A.R., Albrektsson, T.: Temperature threshold levels for heat-induced bone tissue injury: a vital-microscopic study in the rabbit. J. Prosthet. Dent. **50**(1), 101–107 (1983)
5. Pandey, R.K., Panda, S.S.: Drilling of bone: a comprehensive review. J. Clin. Orthop. Trauma **4**(1), 15–30 (2013)
6. Heydari, H., Cheraghi Kazerooni, N., Zolfaghari, M., Ghoreishi, M., Tahmasbi, V.: Analytical and experimental study of effective parameters on process temperature during cortical bone drilling. Proc. IMechE Part H J. Eng. Med. **232**(9), 871–883 (2018)
7. Wiggins, K.L., Malkin, S.: Drilling of bone. J. Biomech. **9**(9), 555–559 (1976)
8. Lee, J., Chavez, C.L., Park, J.: Parameters affecting mechanical and thermal responses in bone drilling: a review. J. Biomech. **71**, 4–21 (2018)
9. Palmisano, A.C., Tai, B.L., Belmont, B., Irwin, T.A., Shih, A., Holmes, J.R.: Heat accumulation during sequential cortical bone drilling. J. Orthop. Res. **34**(3), 463–470 (2016)
10. Tai, B.L., Palmisano, A.C., Belmont, B., Irwin, T.A., Holmes, J., Shih, A.J.: Numerical evaluation of sequential bone drilling strategies based on thermal damage. Med. Eng. Phys. **37**(9), 855–861 (2015)

Part IV
Additive Manufacturing and Casting

Chapter 33
FE-Based Heat Transfer Analysis of Laser Additive Manufacturing on Ti–6Al–4V Alloy

Sapam Ningthemba Singh, Sohini Chowdhury, Md. S. Mujaheed Khan, Manapuram Muralidhar and Yadaiah Nirsanametla

Abstract A thorough understanding of laser-based additive manufacturing process and effect of various process variables such as scanning velocity and laser beam power on melt-pool dimensions and temperature variation is a promising task in design and manufacture of an able product. The present work is focused on comprehending the thermal and melt-pool behavior of a high layer thickness five-layer laser additive manufacturing of Ti–6Al–4V alloy quantitatively. A three-dimensional (3D) nonlinear transient thermal model is developed based on a finite element procedure to simulate single- and multi-layer of Ti–6Al–4V alloy and to estimate melt-pool dimensions and thermal cycles. In this work, temperature-dependent material properties and Gaussian distributed 'disk' heat source model are implemented along with actual process boundary and initial conditions. Also, the influence of laser beam power and laser scanning velocity was examined with respect to melt-pool characteristics and thermal cycles. The laser scanning velocity ranges from 200 to 500 mm s^{-1} and laser beam power from 100 to 400 W are examined. It is observed that the temperature rises for successive layers as the laser power supply continues on consecutive layers. Also, it is obvious that with the rise in temperature, melt-pool dimensions also increase. Furthermore, the melt-pool dimensions increase as the number of deposited layers increases. Time–temperature history and melt-pool evolution in different layers with respect to laser beam power and laser scanning velocity are presented. To verify the effectiveness of the developed model, simulated results are compared with experimentally measured melt-pool profiles and dimensions. A fair agreement between experimental results and computed values is achieved.

Keywords Laser additive manufacturing · FEM · Melt-pool dimensions · Thermal cycles and Ti–6Al–4V alloy

S. N. Singh · S. Chowdhury · Md.S. M. Khan · M. Muralidhar · Y. Nirsanametla (✉)
Department of Mechanical Engineering, North Eastern Regional Institute of Science and Technology (NERIST), Nirjuli, Arunachal Pradesh 791109, India
e-mail: ny@nerist.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_33

33.1 Introduction

Laser-based additive manufacturing (AM) is an advanced manufacturing technique in which an object is fabricated in a layer-wise manner directly from a CAD file. Moreover, it is quite different from conventional machining processes where parent material undergoes subtractive operations to acquire the desired final product [1]. Hence, it gains some significant advantages over conventional machining processes in terms of design freedom, manufacturing flexibility, the ability to manufacture complex and customized shapes. Also, a variety of materials ranging from polymers, metals, alloys to ceramics can be utilized to manufacture desired products either in powder form or the form of the wire material. The main applications of laser-based additive manufacturing processes are in the field of aerospace, automotive, prototyping and healthcare—dental and medical. Ti–6Al–4V alloy is mainly used in the aerospace, automotive and healthcare industries where high strength, high corrosion resistance and light materials are of main necessity [2–4]. Hence, it is of prime interest to investigate and to have good knowledge of laser additive manufacturing of Ti–6Al–4V which will help to predict the nature of the interaction of the laser with the substrate, melt-pool formation and temperature variations during and after manufacturing operation. Data from such analysis can be used to further study mechanical and microstructural properties. This, in turn, will help in designing an efficient process, optimized machines for specific jobs.

Owing to the importance of additive manufacturing of Ti–6Al–4V alloy in different industrial domains, several researchers have carried out experimental investigations to understand the thermo-physics and its effect on final microstructural and mechanical characteristics of the built product. Yadroitsev et al. [5] reported selective laser melting (SLM) of Ti–6Al–4V alloy for single-layer and single-track application to monitor temperature and microstructural changes during the process. The authors implemented a CCD camera to determine temperature distribution and melt-pool dimensions. It was found that the depth and width of melt pool increase with a continuous supply of laser power. Presence of foreign particles or the addition of extra elements in material results in a drastic change in surface tension and shape of the deposited layer during the LAM process. Traces of rare earth elements such as lanthanum hexaboride and boron in the Ti–6Al–4V and slow cooling rate result in very large Ti particles and improved strength. However, the final structure of the material is more hardened due to the presence of these elements [6]. It is possible to maintain constant energy density by adjusting laser intensity, scanning speed, layer thickness, cooling rate, etc. However, changing the process parameters will result in varying shapes and mechanical properties for constant energy density. Bartolomeu et al. [7] examined the influence of heat source (laser power) and scanning velocity on hardness, density and shear strength of Ti–6Al–4V material during a SLM process. Further studies were carried out on the influence of process variables on microstructure. However, they ignored heat transfer and interaction during the whole process. In experimental research on single-track SLM of Ti–6Al–4V material, it is asserted that the process variables

can be varied to control melt-pool dimensions. Controlling melt-pool dimensions is necessary to adjust the final microstructure of the object. However, single-track experiments cannot provide complete information on a multi-layer built object. Contrary to the theoretical concept, when layer thickness varies from 20 to 60 μm, it is found that there is a negligible effect on the melt-pool depth and width when the scanning speed and laser power are kept constant [8].

Chandrakanth et al. [9] compared the effects on the melt-pool characteristics of Ti–6Al–4V with respect to laser power and scanning velocity during the SLM process considering when the material is in powder state and no powder state [9].

Conducting experiments corresponding to every AM criterions is not desirable in terms of resources and expenditure. Hence, it is preferable to simulate AM process using available tools and methods to examine the effect of input factors on the final built object. Numerical simulations can be used to compute melt-pool dimensions and time–temperature distribution in the laser additive manufacturing process. Qingcheng et al. [10] developed a finite element model to determine the thermo-mechanical behavior of Ti–6Al–4V alloy for laser engineered net shaping (LENS) process using ellipsoidal heat source model. The computed results were determined to be in good agreement with the experimental data [10]. Over the past few years, several thermal models have been developed to simulate AM processes to examine the thermal and mechanical effects with respect to process variables. A computational model of 40 layers of the Ti–6Al–4V alloy was studied by Xufei et al. [11] during the solid forming process using generic heat source model. However, performing finite element analysis for 40 layers will increase computational cost. They determined that the initial temperature of the substrate and the cooling rate after the fabrication process are significant factors. The SLM model developed by Ali et al. [12] considered absorptivity, laser spot area and laser power for effective evaluation of optical penetration of laser beam. They also found that there is no or little noticeable increase in the maximum temperature after the fourth layer. Hence, taking more number of the layer to study the thermal analysis and melt-pool formation is not required. Kurian et al. [13] carried out a detailed comparison of FEM result with experimental data during the SLM process for single-track and single-layer stainless steel 316L material [13]. Moreover, the computational model yielded good results when compared with experimental data, and it is within an acceptable range. Similarly, Jun et al. [14] developed a FEM model to analyze the distortion and residual stresses in electron beam-based additive manufacturing of Ti–6Al–4V alloy. In their work, the authors have implemented a uniform distribution heat source model to simulate the AM process. It was found that the distortion level increases initially for 2–3 layers and then decreases as the layers accumulate [14]. However, the heat source in the real world is not uniform. Thermal distribution along the layer can be applied for determining minimum scanning velocity. A minimum energy density should be maintained to avoid incomplete melting and to prevent balling effect [15]. Jihong et al. [16] analyzed that the heat-affected zone (HAZ) is a significant output parameter in the SLM process [16]. However, several investigators neglected HAZ in their analysis [17–22].

The aim of the present work is to develop a FE model for a five-layer thin-walled Ti–6Al–4V alloy during the laser additive manufacturing process to study the effect of laser power and scan velocity. Using the developed model, five different layers are examined with respect to time–temperature history and melt-pool formation using moving Gaussian distributed disk heat source model. To verify the correctness of the developed model, predicted results are compared with experimental values and found in good agreement.

33.2 Theoretical Background

A transient finite element-based model is developed for five-layer SLM process of Ti–6Al–4V to determine the nonlinear, time–temperature history and melt-pool dimensions at each point of the five layers. SLM setup consists of a laser source, a powder feeding mechanism, a base or a building platform, roller scrapper and other auxiliary parts [16]. The transient heat conduction equation for a three-dimensional heat flux can be represented as

$$\boldsymbol{k}\left(\frac{\partial^2 \boldsymbol{T}}{\partial \boldsymbol{x}^2}+\frac{\partial^2 \boldsymbol{T}}{\partial \boldsymbol{y}^2}+\frac{\partial^2 \boldsymbol{T}}{\partial \boldsymbol{z}^2}\right)+\dot{\boldsymbol{Q}}=\boldsymbol{\rho c_p}\left(\frac{\partial \boldsymbol{T}}{\partial \boldsymbol{t}}+\boldsymbol{v}\frac{\partial \boldsymbol{T}}{\partial \boldsymbol{y}}\right), \tag{33.1}$$

where ρ, c_p, T, v, k and $\dot{Q}$ are the density of the material, specific heat, velocity vector, temperature variable, thermal conductivity and heat generation rate (internal) per unit volume. The first term represents the heat conduction process in Cartesian coordinates system along with internal heat generation rate in the substrate. The initial temperature distribution at time $t = 0$ is given by

$$T(x,y,z)|_{t=0}=T_0, \tag{33.2}$$

where T_0 is the ambient temperature and is assumed to be 298 K.

The thermal transfer within the melt pool is largely due to the latent heat of conduction. Heat is dissipated from the layers as well from the substrate to the atmosphere by convection and radiation modes. The necessary boundary conditions are expressed by Newtonian convection (Newton's Law of Cooling) and gray-surface body behavior [23–29], and mathematically represented as

$$k\frac{\partial T}{\partial n}-q+q_{\mathrm{c}}+q_{\mathrm{r}}=0, \tag{33.3}$$

where q_{c} and q_{r} are heat loss due to convection and radiation mode, and q is the rate of heat influx. The basic equation of convection heat loss and radiation heat loss is

$$q_c = h(T - T_0), \tag{33.4}$$

$$q_r = \sigma\epsilon\left(T^4 - T_0^4\right), \tag{33.5}$$

where $h\sigma$ is the coefficient of convective heat transfer, is the Stephan–Boltzmann constant (5.67 J/m^2 s K^4) and ϵ is melt-pool emissivity, T_0 is initial temperature and T is the instantaneous temperature.

The Gaussian distributed disk heat source model that is utilized in the present analysis is given by [30].

$$q = \frac{3\eta P}{\pi r^2}\exp\left(\frac{-3(x^2+y^2)}{r^2}\right), \tag{33.6}$$

where q is the heat flux, η is the laser absorption coefficient, P is the laser power, r is the effective laser beam radius, and x and y are the respective position of the laser beam at a given time.

A computational model was developed using FE-based ANSYS 14.5 software. The dimensions of the substrate are 10 mm × 31 mm × 4 mm and five layers with layer thickness 1 mm as shown in Fig. 33.1. Also, temperature-dependent material properties such as thermal conductivity, density and specific heat are used [31]. The scanning pattern of the laser on successive layers is shown in Fig. 33.2. The process variables used in the present work are given in Table 33.1. Six different data sets are utilized to simulate laser additive manufacturing process and to analyze melt-pool dimensions and temperature distribution.

The deposition of layers is simulated using the element birth and death technique in ANSYS APDL. When the first layer is exposed to the laser, the elements in the other four layers are classified as death elements. After the first layer, the second layer is activated using element birth feature while keeping the last three layers deactivated, and this process is repeated till the fifth layer.

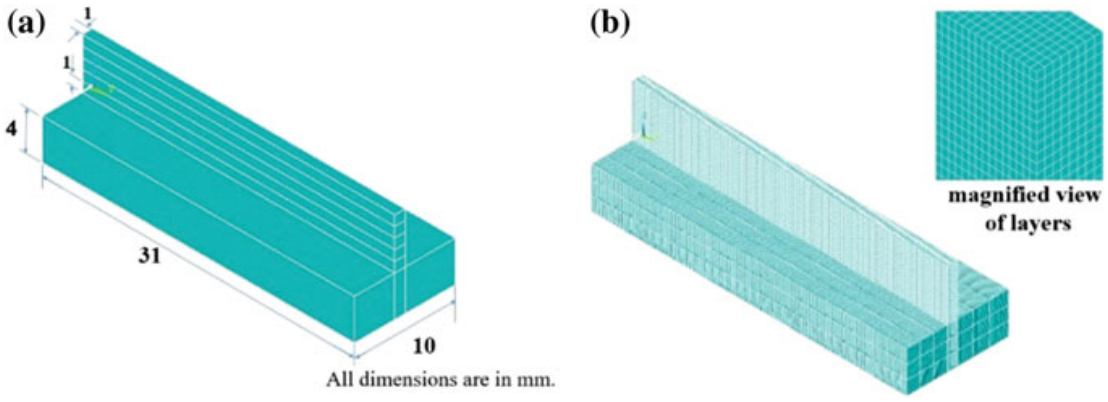


Fig. 33.1 **a** Geometry of the present work, **b** geometry of the present work after meshing

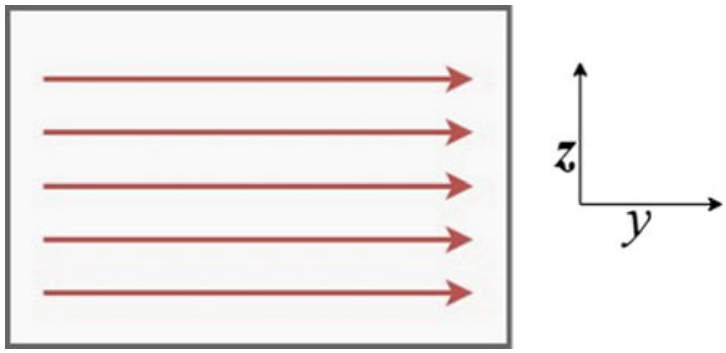


Fig. 33.2 Scanning pattern of the laser on successive layers

Table 33.1 Process parameters used in the present work

Data set no.	1	2	3	4	5	6
Power (W)	100	300	200	200	300	400
Speed (mm s^{-1})	200	200	200	300	400	500

33.3 Results and Discussions

In this section, the influence of laser beam power and scanning velocity on the melt-pool profile and time–temperature distribution at selected points on each layer is presented along with the validation of model predictions with experimental results. Variation of melt-pool dimensions as well as the effect of laser power variation on peak temperature magnitude and overall temperature values is presented. Also, the effect of scanning velocity at constant laser power is also analyzed. To verify the effectiveness of the model, it is calibrated and compared with the experimental data.

For validation of the numerical model results, experimental work is considered from the literature [8]. An SLM experiment was conducted by Yang et al. (2016) which consists of IPG YLR-500 fiber laser, a powder delivery system, a support and building platform along with a process control computer system using Ti–6Al–4V alloy powder with a range of 30–50 μm in an inert environment of Argon. Table 2 depicts the process parameters used in the experimental investigation. From the experimental investigation, it is observed that the melt-pool width ($\sim$70 μm to $\sim$160 μm) and depth ($\sim$10 μm to $\sim$200 μm) increase as the power is increased from 100 to 500 W keeping the other parameters unchanged. But the melt pool decreases when the scanning speed increases from 100 to 1000 mm s^{-1} keeping the laser power and layer thickness constant.

A finite element-based heat transfer model is developed to examine the melt-pool formation and thermal cycles. To confirm the viability of the developed model, the single-layer FEM model is developed initially of 1 mm layer thickness. The single-layer model is compared and validated with the experimental data from literature [8]. The melt-pool depth corresponding to the computed value and the experimental result of data set #3 of Table 33.1 are determined to 0.12 and

0.13 mm. Moreover, a percentage difference in melt-pool depth is estimated to be 5.38% for all processing conditions of Table 33.1 and which is under acceptable range. Comparison of computed (left) and experimental (right) melt-pool shape corresponding to data set #3 of Table 33.1 is presented in Fig. 33.3, and the red color zone represents the melt zone.

With the application of laser power, a sudden rise in temperature is observed with a high-temperature gradient. Figure 33.4 shows the 3D transient temperature distribution at different positions and at different layers corresponding to process variables given in Table 33.1. Figure 33.4a shows the location of the laser source

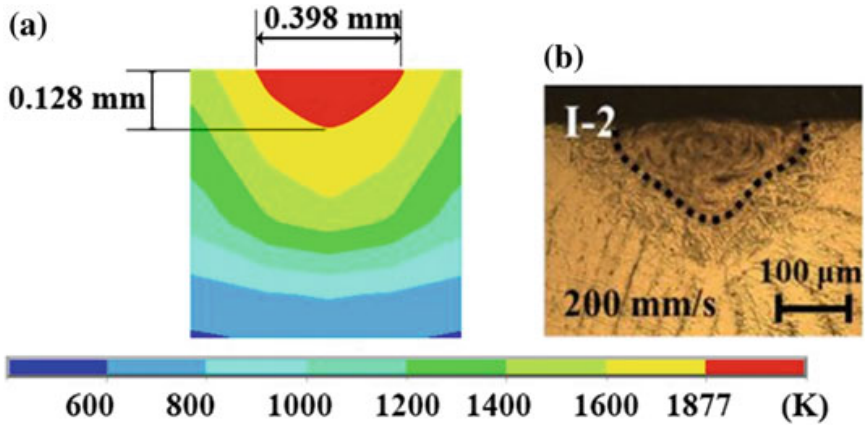


Fig. 33.3 Comparison of computed (left) and experimental (right) melt-pool shape and size corresponding to data set #3 of Table 33.1 of for a single-layer laser additive manufacturing

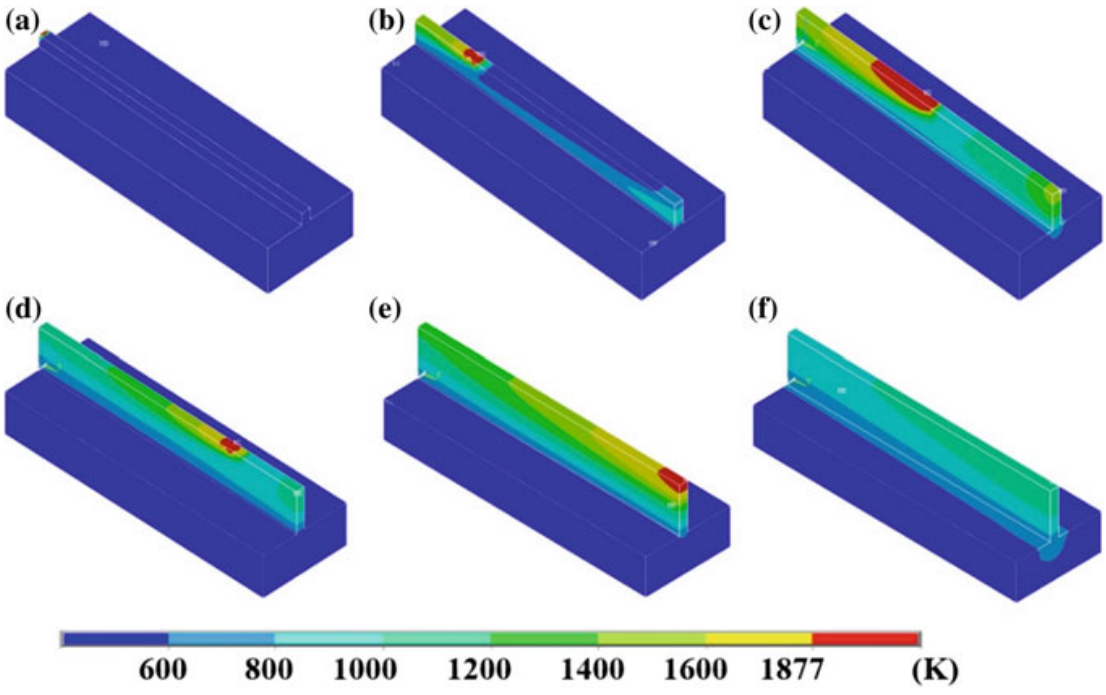


Fig. 33.4 3D transient temperature distribution at different positions and different layers during five-layer laser additive manufacturing of Ti–6Al–4V alloy

which is at the beginning of the first layer corresponding to data set #6. Figure 33.4 (b) represents the location of the laser source which is at a distance of 7.75 mm along the scanning direction of the second layer corresponding to data set #5. Similarly, Fig. 33.4c demonstrates the location of the laser source which is at the center of the third layer corresponding to the data set #4. Figure 33.4d represents the location of the laser source which is at a distance of 23.25 mm along the scanning direction of the fourth layer corresponding to data set #3. Figure 33.4e shows the location of the laser source at the end of the fifth layer corresponding to data set #2. Figure 33.4f represents the cooling nature of laser additive manufacturing of Ti–6Al–4V alloy. It is observed from Fig. 33.4c, d that the melt zone and heat-affected zone are wider for high laser beam power and low scanning velocity.

Figure 33.5 depicts the time–temperature history of each layer at a distance of 7.75 mm along the scanning direction. This figure shows a rapid rise in temperature followed by a gradual decrease in temperature. The sudden rise in temperature is due to sudden impingement of laser power at a point, and a gradual decrease is because of laser source moving away from that specific point. Theoretically, there should be five different peaks for the first layer, four for the second layer and three for the third layer and likewise. This is because laser power passes above a specific point for five, four, three times during the whole process. However, the graph reveals only three visible peaks. It shows that for a point below three layers where the laser is striking, there is little effect of laser power on that point for specific data set. However, other data set represents five peaks during the whole process. This reflects that high laser power affects the temperature at a point, even though layers are away from it. However, the overall temperature of the point still increases as the number of layers increases. After the manufacturing process, there is a rapid decrease in the temperature due to cooling up to 2.5 s, after that the cooling takes place slowly. This is due to the temperature difference between the ambient and builds parts decreases, and heat transfer due to convection decreases sharply.

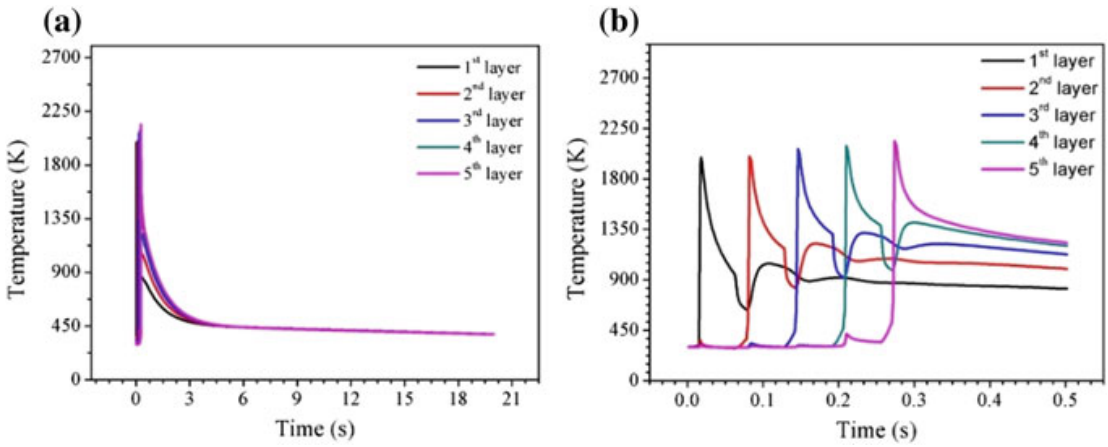


Fig. 33.5 Time–temperature history of each layer during a five-layer laser additive manufacturing of Ti–6Al–4V alloy for data set # 6 of Table 33.1

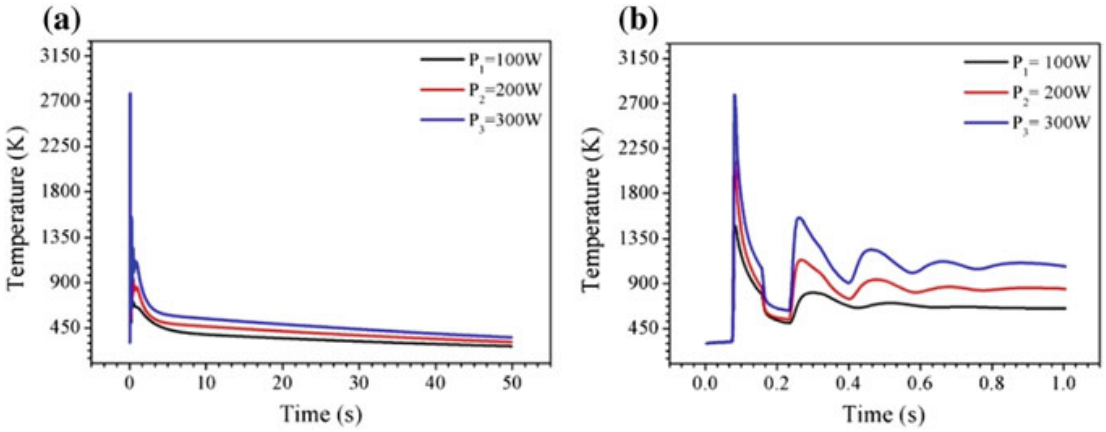


Fig. 33.6 Thermal cycles at constant scanning velocity of 200 mm s^{-1} and different laser beam power: **a** full view and **b** magnified view

The effect of laser power on the time–temperature history of Ti-alloy material is graphically represented in Fig. 33.6. For constant laser scanning velocity of 200 mm s^{-1}, the laser power is varied from 100 to 300 W. Correspondingly, the peak temperatures are 1484, 2113 and 2788 K at the top of the first layer and a distance of 23.25 mm along the scanning direction corresponding to the laser power 100, 200 and 300 W, respectively. There are five peaks for 300 and 200 W laser power; however, for 100 W there are only three peaks. Figure 33.7 shows the increased melt-pool dimensions when laser power varies from 100 to 200 W at constant scanning velocity. It can be identified that there is an incomplete melting and sintered which is represented by the yellow and light green region. This results in a change of microstructure and mechanical properties in the final product.

Figure 33.8 depicts the temperature variation at constant laser power and increasing laser scanning velocity. At constant laser power of 200 W (Fig. 33.8a) and an increasing scanning velocity from 200 to 300 mm s^{-1}, the temperature reduces from 2113 to 1738 K at the top of the first layer and a distance of 23.25 mm along the scanning path. A similar trend is observed when power is

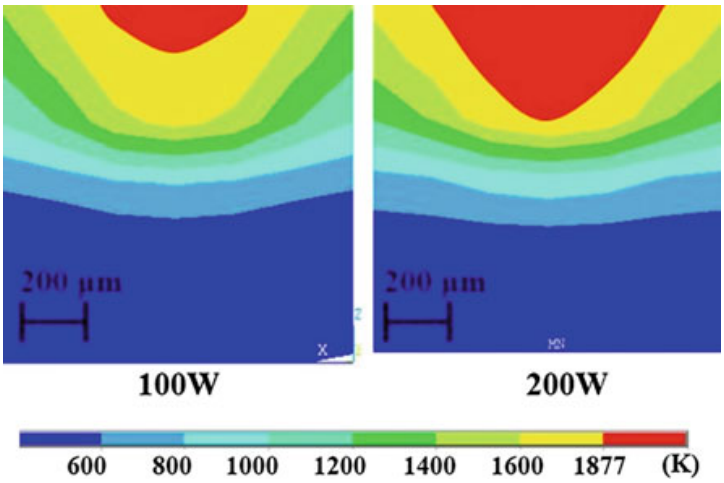


Fig. 33.7 Effect of laser power on the melt-pool formation at a constant scanning velocity of 200 mm s^{-1} in Ti–6Al–4V alloy during laser additive manufacturing

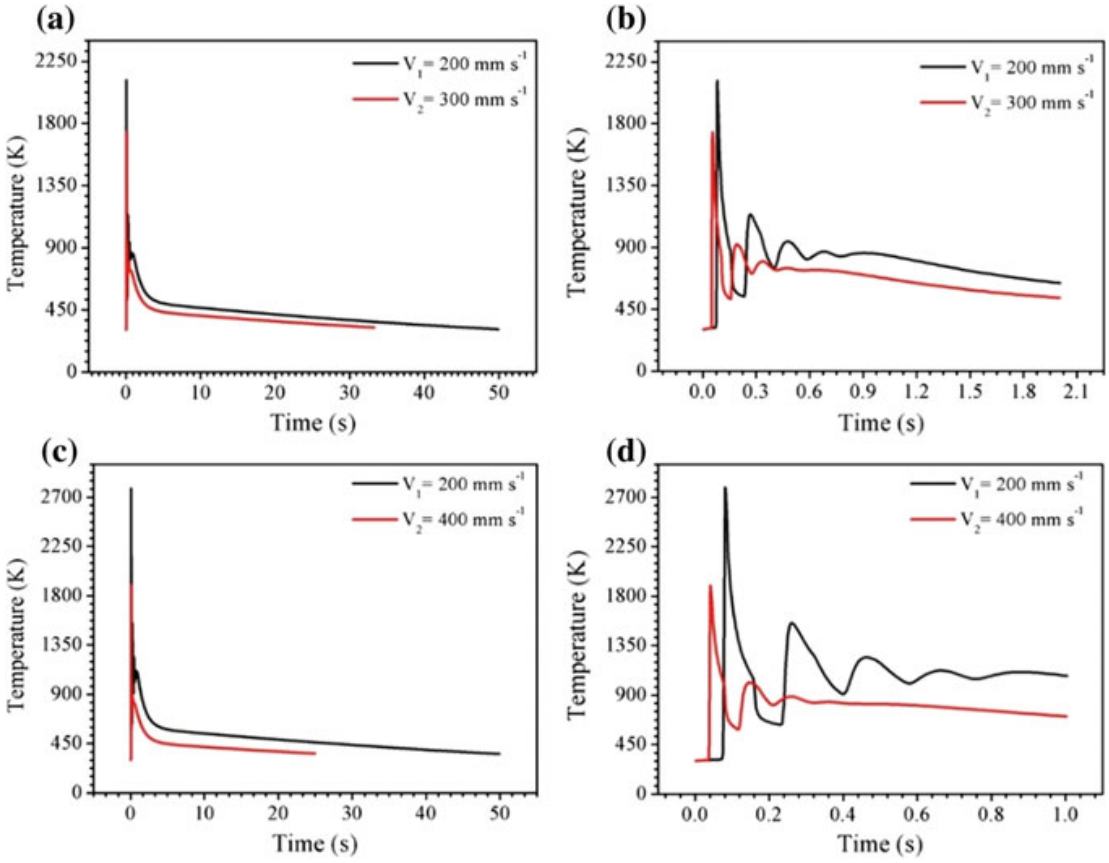


Fig. 33.8 Effect of scanning speed at constant laser beam power **a**, **b** 200 W and **c**, **d** 300 W

300 W (Fig. 33.8c) and scanning velocity is increased from 200 to 400 mm s^{-1}, and temperature decreases from 2788 to 1895 K. The temperature distribution and peak temperatures were estimated using developed model along with melt-pool formation.

33.4 Conclusion

In the present work, a 3D transient FE model of a large layer thickness five-layer laser-based additive manufacturing is developed to understand the complex evolution of melt-pool shape and dimensions as well as transient temperature distribution. The developed model results are validated with the experimentally measured values to verify the effectiveness of the process model. The model is determined to be in good accordance with the experimental data.

The influence of laser beam power and scanning velocity examined five-layer laser additive manufacturing process using a developed model. It is observed that the temperature of subsequent layers rises with an increase in layer number. In the first layer, the substrate played an important part in conducting the heat away from the layers. The results also showed that for low power or high speed, there is the

negligible effect of temperature on the first layer when the laser source passes through the fourth layer. However, if the laser power is high and the scanning speed is low, it is still a significant influencing factor for the first layer even when the laser source passes on the fifth layer. It is also found that the temperature increases up to 2788 from 1484 K when laser beam power increases from 100 to 300 W at the top of the first layer for the same scanning velocity, 200 mm s^{-1}. Moreover, the temperature decreases from 2788 to 1895 K when the laser power is 300 W, and scanning velocity is changed from 200 to 400 mm s^{-1}. The future scope of the present work is the development of a mechanical model to estimate the residual stresses and mechanical properties using the present thermal model. Also, this model will provide a base to develop a process model for a multi-layer, multi-tracks during additive manufacturing.

Acknowledgements The authors thank the TEQIP II grant for financially supporting this work at North Eastern Regional Institute of Science & Technology, Itanagar, Arunachal Pradesh, India.

References

1. International Organization for Standardization: Standard terminology for additive manufacturing—general principles—part 1: terminology, 2015, (52900:2015 (E))
2. Bandyopadhyay, A., Espana, F., Balla, V.K., Bose, S., Ohgami, Y., Davies, N.M.: Influence of porosity on mechanical properties and in vivo response of Ti–6Al–4V implants. Acta Biomater. **6**, 1640–1648 (2010)
3. Vrancken, B., Thijs, L., Kruth, J.P., Humbeeck, J.V.: Heat treatment of Ti–6Al–4V produced by selective laser melting: Microstructure and mechanical properties. J. Alloy. Compd. **541**, 177–185 (2012)
4. Sun, J., Yang, Y., Wang, D.: Parametric optimization of selective laser melting for forming Ti–6Al–4V samples by Taguchi method. Opt. Laser Technol. **49**, 118–124 (2013)
5. Yadroitsev, I., Krakhmalev, P., Yadroitsava, I.: Selective laser melting of Ti–6Al–4V alloy for biomedical applications: temperature monitoring and microstructural evolution. J. Alloy. Compd. **583**, 404–409 (2014)
6. Bermingham, M.J., McDonald, S.D., Dargusch, M.S.: Effect of trace lanthanum hexaboride and boron additions on microstructure, tensile properties and anisotropy of Ti–6Al–4V produced by additive manufacturing. Mater. Sci. Eng., A **719**, 1–11 (2018)
7. Bartolomeu, F., Faria, S., Carvalho, O., Pinto, E., Alves, N., Silva, F.S., Miranda, G.: Predictive models for physical and mechanical properties of Ti6Al4 V produced by selective laser melting. Mater. Sci. Eng., A **663**, 181–192 (2016)
8. Jingjing, Y., Jie, H., Hanchen, Y., Jie, Y., Ming, G., Zemin, W., Xiaoyan, Z.: Role of molten pool mode on formability, microstructure and mechanical properties of selective laser melted Ti–6Al–4V alloy. Mater. Des. **110**, 558–570 (2016)
9. Chandrakanth K.: The effect of laser power and scan speed on melt pool characteristics of pure titanium and Ti–6Al–4V alloy for selective laser melting. Wright State University, MSME Thesis, 1510 (2016)
10. Qingcheng, Y., Zhang, P., Cheng, L., Zheng, M., Chyu, M., Albert, C.T.: Finite element modelling and validation of thermo-mechanical behaviour of Ti–6Al–4V in directed energy deposition additive manufacturing. Addit. Manuf. **12**, 169–177 (2016)

11. Xufei, L., Xin, L., Chiumenti, M., Miguel, C., Li, J., Liang, M., Wei, L., Yunlong, H., Weidong, H.: Finite element analysis and experimental validation of the thermomechanical behaviour in laser solid forming of Ti–6Al–4V. Addit. Manuf. **21**, 30–40 (2018)
12. Ali, F., Mohsen, B., Ehsan, F., Golabi, S.: Finite element simulation of selective laser melting process considering optical penetration depth of laser in powder bed. Mater. Des. **89**, 255–263 (2016)
13. Kurian, A., Arivazhagan, N., Senthilkumaran, K.: Numerical and experimental investigations on laser melting of stainless steel 316L metal powders. J. Manuf. Process. **16**(3), 345–355 (2014)
14. Jun, C., Michael, A.G., Philip, N.: Finite-element analysis and experimental validation of thermal residual stress and distortion in electron beam additive manufactured Ti–6Al–4V build plates. J. Mater. Process. Technol. **237**, 409–419 (2016)
15. Patcharapit, P., Recep, O., Shi-Chune, Y.: Numerical and experimental investigations of micro and macro characteristics of direct metal laser sintered Ti–6Al–4V products. J. Mater. Process. Technol. **240**, 262–273 (2017)
16. Jihong, Y., Shoujin, S., Milan, B., Wenyi, Y.: Experimental investigation and 3D finite element prediction of the heat affected zone during laser assisted machining of Ti–6Al–4V alloy. J. Mater. Process. Technol. **210**(15), 2215–2222 (2010)
17. Peyre, P., Aubry, P., Fabbro, R., Neveu, R., Longuet, A.: Analytical and numerical modelling of the direct metal deposition laser process. J. Phys. D Appl. Phys. **41**(2), 025403 (2008)
18. Gockel, J., Beuth, J., Taminger, K.: Integrated control of solidification microstructure and melt pool dimensions in electron beam wire feed additive manufacturing of Ti–6Al–4V. Addit. Manuf. **1–4**, 119–126 (2014)
19. Fachinotti, V.D., Cardona, A., Baufeld, B., Van der Biest, O.: Finite-element modelling of heat transfer in shaped metal deposition and experimental validation. Acta Mater. **60**(19), 6621–6630 (2012)
20. Romano, J., Ladani, L., Razmi, J., Sadowski, M.: Temperature distribution and melt geometry in laser and electron-beam melting processes: A comparison among common materials. Addit. Manuf. **8**, 1–11 (2015)
21. Labudovic, M., Hu, D., Kovacevic, R.: A three dimensional model for direct laser metal powder deposition and rapid prototyping. J. Mater. Sci. **38**(1), 35–49 (2003)
22. Wang, L., Felicelli, S., Gooroochurn, Y., Wang, P.T., Horstemeyer, M.F.: Optimization of the LENS® process for steady molten pool size. Mater. Sci. Eng., A **474**(1), 148–156 (2008)
23. Beauchamp, M.J., Nordin, G.P., Woolley, A.T.: Moving from mill fluidic to truly microfluidic sub-100-μm cross-section 3D printed devices. Anal. Bioanal. Chem. **409**, 4311–4319 (2017)
24. Monaghan, T., Harding, M.J., Harris, R.A., Friel, R.J., Christie, S.D.: Customisable 3D printed microfluidics for integrated analysis and optimisation. Lab Chip **16**, 3362–3373 (2016)
25. Sweet, E.C., Rudra, R., Mehta, R.L., Lin, L.: Finger-powered 3D printed microfluidic pumps. In: 19th International Solid-State Sensors, Actuators and Microsystems Conference Transducers, pp. 1766–1769, IEEE, Taiwan (2017)
26. Bose, S., Robertson, S.F., Bandyopadhyay, A.: Surface modification of biomaterials and biomedical devices using additive manufacturing. Acta Biomater. **66**, 6–22 (2018)
27. Bandyopadhyay, A., Bose, S.: Additive Manufacturing. CRC Press, US (2015)
28. Bose, S., Ke, D., Sahasrabudhe, H., Bandyopadhyay, A.: Additive manufacturing of biomaterials. Prog. Mater Sci. **93**, 45–111 (2018)
29. Puleo, D.A., Nanci, A.: Understanding and controlling the bone–implant interface. Biomaterials **20**(23), 2311–2321 (1999)
30. Yadaiah, N., Bag, S.: Development of egg-configuration heat source model in numerical simulation of autogenous fusion welding process. Int. J. Therm. Sci. **86**, 125–138 (2014)
31. Boyer, R., Collings, E.W., Welsch, G.: Materials Properties Handbook: Titanium Alloys. ASM International, US (1994)

Chapter 34
Material Perspective and Deformation Pattern of Micro-Sized Metallic Particle Using Cold Gas Dynamic Spray

Abdul Faheem, Faisal Hasan and Qasim Murtaza

Abstract The cold gas dynamic spray is a solid powder deposition phenomenon of coating and repairing with a broad materials range. This deposition technology gained attention not only in institutions but also in industrial aspects. This coating technology is a surface engineering technology which is presently replacing electron, laser beam type of technologies for additive manufacturing. This study provides the pattern of deformation of both particle and substrate under the impact of high particle velocity. An ABAQUS/Explicit, a finite element methodology was used to forecast the coating behaviour of different engineering materials. A FEM approach was used for the simulation of the metallic impact of substrate and particle. A wide range of impact velocities was used to finding the threshold or critical velocity for the proper deposition, phenomenon of rebounding and also to divulge the mechanism of governance, i.e. adiabatic shear instability (ASI). This study focused and revealed how engineering materials that are commonly used in aerospace engineering behaved under different impact conditions using kinetic spray coating technology.

Keywords Cold gas dynamic spray · Numerical simulation · Critical velocity · Finite element methodology

34.1 Introduction

Cold gas dynamic spray or cold spray (CS) or kinetic spray is a new material deposition technology which is procuring attention of surface engineering. This is a newly gaining noticed coating process used for a wide range of materials, i.e. ductile, brittle, ceramics and composite materials, etc., [1] without noteworthy heating of micro- or

A. Faheem · F. Hasan
Department of Mechanical Engineering, AMU, Aligarh 202002, India

Q. Murtaza (✉)
Department of Mechanical Engineering, DTU, New Delhi 110042, India
e-mail: qasimmurtuza@dce.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_34

nano-solid particles. A kinetic spray, firstly, at the Russian's Institute of Applied and Theoretical Mechanics of Science was discovered more than two decades ago and the first patent on it done in 1994 [2]. In CS, adhesion between the particle and substrate occurs due to the impact of the high velocity of micro- or nano-size of particles [3]. This high velocity not only effects the deformation of both of it but also equated a large amount of plastic deformation [4]. The acceleration of particles is obtained by the convergent–divergent nozzle under high pressure with a Mach number in the range of 3 for the impact phenomenon [5]. Apart from heating, the dimension instability and surface protection also be done by alternate technology, i.e. kinetic spray [6]. This metal joining process is similar to cladding in which the fusion of parts occurs to fasten the dissimilar metals [7, 8] (Fig. 34.1).

The major disadvantage of high-pressure cold spray system is the erosion at the throat of the nozzle due to erosion particle which not only affects the working of the nozzle but also the deposition quality [9]. (2) Low-pressure cold spray is a system in which powder is fed at that point in the nozzle where the gas expanded the least pressure. This system is commonly used for metal alloys and ceramics. LPCS systems are generally portable and smaller and also workable on a limited range of velocity (300–600 m/s). Due to the portability of LPCS, it does not need pressurizer feeder (Fig. 34.2).

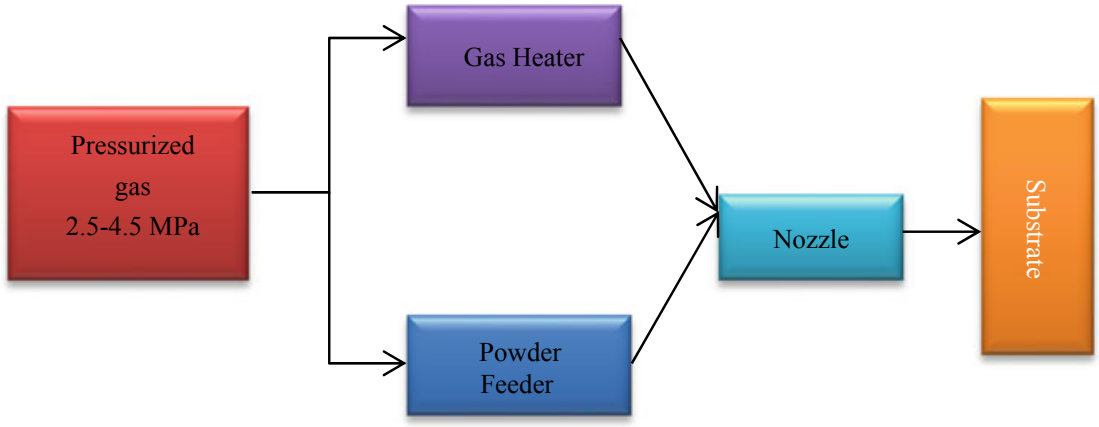


Fig. 34.1 High-pressure cold gas dynamic spray system

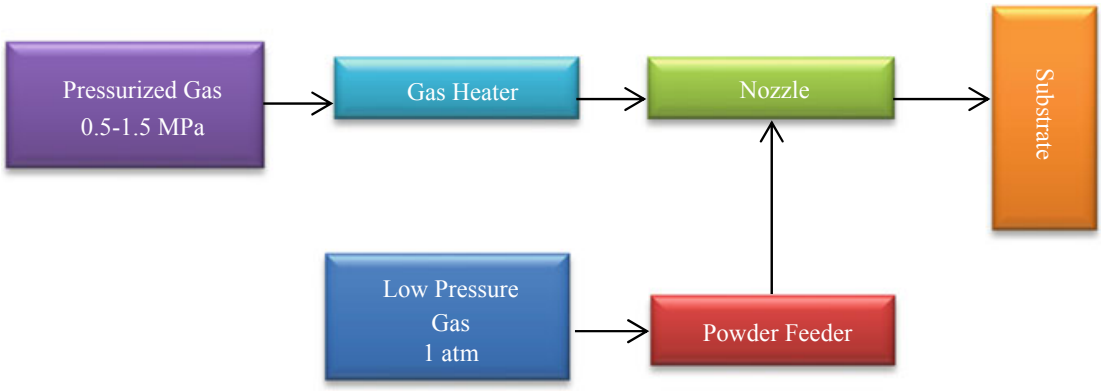


Fig. 34.2 Low-pressure cold gas dynamic spray system

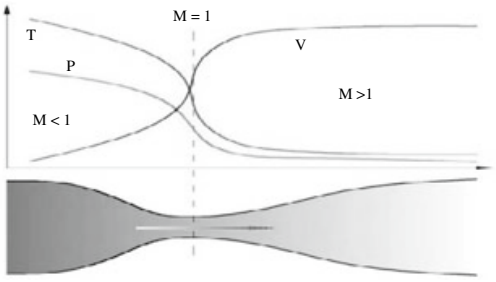


Fig. 34.3 Pressure, velocity and temperature variation in convergent–divergent nozzle

Generally, the pressure of working gas is between 0.5 and 1 MPa, and the flow rate is about 0.5–2 m^3/min. In LPCS, the design of the nozzle at exit is limited to Mach number less than 3, and inlet pressure is also less 1 MPa. If it exceeds, then atmospheric pressure is not able to supply powder particles into the nozzle. Due to this restriction, only low velocities particle reached into the nozzle. There are different variants of the standard apparatus of cold spray like kinetic metallization, pulsed gas dynamic spraying and vacuum cold spray.

KM is a variant of cold gas dynamic spraying that uses a convergent nozzle to achieve Mach number 1 at the exit of the gas velocity and small divergence for the compensation of friction impact. The deposition efficiency of micro- or nano-sized particles either preheated or at normal temperature solely depends on the mechanical and thermal properties of the engineering materials [10] (Fig. 34.3).

34.2 Numerical Methodology

Modelling of high velocity has been done by ABAQUS/Explicit using the code of finite element analysis (FEM). FEM has been commonly adopted for impact analysis in the kinetic spray. Adiabatic condition has been assumed for all kind of impact phenomenon. Particle is in range of 10–20 μm with the axis-symmetric condition for the impacting single particle on the substrate. Aluminium, copper and titanium have been used for the high-velocity impact simulation. Both the particle and substrate have not been preheated, and the impact problem was assumed to be adiabatic. Elements for the geometries are four-node bilinear quad elements and also with reduced integration and hourglass control (CAX4R) [11]. Reference strain rate for all the materials is taken as 1. Johnson–Cook (J-K) plasticity approach has been used which deals with thermal softening, strain and strain rate hardening. The elasticity response assumed to be linear for both the materials. Considerations of gravity, friction and rebounding force have been neglected in this study. The

damage parameters are not considered in this study. For the J-K approach, the values of equivalent plastic strain (PEEQ) solely depend at the integration points. An isotropic condition has been assumed for both thermal and mechanical properties of the material. The boundary conditions for constraining the substrate and particle motion have been shown in Fig. 34.4. For both the particle and substrate, axis-symmetric boundary condition has been used. Fixed boundary condition is used for the right side of the substrate's surface. J-K plasticity approach has been used for the plastic deformation of the particle as well as the substrate. This model gives the rate dependency of material during plastic deformation. Thermal softening, strain hardening and strain rate sensitivity are also defined for the model. The equivalent flow stress is given by as

$$\sigma_e = \left(A + B\epsilon_p^n\right)(1 + C\ln\epsilon^*)(1 - (T^*)^m) \tag{34.1}$$

where A and B are yield stress and work hardening parameter of the material. T^* is the homologous temperature. Different boundary conditions were used from ABAQUS/Explicit, i.e. symmetry about a plane where X is constant (XSYMM) and ENCASTRE where $U1 = U2 = U3 = 0$ (Table 34.1).

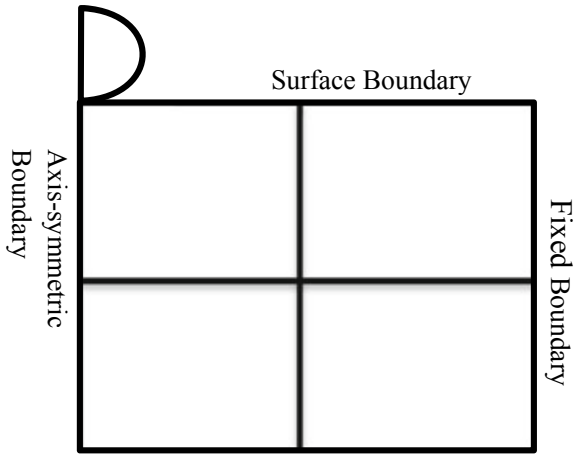
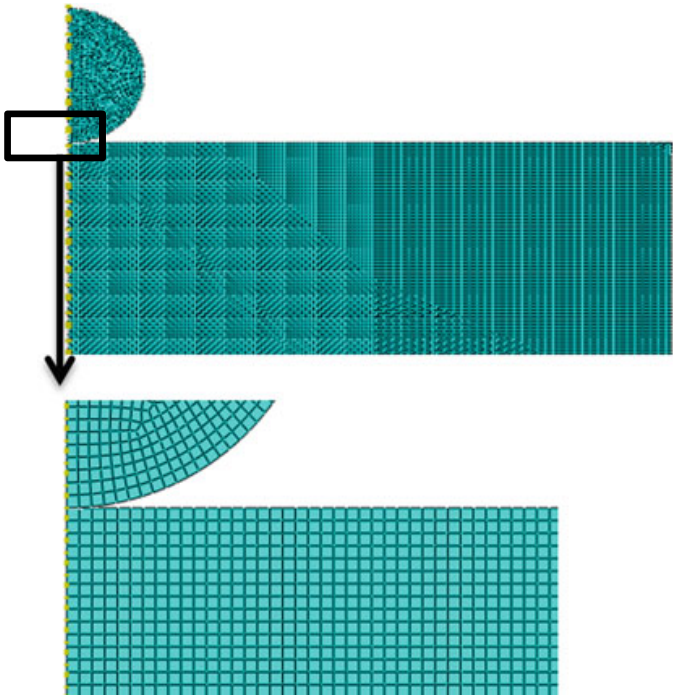


Fig. 34.4 Boundary conditions for the axis-symmetric single particle impact

Table 34.1 Properties of different materials

Material	Cu	Al	Ti
Density (kg/m^3)	8960	2700	7870
Heat capacity (J/kgK)	383	920	528
Young's modulus (GPa)	124	65.762	116
Poisson's ratio	0.34	0.3	0.34
A (MPa)	90	148.361	806.57
B (MPa)	292	345.513	481.61
N	0.31	0.183	0.319
C	0.025	0.001	0.0194
M	1.09	0.859	0.655
Melting point (K)	1356	930	1923
Reference temperature (K)	298	298	298



(a) Mesh Distribution in axis-symmetry Single particle impact. Number of elements and nodes are 12840 and 13075 respectively for 20 micrometer copper particle.

(b) Magnified view of Mesh at the contact zone of particle and substrate.

Fig. 34.5 Meshing of axis-symmetric impact behaviour of cold spray

Meshing plays a very crucial role in every computational problem. It also plays a vital role in the estimation of computational cost. Different meshing has been used like the diameter of particle (d_p)/32, d_p/35 and d_p/40. It was found that after d_p/40, there were no improvements in the results occurred. So, d_p/40 meshing size as shown in Fig. 34.5 has been used for all the computational impact problem in cold gas dynamic spray.

34.3 Results and Discussion

Copper particle having a diameter of 10 μm at a diameter of particle (d_p)/40 meshing at different velocities and also different values of equivalent plastic strain (PEEQ), compression ratio and flattening ratio, etc., were found for checking the adequate bonding between them. Figures 34.6 and 34.7 show that the flattening ratio and compression ratio are totally depended on the velocities and mechanical properties. Figure 34.8 also shows that values of PEEQ solely depend on the velocity of particle and substrate materials. Thermal properties also play a very important role for the deformation of the particle, i.e. if the specific heat of the material is low, then the particle will plastically deform at low velocity.

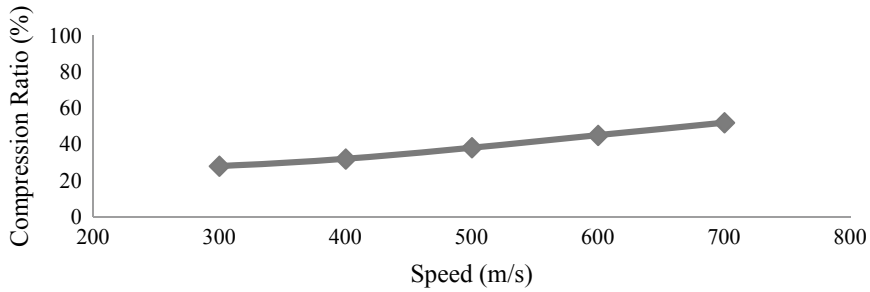


Fig. 34.6 Variation of compression ratio vs increasing the speed of 10 µm copper particle with meshing size $d_p/40$ using the lagrangian approach

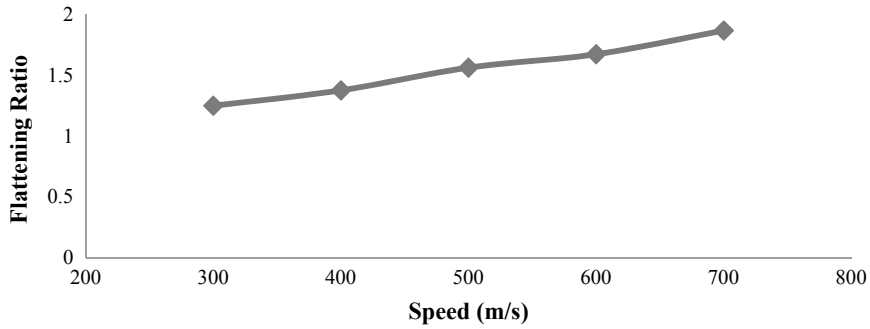


Fig. 34.7 Variation of flattening ratio by increasing the speed of 10 µm copper particle with meshing size $d_p/40$ using the lagrangian approach

34.3.1 Deformation Behaviour of Soft Particle on Hard Substrate

It has been seen from Fig. 34.9 that the deformation behaviour of a soft particle like aluminium when impacted on the hard substrate than splattering phenomenon occured. When micro-sized particle impacts at a certain angle like 30°, 45° and 60° especially when the particle made of soft material, splattering of the particle occurs on the surface of substrate with minute crater.

Formation of the jet by the deformation of the particle also been seen that removes the oxide layer from the surface. As shown in Fig. 34.9, when 20 µm aluminium particle is impacted at a different angle, then deformation behaviour is totally different from the perpendicular impact.

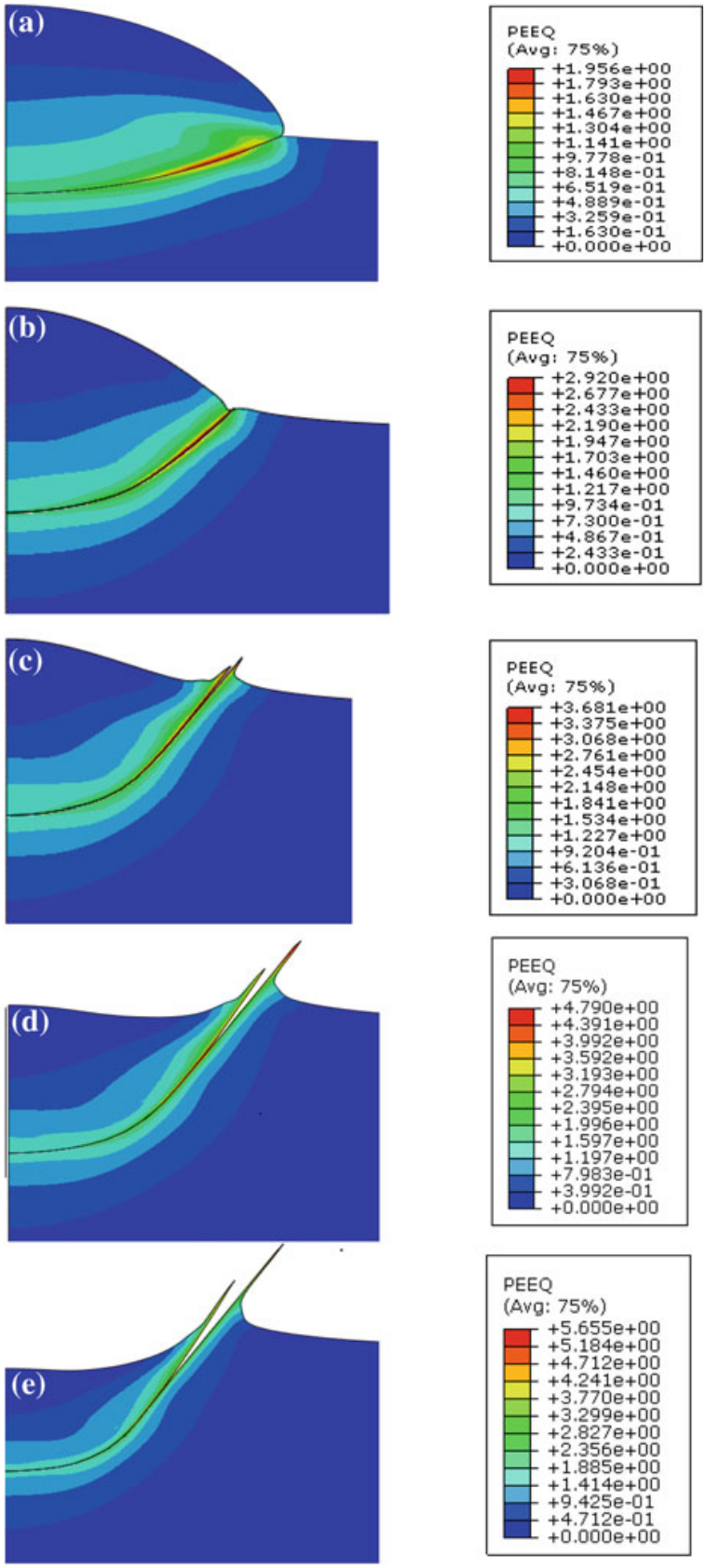


Fig. 34.8 Contour of 10 µm diameter of the copper particle with copper substrate and PEEQ at different velocities **a** 300, **b** 400, **c** 500, **d** 600 and **e** 700 m/s respectively

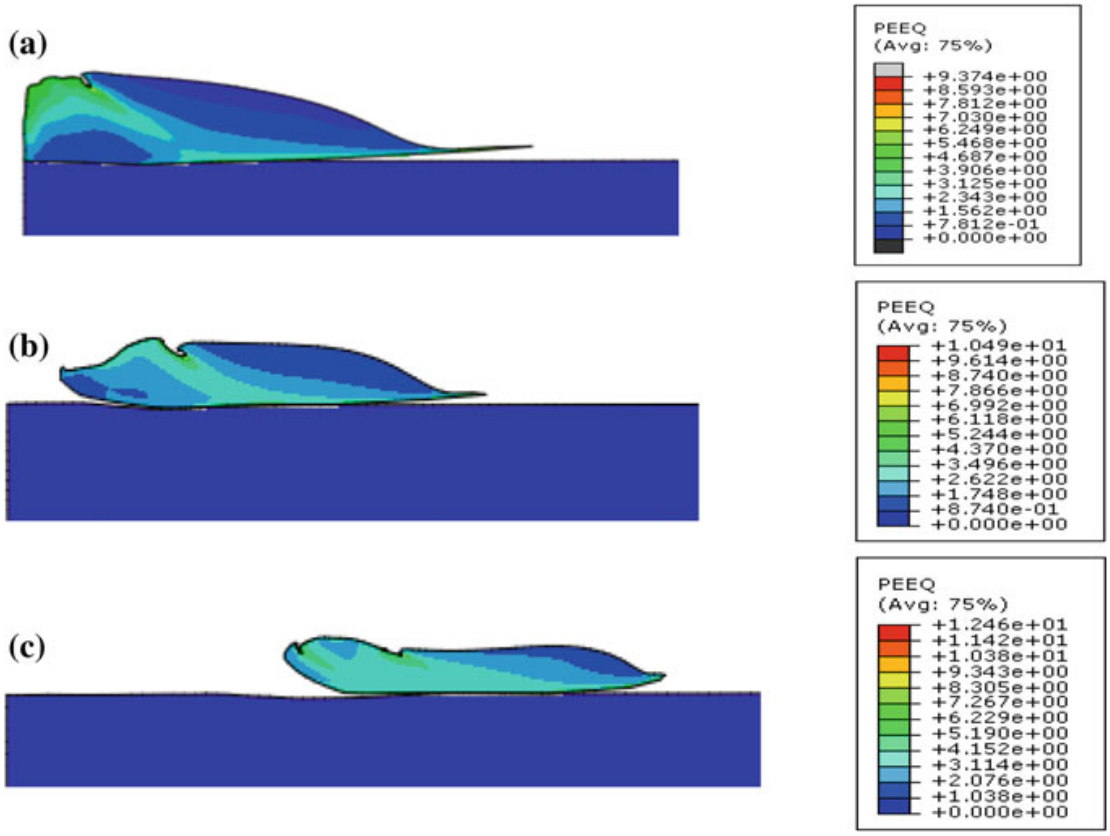


Fig. 34.9 Contour and values of PEEQ for 20 μm aluminium particle on titanium substrate at an angle of **a** 30°, **b** 45° and **c** 60° at 750 m/s, respectively

34.4 Conclusion

Different ranges of micro-sized particles have been probed during the plastic deformation process using kinetic spray technology. It has been investigated that for same particle and substrate, the critical velocity which is required for proper bonding, contour and crater formation and also for the adequate penetration is in the range of 500–560 m/s. Critical velocity has been very important for the adequate metal coating on the substrate by cold gas dynamic spray.

On the contrary, using soft particle and impacting on the hard substrate, it has been probed that huge amount of increment in the critical velocity or adequate penetration requires more particle velocity as compared to same particle/substrate. In this impact condition, homogeneity and uniformity of metal coating have occurred which indicates better coating. Lastly, the inclination angle also played a crucial role in the splattering and uniformity in metal coating using cold spray.

References

1. King, P.C., Bae, G., Zahiri, S.H., Jahedi, M., Lee, C.: An experimental and finite element study of cold spray copper impact onto two aluminum substrates. J. Therm. Spray Technol. **19** (3), 620–634 (2010)
2. Alkhimov, A.P., Kosarev, V.F., Papyrin, A.N.: A method of cold gas-dynamic deposition. Dokl. Akad. Nauk SSSR **35**(12), 1062–1065 (1990)
3. Yin, S., Suo, X., Guo, Z., Liao, H., Wang, X.: Deposition features of cold sprayed copper particles on preheated substrate. Surf. Coat. Technol. **268**, 252–256 (2015)
4. Li, W., Liao, H., Li, C., Li, G., Coddet, C., Wang, X.: On high velocity impact of micro-sized metallic particles in cold spraying. App. Surf. Sci. **253**, 2852–2862 (2006)
5. Yin, S., Wang, W., Li, X., Li, Y.: Numerical study on the effect of substrate size on the supersonic jet flow and temperature distribution within the substrate in cold spraying. J. Therm. Spray Technol. **21**(3–4), 628–635 (2012)
6. Grujicic, M.: Adiabatic shear instability based mechanism for particles/substrate bonding in the cold-gas dynamic-spray process. Mater. Design. **25**, 681–688 (2004)
7. Klassen, T., Kreye, H.: On parameter selection in cold spraying. J. Therm. Spray Technol. **20** (6), 1161–1176 (2011)
8. Yin, S., Liu, Q., Liao, H., Wang, X.: Effect of injection pressure on particle acceleration, dispersion and deposition in cold spray. Comput. Mater. Sci. **90**, 7–15 (2014)
9. Assadi, H., Ga, F., Stoltenhoff, T., Kreye, H.: Bonding mechanism in cold gas spraying. Acta Mater. **51**, 4379–4394 (2003)
10. Klinkov, S.V., Kosarev, V.F.: Measurements of cold spray deposition efficiency. J. Therm. Spray Technol. **15**(9), 364–371 (2006)
11. ABAQUS/CAE User Manual 6.14. Hibbitt, Karlsson and Soerensen, Pawtucket

Chapter 35
Towards an Improved Understanding of Stereolithography Process—A Computational Study

Usharani Rath and Pulak M. Pandey

Abstract Amongst various futuristic manufacturing technologies, additive manufacturing is fast emerging from the lap of laboratory-scale testing to full-scale commercial applications. Stereolithography of polymeric materials has the potential to offer cost-effective and accurate solutions for many industrial sectors such as automotive, aerospace and electronics. However, the process of understanding and comprehending the underlying physics of the liquid additive manufacturing process is still in its infantile stage. In the present work, an attempt was made to computationally model the laser beam and the photosensitive resin interaction in the stereolithography process. The laser source was considered as a Gaussian heat source, and the liquid vat was considered as a homogenous isotropic fluid domain. Finite element modelling approach was utilized to evaluate the thermal changes occurring during the photo-initiated curing process. The resultant process quality indicators such as transient temperature distribution across the liquid vat, conductive and radiative heat flux, laser intensity on the liquid layer and its diminishing impact across the layer thickness were computed and compared. Maximum local temperature as high as 390 K was found on the irradiated location which cools down rapidly due to heat conduction and diffusion. However, the temperature reduces exponentially with the height of resin vat, and it takes around 4 s for the irradiated point to cool down to the ambient temperature and gets cured. Approximately 1.1×10^6 W/m^2 of laser intensity is produced on the bottom surface of laser. Some selected results were validated with contemporary literature and found to be satisfactorily consistent.

Keywords Additive manufacturing · Stereolithography · Finite element modelling

U. Rath · P. M. Pandey (✉)
Rapid Prototyping Laboratory, Department of Mechanical Engineering,
Indian Institute of Technology Delhi, New Delhi 110016, India
e-mail: pmpandey@mech.iitd.ac.in

U. Rath
e-mail: usha.rrath@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_35

Nomenclature

ρ	Density of PMMA resin
C_p	Specific heat capacity of PMMA resin
K	Thermal conductivity of PMA resin
Q	Heat generated during photo-curing reaction
I_s	Laser intensity at the bottom surface of resin vat
I_0	Original laser intensity
s	Coordinate of irradiated point
w	Laser spot width
I_Z	Laser intensity at a height z from the bottom surface of resin vat
D_P	Penetration depth
n	Normal direction to boundary 1–5
q	Heat flux due to radiation
ϵ	Emissivity
σ	Stefan Boltzmann's constant
T_{amb}	Ambient temperature
T	Transient temperature

35.1 Introduction

3D printing technology has been a watershed across the manufacturing industry in the last two decades. Some distinct features, such as ability to produce any complex three-dimensional geometries from a CAD file employing a layer-by-layer approach, zero material wastage, nearly net shape production of parts and enhanced automation leading to improved process capability, make these set of additive technologies futuristic and transformational. This solid free-form fabrication technique uses layer-by-layer approach to obtain physical objects from a CAD model by generating material deposition path or laser scanning path.

Stereolithography is one of the pioneers and widely used 3D printing technique utilizing photosensitive liquid resin as the feedstock. The pre-programmed laser cures the liquid resin vat in a layer-by-layer approach resulting from the chemical crosslinking of the monomer resin to produce solid 3D polymeric components as designed by the input CAD model. It produces physical 3D objects with highest precision and accuracy as compared to other commercially available 3D printers such as selective laser sintering or fused deposition modelling.

A good number of research expeditions are reported on stereolithography process and its capabilities. However, the process modelling and analysis of underlying physics of stereolithography process are still in its early stages, and very few research contributions are reported on the same. Huang et al. [1] generated dynamic finite element code for evaluating the shrinkage effects in constrained-surface stereolithography process. They suggested that the surface tension of liquid resin

and shape of physical model may give rise to the inaccuracies of 3D-printed parts. Lee et al. [2] utilized artificial neural network to build up a model that can predict the effects of process parameters on the build accuracy of stereolithography process. They also mentioned that understanding the process physics can improve the process performance of the photo-curing process. They concluded that reduced layer thickness and increased hatch spacing improved the part performance. Bartolo [3] compared the phenomenological approach with the mechanistic approach for the stereolithography process. They suggested that the photo-curing process is an exothermic reaction increasing the local temperature of the irradiated region which gradually diminishes through conduction, convection and diffusion. Furthermore, they also implied that the fraction variation contour has similar resemblance as that of laser intensity which is conical. It was also explained that the local temperature and light intensity can lead to increased curing reaction and reduced printing time. Jiang et al. [4] utilized a dynamic finite element-based method to anatomize the shrinkage effects occurred during conventional stereolithography and mask-based stereolithography processes. They could also investigate the presence of residual stresses during the stereolithography process. Emami et al. [5] developed an analytical model for scanning-projection stereolithography process. They evaluated and compared the total laser energy intensity received by projection stereolithography and scanning-projection stereolithography processes during curing of each layer. They also indicated that SPSL process produces better accuracy than the PSL process. Liravi et al. [6] adopted cohesive zone model and conducted finite element simulation to predict the separation force existing between the cured part and the resin tray. Bae et al. [7] proposed a segregation model to determine optimum process parameters for direct ceramic stereolithography. The model was developed from the degree of segregation and the time constants such as writing time and settling time. Bikas et al. [8] recommended that the major cause of poor quality and inaccurate parts in additive manufacturing lies in the limited available modelling and optimization approaches.

In the light of the literature survey, it was realized that the quantum of research on the modelling and simulation of stereolithography process is scanty. However, from the research contributions by Bartolo [3] and Jiang et al. [4], it has been established that modelling of the stereolithography process can lead to better understanding of the underlying physics and improve the process performance. However, none of the research papers reported the transient temperature distribution along the resin surface and along the depth. Furthermore, findings regarding the laser intensity variation along the depth of resin are also very few.

Thus, in the wake of the aforementioned literature explorations, the objective of the present research investigation was formulated to vindicate some of the underlying physics of the stereolithography process. In the present study, the finite element analysis of the bottom-up stereolithography process was carried out for a single-layer curing of resin vat. As an outcome, the transient temperature distribution, laser intensity variation, conductive and convective flux contours, transient residual stress distribution and variation on the resin surface were evaluated and analysed for improved part quality and better process performance.

35.2 Process Modelling

A finite element-based commercial multi-physics software package, COMSOL, which operated under Windows 64-bit operating system, was used for the simulation study. The physical model of the present study primarily consists of a methyl methacrylate resin vat of 10 cm length, 5 cm width and 3 mm height. The fluid properties of resin and specifications of laser used in bottom-up stereolithography set-up are presented in Tables 35.1 and 35.2. The UV laser was considered as a moving heat source which travels across the bottom surface of resin vat. The X-coordinate of laser beam follows a ramp function with a slope of 1/50, so that the laser moves 10 cm distance along X-direction in 5 s.

35.2.1 Assumptions

There are broadly five assumptions considered in the present modelling expedition in line with the spirit of some of the reviewed literature [3, 4] and the contemporary understanding of stereolithography process.

1. The model was considered as a homogenous and isotropic fluid domain.
2. The effect of optical scattering was not considered.
3. Convection effects were taken as negligible as air is quiescent inside the stereolithography chamber.
4. The heat generation was considered to be only due to heat of polymerization as the stereolithography process is governed by photo-curing reaction.
5. The absorption of UV laser intensity was considered to be governed by Beer–Lambert's law.

Table 35.1 Resin fluid properties

Serial number	Properties	Value
01	Density (g/cm^3)	1.12
02	Dynamic viscosity (Pa s)	890
03	Thermal conductivity (W/mK)	0.25
04	Specific heat (J/kgK)	1500
05	Ambient temperature (K)	298.15
06	Emissivity	0.9

Table 35.2 Laser specifications

Serial number	Specifications	Value
01	Power (mW)	120
02	Wavelength (nm)	405
03	Spot size (mm)	0.155
04	Radiation	Ultraviolet

35.2.2 Governing Equations

A set of governing equations were used in the current study to model the physics of heat transfer due to laser radiation. There are two equations governing the three-dimensional transient energy equation and absorption of UV radiation (Beer–Lambert's law) by the liquid resin vat, given by Eqs. 35.1, 35.2 and 35.3, respectively. The governing Equation 35.1 is solved considering the heat source 'Q' as a Gaussian heat source.

$$\rho C_{\mathrm{p}} \frac{\partial T}{\partial t} + \rho C_{\mathrm{p}} u \nabla T + \nabla \cdot (-K \nabla T) = Q \tag{35.1}$$

$$I_{\mathrm{s}} = I_0 \exp[-2 \times (s/w)^2] \tag{35.2}$$

$$I_Z = I_0 \exp\left(\frac{-z}{D_{\mathrm{P}}}\right) \tag{35.3}$$

where ρ, C_{p} and K are the density, specific heat capacity and thermal conductivity of liquid resin used, respectively. Q is the heat generated from the photo-curing exothermic reaction. I_{s} and I_0 are the laser intensity at the resin bottom surface and the original laser intensity, respectively. s and w are the coordinate of irradiated point and laser spot width, respectively. I_Z is the laser intensity at a height z from the bottom surface of resin vat. D_{P} is a material property of the photo-polymeric resin, known as penetration depth.

35.2.3 Beer–Lambert's Law

Beer–Lambert's law is utilized to establish a relation between the depth of penetration of a light source in a material and the attenuation coefficient of the material. The attenuation coefficient of the material is inversely proportional to the material property $(D_{\mathrm{P}})$of the medium through which the material is travelling. A material or medium with higher attenuation coefficient is supposed to have reduced depth of penetration of light source travelling through it. Equations 35.2 and 35.3 govern the Beer–Lambert' law at the resin vat bottom surface and at a specific height from the bottom surface.

35.2.4 Domain and Boundary Conditions

Various boundary and initial conditions were applied in the present model for the UV radiation and subsequent phenomena such as transient heat conduction,

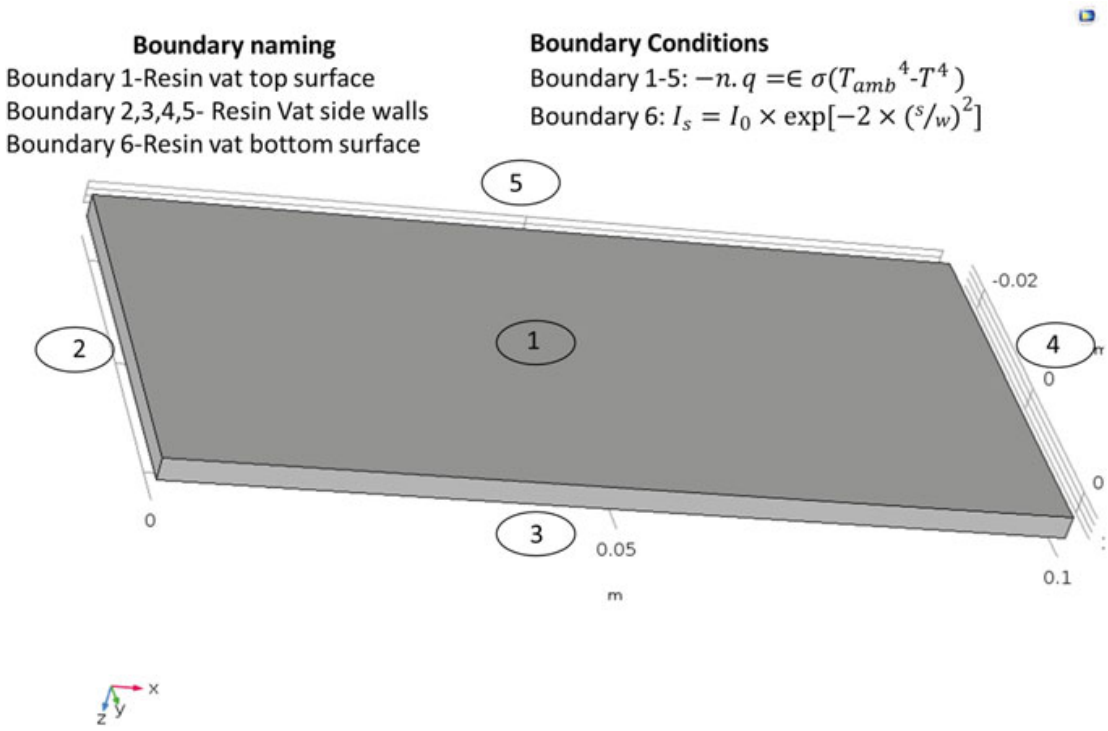


Fig. 35.1 Geometric model with boundary and initial conditions

convection and diffusion to happen. The boundary and initial conditions are clearly depicted from Fig. 35.1. A single rectangular fluid domain was considered for the simplicity of the model with six boundaries. Boundary 1 was considered as the top surface of the resin vat, and boundaries 2, 3, 4 and 5 were modelled as the side surfaces of the resin vat. Boundary 6 was considered as the bottom surface of the resin vat, on which the laser light was irradiated.

35.2.5 Meshing

A customized meshing chronology was used to discretize the three-dimensional geometric model of resin vat. Tetrahedral meshing was done with a maximum element size as 0.001 mm for the entire domain for the convergence of the solution. A total of 536,841 numbers of elements were generated. The model generation and succeeding analysis were done using COMSOL commercial software package. The simulation was performed for an irradiation time of 5 s with 0.1 as the time step and 0.001 as the relative tolerance of the simulation study.

35.3 Results and Discussion

The dynamic temperature variation on the bottom surface of resin vat is represented in Fig. 35.2. The maximum local temperature obtained during the curing process at the point of irradiation was predicted to be 390 K, which diminishes to ambient temperature by heat conduction and diffusion in a very short span of time. Furthermore, it can also be clearly visualized that the temperature rise due to irradiation at each irradiated point is maximum and reduces radially as per the Gaussian distribution. The variation of temperature with the height from the bottom surface of resin vat at a particular irradiated point is represented in Fig. 35.3. The decrease in temperature with height signifies that the absorption of UV radiation on resin vat is according to the Beer–Lambert's principle. Nonetheless, further slight increase in temperature might be caused due to the diminished conduction and diffusion as compared to the bottom surface. The variation of temperature with time at a particular irradiated point (boundary 2) on the resin bottom surface is also shown in Fig. 35.4. The curing profile of the photo-polymeric resin is in accordance with the star-weave build structure, which resembles a conical shape [9]. The laser intensity distribution follows a conical pattern as presented in Fig. 35.5, validating the work done by Bartolo [3]. Moreover, the laser intensity reduces from the bottom to top surface of resin vat in accordance with the Beer–Lambert's principle as represented in Eq. 35.3 and cures the resin irradiated beyond the critical laser intensity. The value of critical laser intensity varies depending on the resin and

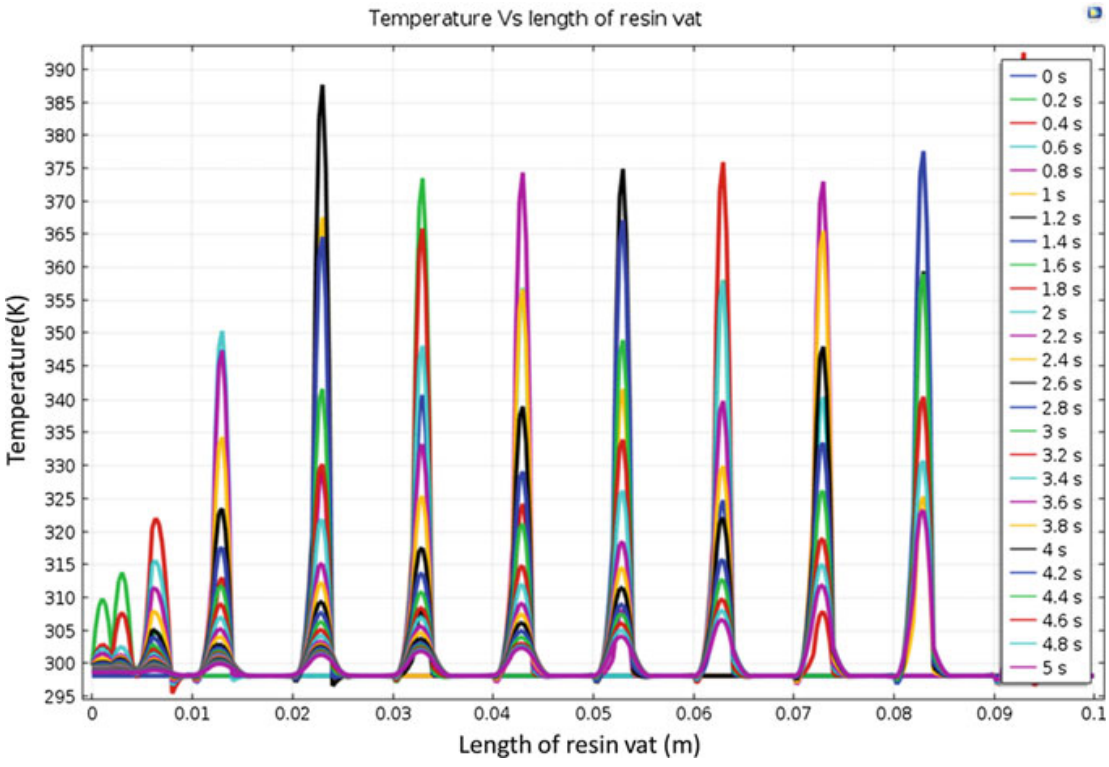


Fig. 35.2 Temperature variation on the bottom surface of resin vat

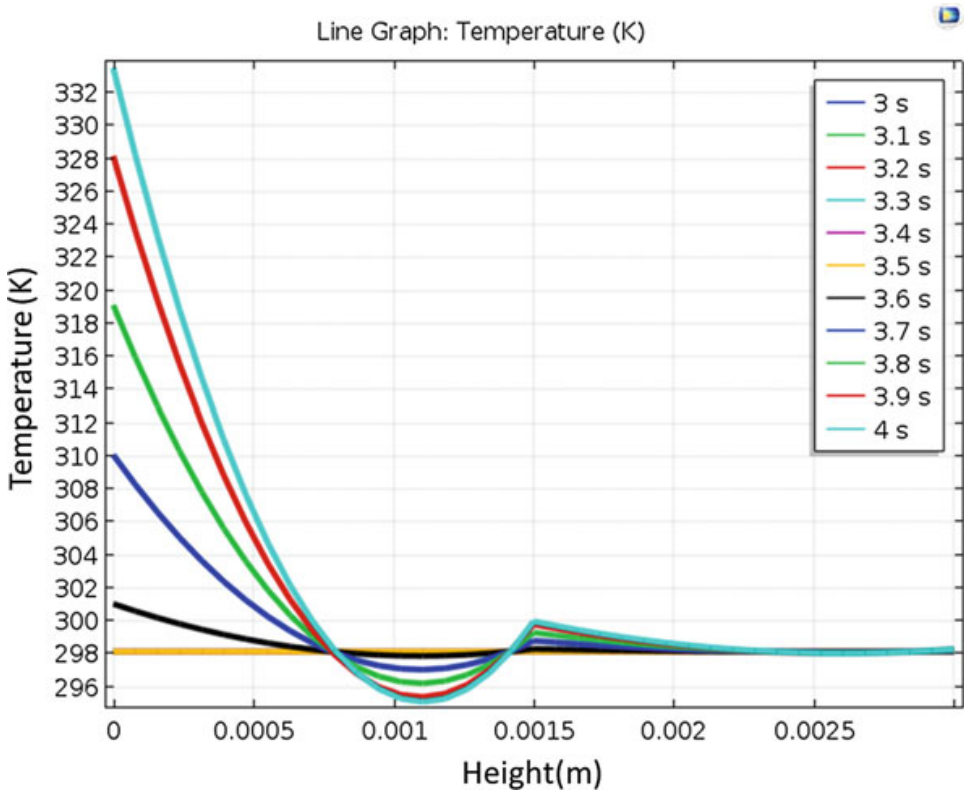


Fig. 35.3 Temperature variation with height of resin vat

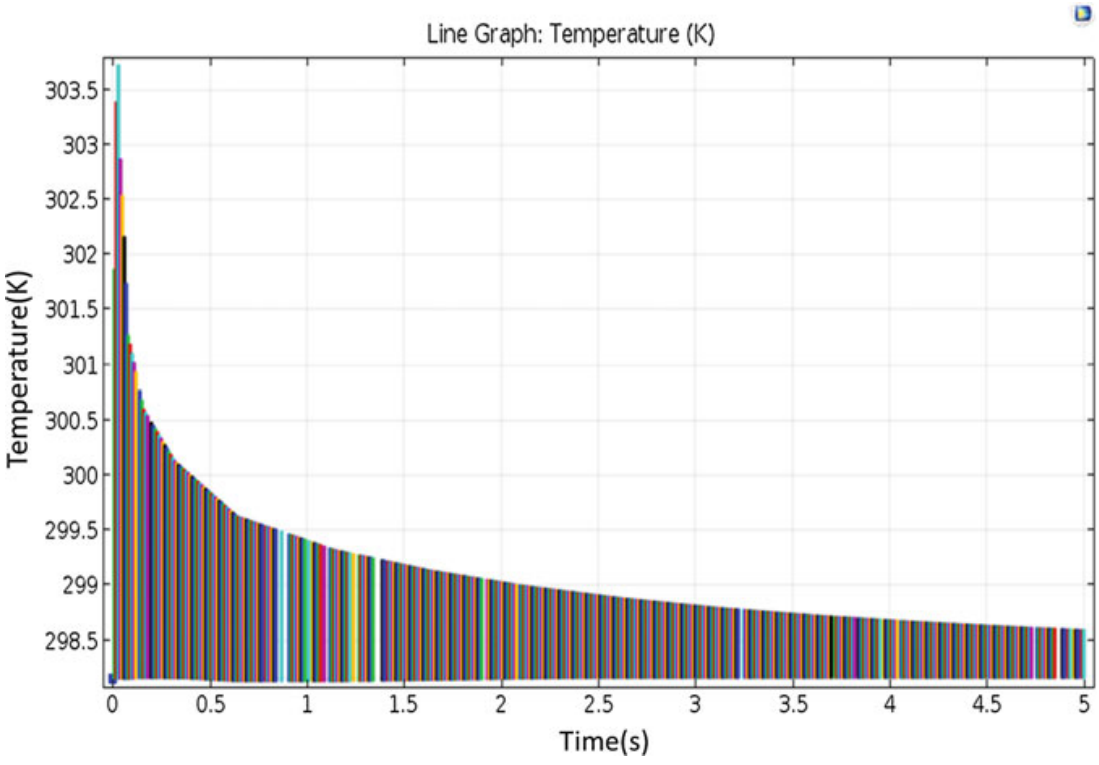


Fig. 35.4 Temperature variation with time

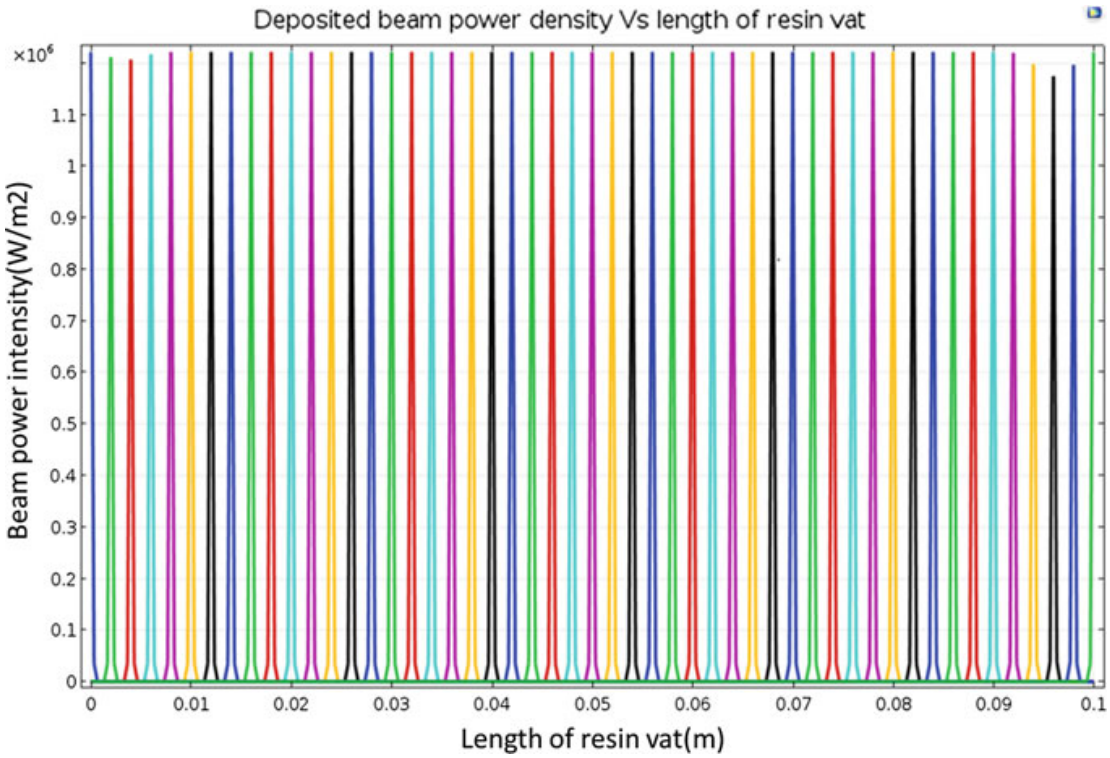


Fig. 35.5 Laser intensity variation on the bottom surface of resin vat

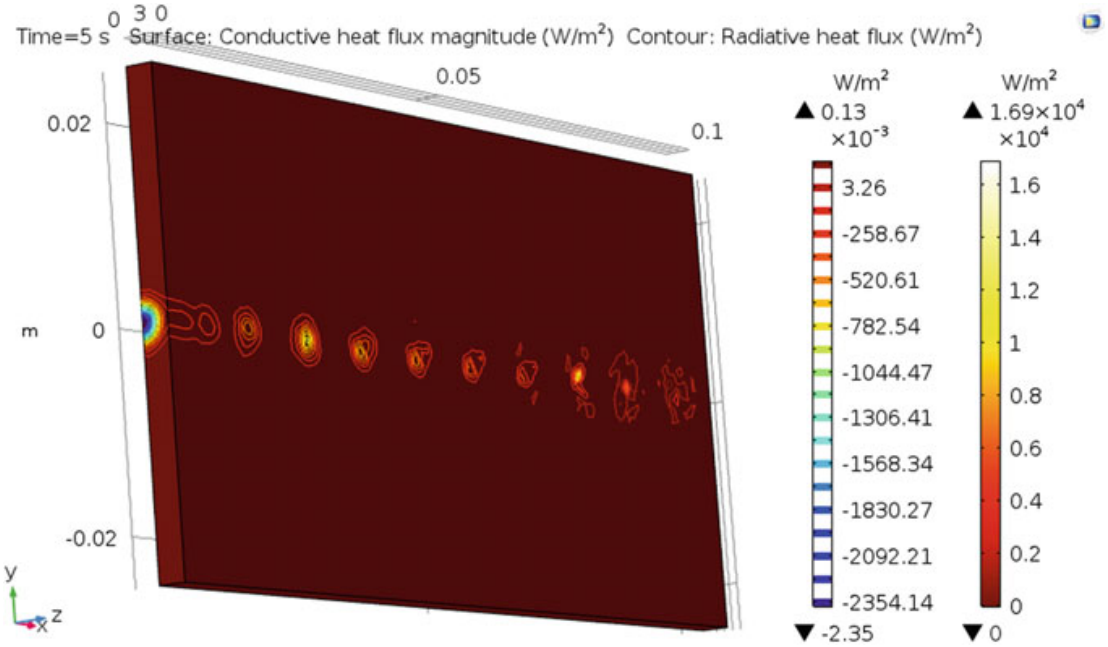


Fig. 35.6 Conductive heat flux contour

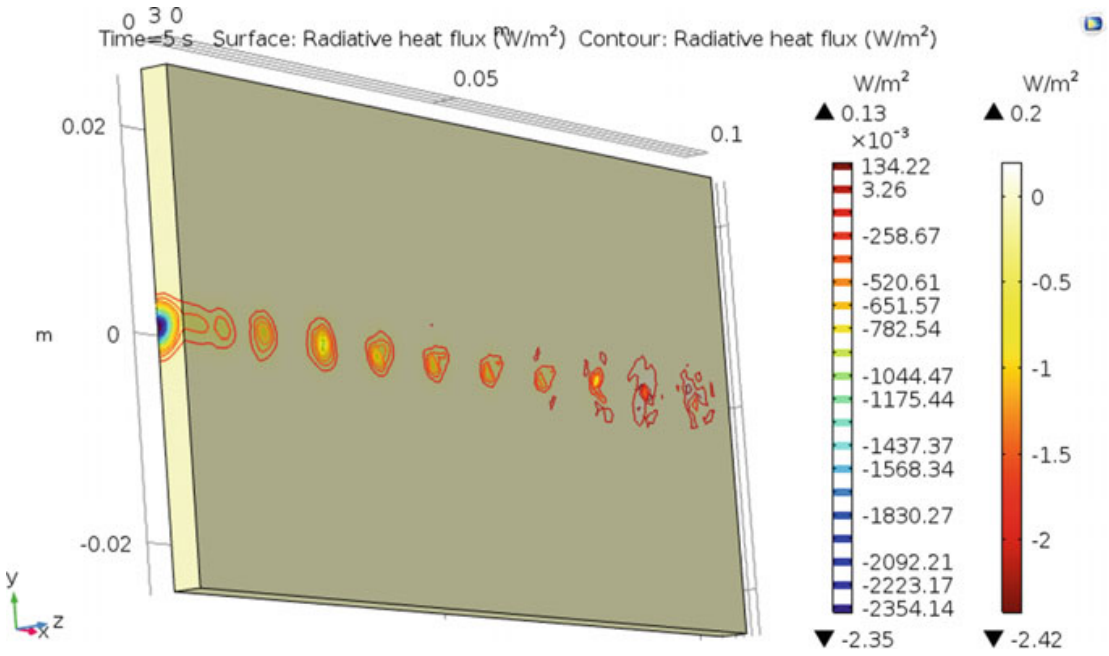


Fig. 35.7 Radiative heat flux contour

machine specifications. It can be evaluated by conducting a good number of experiments and plotting the semi-logarithmic graph between cure depth and logarithm of maximum exposed laser energy. The local heat generation and subsequent temperature rise due to exothermic photo-curing reaction decline expeditiously by heat conduction and diffusion. The conductive heat flux and radiative heat flux contours are depicted from Figs. 35.6 and 35.7, respectively.

35.4 Conclusions

A macroscopic model established by finite element method using the governing equations for the physical changes occurring during the stereolithography process was proposed. The following inferences can be deduced from the research work.

1. The transient temperature history on the surface and along the depth of the liquid resin irradiated by ultraviolet laser light with 405 nm wavelength was analysed.
2. Irradiated zone of resin undergoes a local increase in temperature as high as 390 K, which reduces quickly due to the advent of conduction and diffusion.
3. Laser intensity cures the laser across the depth in a conical manner, resulting in the prediction of fractional conversion contour of liquid resin to the green-state/semi-solid resin as conical [3]. The bottom surface of resin vat was irradiated with a laser intensity of 1.1×10^6 W/m^2.

The simulation approach can be improved further by considering the chemical kinetics of the photo-curing reaction, which involves the monomer and photo-initiator concentration in the liquid resin.

References

1. Huang, Y., Jeng, J., Jiang, C.: Increased accuracy by using dynamic finite element method in the constrain-surface stereolithography system. J. Mater. Process. Tech. **140**, 191–196 (2003)
2. Lee, S.H., Park, W.S., Cho, H.S., Zhang, W.L.M.C.: A neural network approach to the modelling and analysis of stereolithography processes. Proc. Inst. Mech. Eng. Part B J. Eng. Manuf. **215**, 1719–1734 (2001)
3. Bartolo, P.J.: Mechanistic approaches. Ann. CIRP **55**(1) (2006)
4. Jiang, C., Huang, Y.-M., Liu, C.-H.: Dynamic finite element analysis of photopolymerization in stereolithography. Rapid Prototyp. J. **12**(3) (2012)
5. Emami, M., Barazandeh, F., Yaghmaie, F.: An analytical model for scanning-projection based stereolithography. J. Mater. Process. Tech. **219**, 17–27 (2015)
6. Liravi, F., Das, S., Zhou, C.: Computer-aided design separation force analysis and prediction based on cohesive element model for constrained-surface stereolithography processes. Comput. Des. **69**, 134–142 (2015)
7. Bae, C., Halloran, J.W.: A segregation model study of suspension-based additive manufacturing. J. Eur. Ceram. Soc. **38**(15), 5160–5166 (2018)
8. Bikas, H., Stavropoulos, P., Chryssolouris, G.: Additive manufacturing methods and modelling approaches: a critical review. Int. J. Adv. Manuf. Technol. **83**, 389–405 (2016)
9. Rahmati, S., Ghadami, F.: Process parameters optimization to improve dimensional accuracy of stereolithography parts. Int. J. Adv. Des. Manuf. Technol. **7** (2014)

Chapter 36
Experimental Investigations for Effects of Raster Orientation and Infill Design on Mechanical Properties in Additive Manufacturing by Fused Deposition Modelling

Pushpendra Yadav, Ankit Sahai and Rahul Swarup Sharma

Abstract Manufacturing and designing of various functional components in the fields of engineering, automobile and aerospace are being done by additive manufacturing (AM) technology. These activities must be supported by the knowledge of process parameters that may influence mechanical properties of industrial products. Fused deposition modelling (FDM) is one of the most popular AM technologies in which the quality of part depends on selection of process variables. The present work focuses on two input process parameters, i.e. raster orientation and infill density and its effect on the mechanical properties of 3D-printed samples. The PLA samples were prepared in *XY* and *XZ* orientation with infill density of 20, 40, 60, 80 and 100%. The tensile testing is conducted to evaluate the effects of these process parameters on tensile strength, ultimate strength, etc. Further, fractography was performed and it is concluded that dimple size and distribution on fractured surface were affected by the infill density. Printing time was also checked and compared for specimens printed for different parameters, which shows printing time is least for *XY* orientation and maximum for *XZ* orientation. Printing time increases as the infill density increases from 20 to 100%.

Keywords 3D printing · FDM · Tensile testing · Infill density · Fractography

P. Yadav · A. Sahai (✉) · R. S. Sharma
Department of Mechanical Engineering, Faculty of Engineering, Dayalbagh Educational Institute Dayalbagh, Agra 282005, India
e-mail: sahaiankit@dei.ac.in

P. Yadav
e-mail: er.pushpendrayadav@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_36

36.1 Introduction

3D printing is considered as the future of the technology in the world and is a manufacturing phenomenon that will revolutionize industries, factories and mass production in the coming years. With the advent of Industry 4.0, 3D printing has become a buzzword in almost all industrial sectors. Although 3D printing technologies took birth in the 1980s, mainly by the name of rapid prototyping, they got a fresh impetus in this century and have been growing at an exponential rate in last few years [1]. This technology allows designers to quickly and inexpensively produce prototypes without tooling or machining which act as catalysts to students, researchers, scientists and innovators to rapidly produce prototypes based on their imagination and thus 3D printing is helping users across all age groups to enhance product manufacturing by simplification in working. Some economists have considered this a third industrial revolution following mechanization and assembly line production [2]. This industry has grown quickly in recent years to an annual value of over $5 billion. Sales of 3D desktop printers were 160,000 in 2014 and grew to nearly 300,000 in 2015 [3]. Some predictions have shown that the 3D printing total market revenue will be between $7 and $23 billion per year by 2020. Because of this massive expansion and growing significance, FDM 3D printing is being studied in the areas of medical, aerospace, building equipment, materials, preparation techniques, etc. and is still the subject of research [4–8]. 3D printing is far different from traditional manufacturing processes and saves lots of build time, production waste and can be easily worked by ordinary workers as compared to skilled machining, forging, casting, etc. where skilled workers are required to complete the job. But with any manufacturing process, it is important to focus on resource saving and environmental issues and with the new technologies coming up in 3D printing, it is necessary to pay more attention to its sustainability in context to the present scenario before we integrate 3D printing with industrial set-up. The sustainability studies will also help scientists, researchers and other decision makers to understand the 3D printing process and its implications in different applications or products [7]. Therefore, it is extremely important to understand the influence of process variables or parameters on product development. These variables can be classified in: material properties, build specifications, part positioning and orientation and environment [8]. A similar study was developed by Ziemian et al. [9] in which the group evaluated the effect of raster orientation in the direction of the strain and developed the optimal levels for improving the strength of the specimen. The most studied parameters, in a static situation, have been build orientation, raster angle, raster width, raster-to-raster air gap and nozzle diameter or slice height [10, 11]. Some researchers have also analysed the effect of 3D printing in Engineering education which is suitable for classrooms and public spaces and encourages action-based learning, exploration and innovation, to accelerate the acceptance and involvement of 3D printing technologies in the manufacturing sector [12, 13].

Not many people have focused the research on the effect of infill density on the mechanical properties of PLA materials for different raster orientations. Therefore, the present work focuses on the effect of infill density on mechanical properties in *XY* and *XZ* orientation and its fractography analysis using scanning electron microscopy (SEM).

36.2 Fused Deposition Modelling Process

One of the most exciting and widely used additive manufacturing technologies is fused deposition modelling (Fig. 36.1) which is based on the material extrusion process. In FDM process, the polymer filament is fed to the liquefier chamber where it gets heated to a particular temperature depending upon the melting point of the material used for printing. Once the material is heated, it gets softens and melts and after that molten material is pushed through the nozzle from the liquefier with some suitable velocity. The extruded material in molten form is deposited on the platform layer by layer which slowly gets cooled at room temperature. The layer thickness and other parameters depend upon the G-code generated by the slicing software.

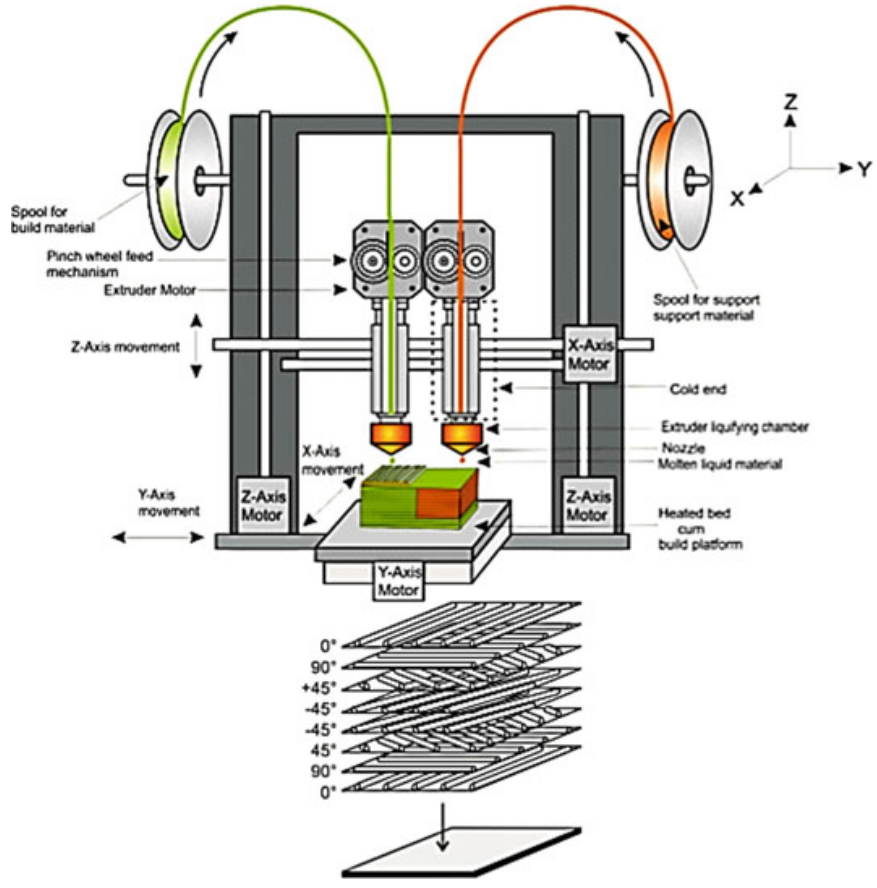


Fig. 36.1 Complete process of FDM process machine set-up

These 3D printers working in FDM technology are easy to design, build and quick to work upon. The students, researchers and even normal household users can easily install and work to produce the desired products using the different parameter combinations and of different shapes and sizes. These printers can be a great learning experience as the users can visualize, model and print in quick succession to get the rapid prototype of the specimen or product.

36.3 Experimental Details

In the present work, raster orientation and infill density are input parameters and based on their different combinations, mechanical properties like tensile strength and yield strength are calculated. The tensile test specimens were prepared based on ASTM D638, standard dimensions are used corresponding to Type-1 specimen in Fig. 36.2, and each sample set consisted of five specimens, Fig. 36.3. The geometry of the 3D-printed specimens was modelled using SolidWorks software exported as an STL file and imported to the 3D printing software. A 50 KN universal testing machine was used for tensile testing with speed of 5 mm/min, and fractography was performed using scanning electron microscope (SEM). The material used for current experiment is PLA.

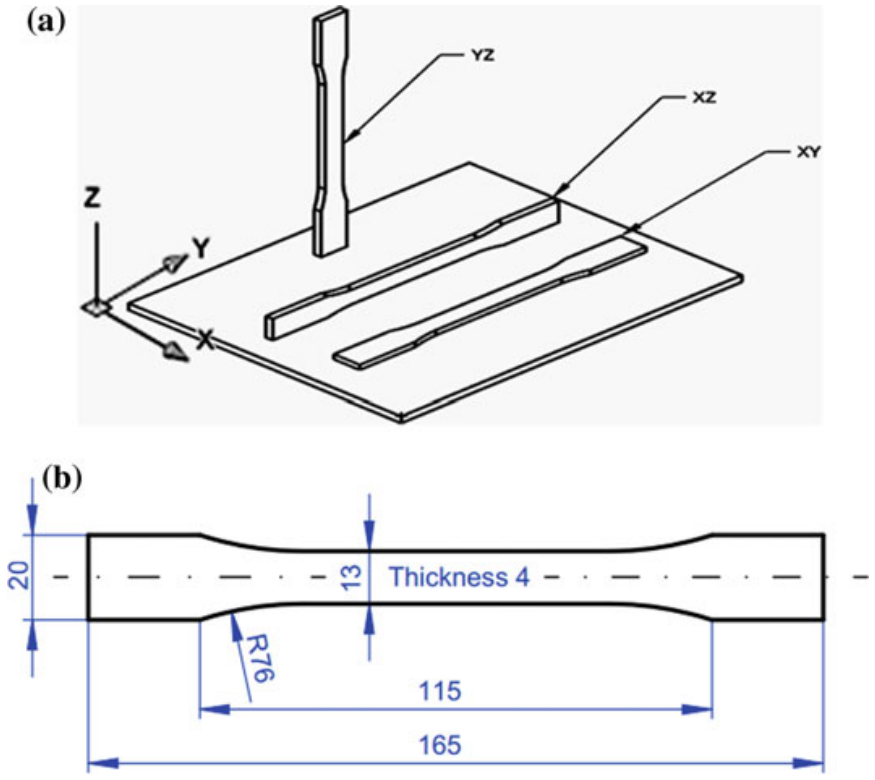


Fig. 36.2 Different raster orientations *XY*, *XZ*, *YZ* and ASTM D638 Type-1 tensile test specimen

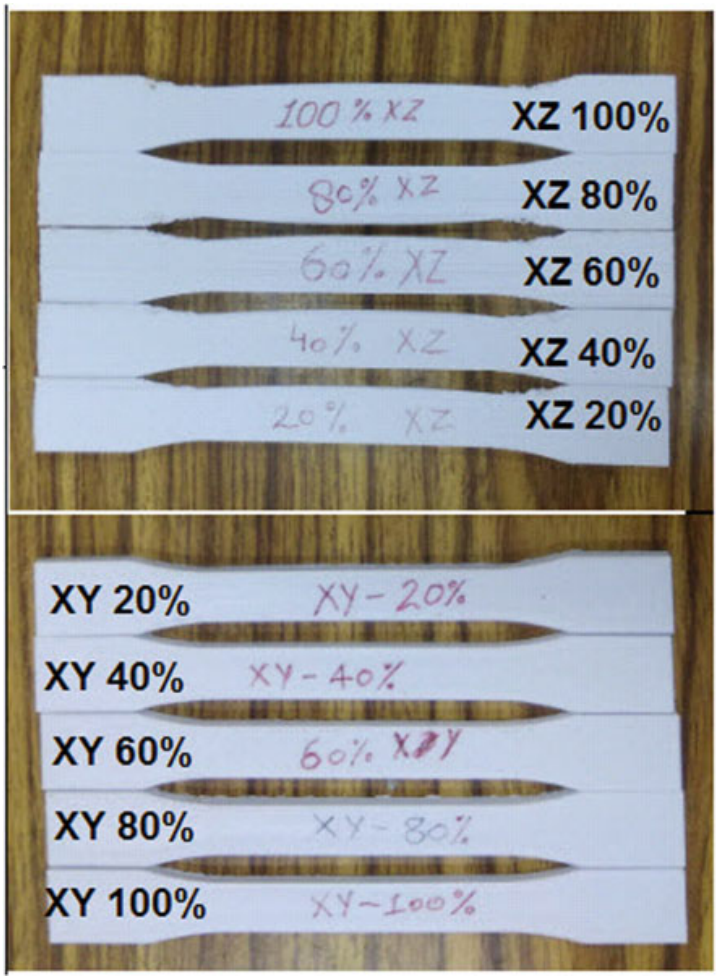


Fig. 36.3 Printed specimens in each orientation of *XY*-and *XZ*-directions

Before performing fractography, tensile tested samples were cut and gold (75%) and palladium (25%) coating was done for better image analysis. During experimentation, mechanical properties were evaluated for raster orientation *XY* and *XZ* as mentioned in Tables 36.1, 36.2 based on certain fixed parameters such as infill pattern, layer thickness, temperature and air gap while infill density and raster orientation were changed, Table 36.3.

Table 36.1 Mechanical properties of 3D-printed PLA in *XY*-raster orientation and different infill densities

Raster orientation—*XY*					
Infill density (%)	20	40	60	80	100
Area(mm^2)	52	52	58.08	52.8	52
Gauge length (mm)	50	50.071	50.062	50.091	50
Peak load (kN)	1.183	1.401	1.424	1.859	2.345
Peak stress (MPa)	22.746	26.946	24.518	35.203	45.089
Yield stress (MPa)	17.583	18.81	19.754	28.476	34.428
Yield load (kN)	0.914	0.978	1.147	1.504	1.79
Elongation at break (%)	7.435	6.922	4.4	3.1	14.678

Table 36.2 Mechanical properties of 3D-printed PLA in *XZ*-raster orientation and different infill densities

Raster orientation—*XZ*					
Infill density (%)	20	40	60	80	100
Area (mm^2)	50.92	54.6	54.8	49.4	52
Gauge length (mm)	50	50	50	50	50
Peak load (kN)	1.5	1.652	1.836	1.939	1.684
Peak stress (MPa)	29.46	30.254	33.496	39.256	32.382
Yield stress (MPa)	22.212	23.492	24.896	31.33	28.028
Yield load (kN)	1.131	1.283	1.364	1.548	1.457
Elongation at break (%)	5.127	6.422	5.586	4.818	3.25

Table 36.3 Printing parameters of PLA 3D printing

Fixed variables		Control factors	
Speed	60 mm/s	Infill density	20%
Layer thickness	0.2 mm		40%
Number of shells	3		60%
Air gap	0		80%
Infill pattern	Rectilinear		100%
Nozzle temperature	210 °C	Raster orientation	*XZ*
Bed temperature	60 °C		*XY*

36.4 Results and Discussion

The experimentation was mainly focussed on the evaluation of mechanical properties with the variation in raster orientation and infill density. The stress–strain curve in Fig. 36.4 indicates that PLA-printed specimen in *XY* plane exhibits the maximum tensile stress 45.09 MPa and elongation as 14.68%. In *XZ* orientation,

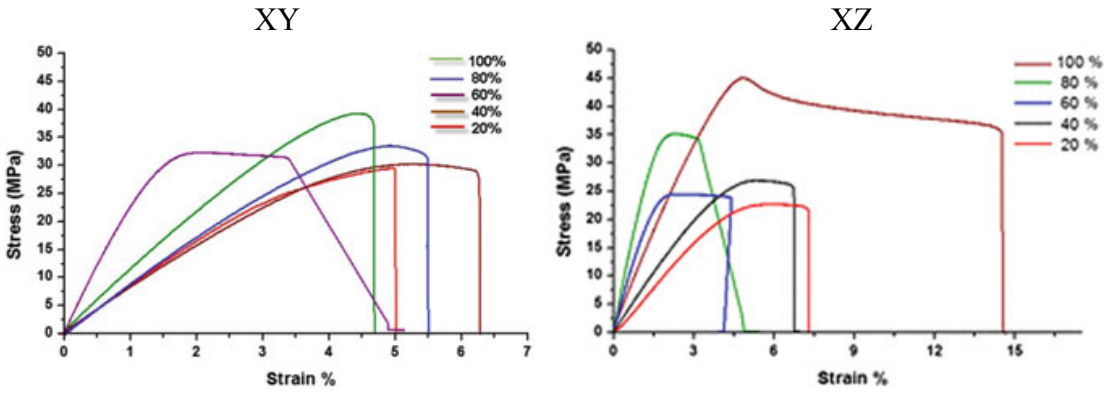


Fig. 36.4 Stress–strain curve for different infill densities in *XY* and *XZ*-raster orientations

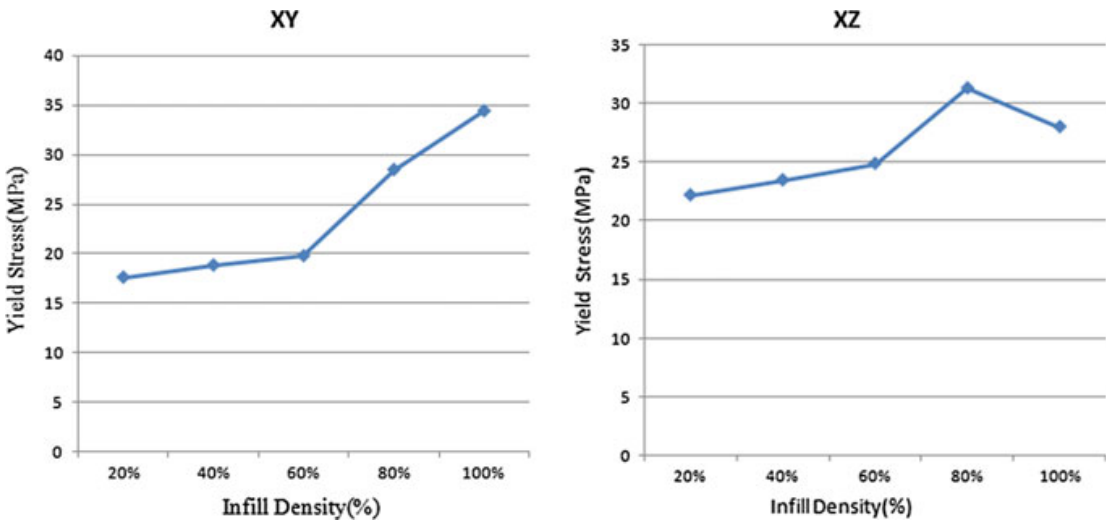


Fig. 36.5 Variation of yield stress for different infill densities and raster orientations

maximum tensile strength is 39.25 MPa for 100% infill density while maximum elongation is 6.3% for 40% infill density. The graph clearly indicates that the value of tensile strength in *XZ* for all infill densities are close to each other while in *XY* orientation the trend is completely different and the values range from 21 to 45 MPa. Figure 36.5 displays the variation of yield strength for *XY* and *XZ* orientation with different infill densities.

36.5 Fractography

In product design and manufacturing, failure analysis is an important part and, therefore, in this work the fractured surface of PLA specimens is also analysed using SEM for different raster orientations and infill densities, Fig. 36.6. It is depicted that ductile fracture is observed in *XZ* orientation as fibre is deformed in different planes leading to trans-layer tensile failure. The fracture gets transferred from one layer to another in *XZ*-direction. Specimens with *XY* orientation displayed intermediate brittle-ductile fracture behaviour. Failure is generally caused by raster pulling and rupturing as clearly depicted in Fig. 36.6. The fractured surface exhibits different fracture modes indicating hairline and major cracks marked with arrows at different location of fractured surface. In *XY* orientation, yield stress and elongation at break are more as compared to *XZ*, thereby exhibiting prolonged necking while applying tensile load as depicted in Tables 36.1 and 36.2. The strain (%) is more in *XY* orientation compared with *XZ* orientation. *XY* orientation shows brittle and ductile fracture which can be attributed to cooling rate behaviour during the solidification of material. The *XZ* orientation displays ductile fracture with dimples visible at different locations of fractured surface.

Figure 36.7 displays the printing time for different infill densities and raster orientations for PLA material. The printing time and number of layers for the tensile

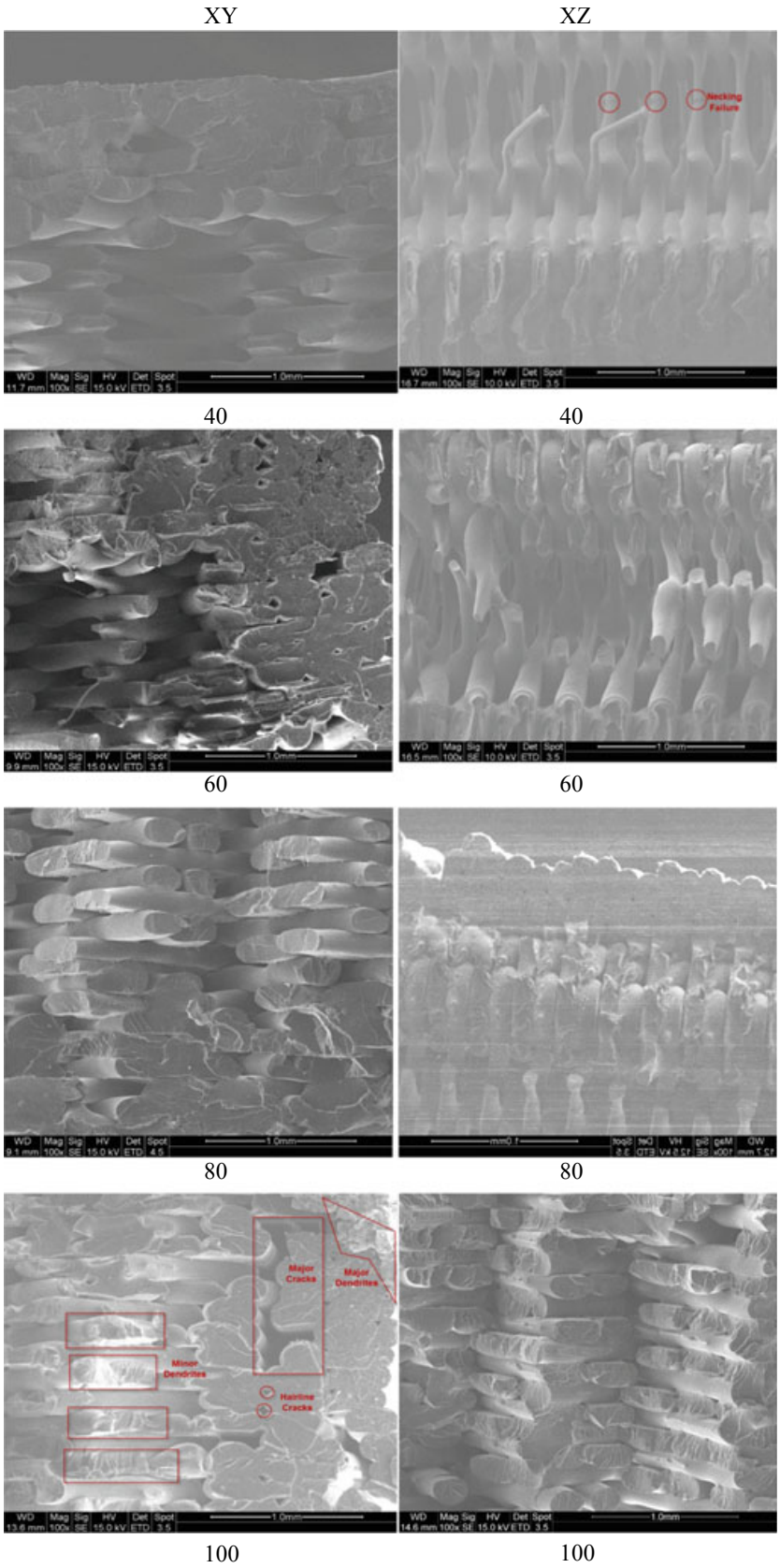


Fig. 36.6 SEM image fractured surface of *XY* and *XZ*-raster orientation and different infill density

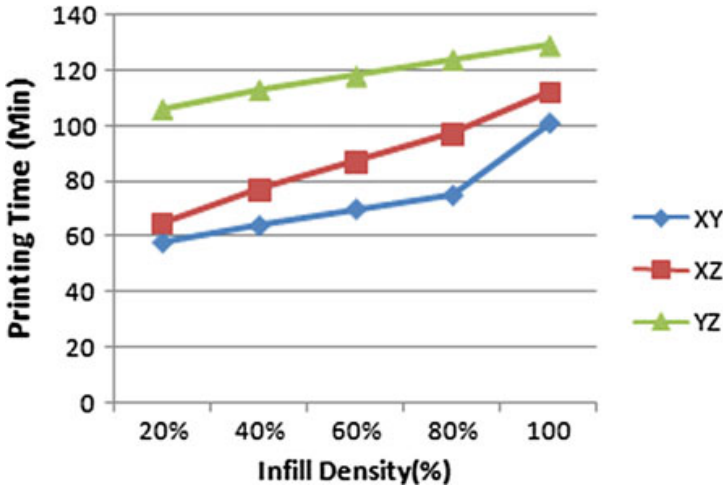


Fig. 36.7 Variation of printing time for various infill densities and raster orientations

samples depend upon the various process parameters such as raster orientation, layer thickness and feed rate. Also, the printing time is directly related to manufacturing cost [9]. Raster orientation *YZ* exhibited the longest printing time as compared to *XY* samples which have the shortest time. Printing time also decreases as feed rate is increased for *XY* orientation. In case of the *YZ* orientation, printing time remained almost constant for high feed rate values.

36.6 Conclusion

The research outlines the following conclusions:

- The maximum tensile stress in *XY* orientation is 45.09 MPa for 100% infill density with elongation as 14.68%, and similarly for the *XZ* orientation maximum tensile stress 39.256 MPa for 80% and elongation as 6.422% for 40% infill density.
- In *XY*-raster orientation, yield stress goes on increasing at different infill densities from 20 to 100% in step of 20. The maximum rate of change of yield stress is **0.958** with reference to 20% infill density. Similarly, in *XZ*-raster orientation, yield stress is increasing but at a smaller rate of change is **0.2618** with reference to 20% infill density. This clearly indicates that from yield stress point of view *XY* orientation is stronger than *XZ* orientation. This is also supported from Tables 36.1 and 36.2 that yield load is more at *XY* orientation as compared to *XZ* orientation.
- *XY* orientation has least printing time as compared to *XZ* and *YZ* orientation, with *YZ* taking maximum printing time.
- Fracture morphologies are affected by the raster orientation direction and infill density with *XY* orientation displayed intermediate brittle-ductile fracture behaviour while *XZ* displayed ductile fracture.
- Failure after tensile loading exhibits hairline and major cracks in *XY* orientation.

Acknowledgements The authors acknowledge the help from ACMS, IIT Kanpur for allowing to perform tensile testing and fractography. Support from TEQIP-III, Faculty of Engineering, DEI and MHRD NMEICT is highly appreciated.

References

1. Anderson, I.: Mechanical properties of specimens 3D printed with virgin and recycled polylactic acid. 3D Print. Addit. Manuf. **4**(2), 110–115 (2017)
2. Huang, S., Liu, P., Mokasdar, A.: Additive manufacturing and its societal impact: a literature review. Int. J. Adv. Manuf. Technol. 67:1191–1203 (2013)
3. Dawoud, M., Taha, I., Ebeid, S.: Mechanical behavior of ABS an experimental study using FDM and injection molding techniques. J. Manuf. Process. **21**, 39–45 (2016)
4. Joshi, S.C., Sheikh, A.A.: 3D printing in aerospace and its long-term sustainability. Virtual Phys. Prototyp. **10**, 175–185 (2015)
5. Lee Ventola, C.: Medical applications for 3D printing: current and projected uses. Pharm. Ther. **39**(10), 704–711 (2014)
6. Ford, S., Despeisse, M.: Additive manufacturing and sustainability: an exploratory study of the advantages and challenges. J. Clean. Prod. **137**, 1573–1587 (2016)
7. Kim, K., Lee, D.S., Kim, Y.C.: Sustainable design process based on 3D printing technology. Indian J. Sci. Technol. **9**(46), (2016)
8. Montero, M., Roundy, S., Odell, D., Ahn, S.H., Wright, P.K.: Material characterization of fused deposition modeling (FDM) ABS by designed experiments. In: Proceedings of Rapid Prototyping and Manufacturing Conference, pp. 1–21. SME, Cincinnati, Ohio, (2001)
9. Ziemian, C.W., Sharma, M.M., Ziemian, S.N.: Anisotropic mechanical properties of ABS parts fabricated by fused deposition modelling. In: Gokcek, M. (ed.) Mechanical Engineering, pp. 159–180. In Tech, New York, NY, USA (2012)
10. Bellini, A., Güçeri, S.: Mechanical characterization of parts fabricated using fused deposition modeling. J. Rapid Prototyp. **9**, 252–264 (2003)
11. Yadav, P., Kapoor S., Sahai A., Sharma, R.S.: Effect of build orientation and infill density on mechanical properties of 3D printed parts by Fused Deposition Modeling. In: 34th National Convention of Mechanical Engineers & National Seminar on. Mechanical Engineering and Emerging Challenges for Manufacturing, pp. 118–124, The Institution of Engineers India, New Delhi (2018). ISBN: 978-81-936149-8-3
12. Sharma R.S., Singhal I., Gupta S.: Innovative training framework for additive manufacturing ecosystem to accelerate adoption of three-dimensional printing technologies. 3D Print. Addit. Manuf. **5**(2), (2018)
13. Yadav, P., Singhal, I., Sahai, A., Sharma, R.S.: Intensifying hands-on learning and experimentation of fused deposition modelling three dimensional printers. In: 7th International and 28th AIMTDR Conference, Anna University, Chennai, India (2018)

Chapter 37
Development of Methods for Machine Intelligence Quotient Measurement for Piston Casting by Die Casting Machines: A Study

Rupshree Ozah and Muralidhar Manapuram

Abstract Pistons are the most important parts of an automobile engine. Pistons are manufactured by various die casting processes like sand casting, gravity die casting, pressure die casting and automatic die casting. Manufacturing assets and processes have different intelligent attributes based on embedded intelligent systems. Not much of work seems to be available for measurement and quantification of machine intelligent quotient (MIQ) for metal casting manufacturing machines. To assess the intelligence of die casting machines, three methods are developed. High MIQ indicates the lower interaction between human and die casting machine, i.e., the machine can take comparatively higher percentage of decisions in machine operations. In the present work, an attempt has been made to develop a stepwise procedure for MIQ measurement by three methods, i.e., human–machine cooperative system, geometric polytope method and multiple perspective analysis methods. MIQ for three die casting machines is computed based on human–machine cooperative method, and it has been found that automatic die casting machine has the highest MIQ (109.31) followed by pressure die casting machine (59.69) and gravity die casting machine (26.44) the lowest.

Keywords MIQ · Die casting machine · Piston casting · Human–machine cooperative system · Geometric polytope method · Multiple perspective analysis method

R. Ozah (✉) · M. Manapuram
Department of Mechanical Engineering, North Eastern Regional Institute of Science and Technology, Itanagar, Arunachal Pradesh 791109, India
e-mail: rupshreeozah91@gmail.com

M. Manapuram
e-mail: mm@nerist.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_37

37.1 Introduction

Automobile industry is the fast growing and technology-intensive industry. India is the second-largest casting producing country in the world. Globalization has led to increase the pressure on engineers and technologists to increase the efficiency of casting manufacturing processes and assets to edge-over other manufacturing processes and cost effectiveness. With the advent of computers, Internet, intelligent techniques, Industry 4.0 and cyber physical system's decision-making attributes of machines have been increased significantly. But there are no systematic methods to measure MIQ of machines. Moreover, efforts of researchers to embed human intelligent attributes by various software programs and hardware devices enable to enhance MIQ. Various tools [1] used to communicate or to transfer intelligence to the machine are fuzzy logic, genetic algorithm, vision system, charge couple device (CCD) cameras, sensors and actuators.

Intelligence is always studied in humans but has been observed in both non-human animals and plants. Intelligence in the machine is also called as artificial intelligence. Engineers and scientists are trying to impart human intelligence attributes to machines to improve machine intelligence [2, 3]. Not of much work seems to be available for measurement of MIQ. Hence in this paper, a stepwise procedure is developed to measure the machine intelligence quotient. MIQ for three types of die casting machines (gravity die casting machine, pressure die casting machine and automatic die casting machine) for the manufacture of piston casting is also considered by human–machine cooperative system. In addition, two more methods for measuring MIQ have been considered and discussed in brief.

37.2 Machine Intelligent Quotient and Its Sub-systems

Intelligence is always a hard topic to express and there is no clear definition for it. It may be defined as the ability to respond rapidly and successfully to a new situation. Intelligence requires the ability to sense the environment, make decisions and control action [4]. Various attributes of intelligence are adaptability, self-maintenance, communication, autonomy, learning, self-improvement, anticipation, goal seeking, inference and creativity. Intelligent machine is a machine that learns and can solve problems by reasoning and inference. Intelligent controller is any control system that can deliver a degree of human-like response to stimuli. Nowadays, intelligent factories have been evolving with intelligent machine systems along with intelligent human–machine interfaces.

Intelligent systems can be broadly classified into four types viz. biological intelligence, artificial intelligence, machine intelligence and computational intelligence. Intelligent quotient is developed in human beings in the form of intelligence quotient (IQ) which can be defined as the ratio of mental age and chronological age multiplied by hundred.

Machine intelligence is another term used to measure the intelligence of the machine, i.e., theory and development of the computer systems facilitating to perform the tasks normally which requires human intelligence.

Machine intelligence is the process of analyzing, organizing and converting data into knowledge, where (machine) knowledge is defined as the structured information acquired and applied to remove ignorance and uncertainty about a specific task pertaining to the intelligent machine [5]. Machine intelligence consists of two components, i.e., control intelligence and interface intelligence. Interface intelligence indicates the degree of intelligence of human–computer interaction that may be human to machine interaction and vice versa. Machine intelligence quotient can be considered as a union of machine control intelligence and machine interface intelligence [6]. High machine intelligent quotient indicates the number of decisions taken by machines is more than human, i.e., machine to human interaction is more than human to machine interaction. Nowadays, machines are taking complex decisions without any interaction of human in fully automatic process. The taxonomy of intelligence and its meaning and the paradigms are presented in Table 37.1.

Intelligence attributes consist of activity attributes (cognition, memory, logical thinking, creative thinking and evaluation), material attributes (images, symbols, non-verbal perceptions), product attributes (classes, relations, systems, unitary items, transformations and implications) and their combinations. There is a continuous effort by the engineers and the technologies to impart human intelligent attributes to machines and industrial shop-floor systems [7, 8].

The human intelligence has been gradually increasing linearly with time, but the machine intelligence has not been significant during 1950–2000 [9]. In Fig. 37.1, human intelligence and machine intelligence with time are shown. With the advent of computer, artificial intelligence, sensors, actuators, neural networks, the intellectual power of machines has been steeply increased. The point where the machine

Table 37.1 Taxonomy of intelligence and its paradigms

Types of intelligence	Literal meaning	Paradigms of intelligence
Biological intelligence	Naturally grown organism, human knowledge plus sensory input	Human and other developed species brain
Artificial intelligence	Cognitive inspired artificial models, knowledge tidbits and sensor data	Intelligent systems, knowledge systems, decision-making systems
Machine intelligence	Complex machines, PLC controlled machines	Robot, neural network
Computational intelligence	Computational methodologies and software systems, mathematical computation plus sensor data	Expert system, Fuzzy system, robotic nose, robotic arm, intelligent agent-based systems, Internet of things, cloud manufacturing, genetic algorithm

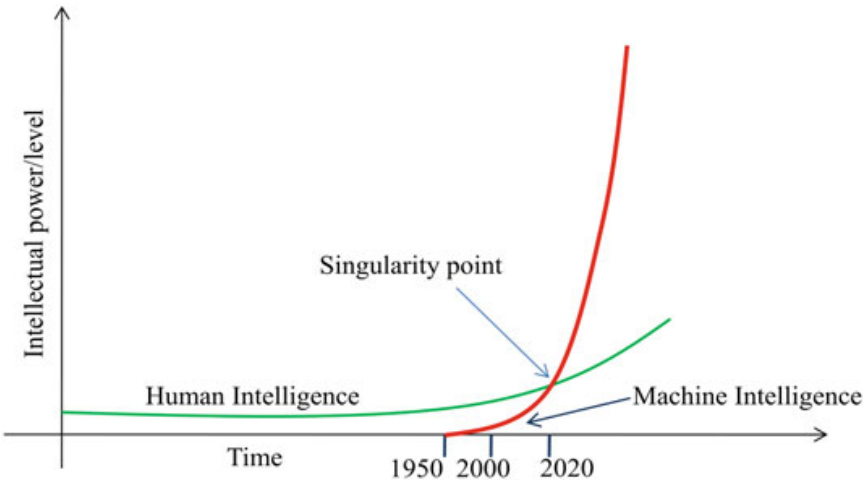


Fig. 37.1 Evolution of human and machine intelligence

intelligence outperforms the human intelligence is termed as singularity. In this century, machine intelligence is likely to supersede human intelligence.

37.3 Die Casting Machines for Piston Manufacturing

Pistons are the short piece of metal that moves up and down under pressure inside a cylinder of an engine and are made gas tight by piston rings. Pistons are usually made of Al–Si alloy and manufactured by different types of die casting machines viz. gravity, pressure and automatic based on the quantity, quality and lead times. The number of automobiles manufactured [10, 11] and the number of pistons required are quantified during 2008–2016 and presented in Table 37.2. The numbers of automobiles and pistons produced during a decade period are 157.1 million and 238.2 million, respectively. The annual requirement of pistons for automobile industry is around 60.12 million.

Al–Si pistons are generally manufactured by die casting processes, i.e., gravity die casting, pressure die casting and automatic die casting process. In die casting machine, molten metal enters the mold cavity under gravity [12]. Pressure die casting is characterized by forcing molten metal under high pressure into a mold cavity. Gravity die casting (GDC) is preferred in medium to large volumes and of a heavier section thickness than high pressure die casting, but thinner sections than sand casting. It has very good dimensional accuracy and also good cast surface finish. Pressure die casting cycle time is less, reliable and cost-effective manufacturing process for the production of high volumes and cast components that are near net-shaped having tight tolerances. In automatic die casting, more operations castings are performed by machines like automatic die coating, automatic pouring and ejection without any human interaction. The tasks of die casting machine,

Table 37.2 No. of automobiles and pistons produced in India during 2008–17

Type of automobile	Existing automobiles plying (2008–16) (Million) (1)	No. of pistons/ automobile (2)	Total no. of pistons for existing automobiles (3 = 1 * 2)	10% annual spare parts for existing automobiles (Million) (4 = 3 * 10%)	2017 annual automobile production (Million) (5)	No. of piston for 2017 (Million) (6)	Total no. of pistons required annually (Million) (7 = 4 + 6)
Two wheelers	125.6	1	125.6	12.56	19.9	19.9	32.46
Three wheelers	6.7	2	13.4	1.34	0.8	1.6	2.94
Four wheelers	24.8	4	99.2	9.92	3.7	14.8	24.72
Total	157.1		238.2	23.82	24.4	36.3	60.12

Table 37.3 Die casting machine/human task chart

Task performed by			
Die casting machine	Gravity die casting machine	Pressure die casting machine	Automatic die casting machine
Machine on/off	Human	Human	Human
Die lubrication	Human	Human	Machine
Core insertion	Human	Machine	Machine
Closing mold	Machine	Machine	Machine
Pouring the metal	Human	Human	Machine
Opening the mold	Machine	Machine	Machine
Ejection of the casting	Human	Human	Machine
Extraction of the casting	Human	Human	Machine

sequence of operations, task performed by human or machine for gravity, pressure die and automatic die casting machines are presented in Table 37.3.

37.4 Methods for MIQ Measurement for Die Casting Machines

In the present research, three methods for computation of MIQ of die casting machines have been considered. These include machine–human cooperative system, geometric polytope method and multiple perspective analysis method.

37.4.1 Method for MIQ Measurement for Die Casting Machines by Human–Machine Cooperative System

This system basically estimates the mental workload and data quantity and their interface complexity. Stepwise procedure to estimate MIQ is as follows:

Step 1: To find out the input and output variables of the target machine, i.e., the die casting machines and decide the task performed by either machine or human.
Step 2: To define the die casting intelligent task set,

$$T = \{T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8\} \tag{37.1}$$

where $T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8$ are the intelligent tasks of die casting machine performed by machine and human. These are the set of tasks required to control and operate the event.

Step 3: To determine the intelligent task cost,

$$t = \{t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8\} \tag{37.2}$$

where $t_1, t_2, t_3, t_4, t_5, t_6, t_7, t_8$ are the task intelligent costs. Task intelligent costs t_i are intelligence required to execute the task T_j which is varying from 6 to 22 for the tasks 1–8.

Step 4: To prepare data transfer matrix $D = d_{ij}$, for transferring data from task T_i to T_j.

$$D = d_{ij} = \begin{bmatrix} d11 & d12 & d13 & d14 & d15 & d16 & d17 & d18 \\ d21 & d22 & d23 & d24 & d25 & d26 & d27 & d28 \\ d31 & d32 & d33 & d34 & d35 & d36 & d37 & d38 \\ d41 & d42 & d43 & d44 & d45 & d46 & d47 & d48 \\ d51 & d52 & d53 & d54 & d55 & d56 & d57 & d58 \\ d61 & d62 & d63 & d64 & d65 & d66 & d67 & d68 \\ d71 & d72 & d73 & d74 & d75 & d76 & d77 & d78 \\ d81 & d82 & d83 & d84 & d85 & d86 & d87 & d88 \end{bmatrix} \tag{37.3}$$

Step 5: To prepare the task allocation matrix for gravity die casting machine, pressure die casting and automatic die casting machines. The element of this type of matrix can only have binary numbers, i.e., 0 or 1. If the machine performs a task T_i then a_{i1} is 1, and if human performs the task a_{i2} will be 1. If the task T_i is neither assigned by the machine nor human, then it will be 0. Hence, $a_{i1} + a_{i2} + a_{i3} = 1$ for all $1 \leq i \leq n$ task allocation matrix will be $3 \times n$.

$$A = \begin{pmatrix} a_{i1} & a_{j1} & 0 \\ \vdots & \ddots & \vdots \\ a_{i8} & a_{j8} & 0 \end{pmatrix} \tag{37.4}$$

Step 6: To prepare the task allocation matrix considering the task performed by machine or human. To prepare task allocation table considering the task performed by either human or machine, let us consider the machine–human interface (C_{mh}) and human–machine interface complexity (C_{hm}). The values can be from 0.01 to 0.09 depending on the human–machine interaction. Higher values are assigned where more human interaction is required and vice versa. The interface complexities are to be considered for three die casting machines.

Step 7: To prepare the task allocation graph by considering the set of tasks and their interactions like machine to machine, human to machine and machine to human.

Step 8: To compute the machine control intelligent quotient C, i.e., the sum of all intelligent task costs [10].

$$C = \sum_{i=1}^{n} a_{i1} t_i + \sum_{i=1}^{n} a_{i2} t_i \tag{37.5}$$

where C is the control intelligence quotient (CIQ) of the overall control system containing the human and machine. C is the sum of all task intelligence costs (both machine and human).

$$C = t_1 + t_2 + t_3 + t_4 + t_5 + t_6 + t_7 + t_8 \tag{37.6}$$

Step 9: To compute the human intelligence quotient (H). It is the sum of task intelligence costs allocated to the human and interface intelligence costs. We can obtain MIQ by subtracting H from C.

$$H = \sum_{i=1}^{n} a_{i2} t_i + C_{\text{mh}} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i1} a_{j2} d_{ij} + C_{\text{hm}} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i2} a_{j1} d_{ij} \tag{37.7}$$

Step 10: MIQ is divided into two components, i.e., machine control intelligence and interface intelligence.

$$M = M_{\text{i}} + M_{\text{c}} \tag{37.8}$$

$$M_{\text{i}} = -\left\{ C_{\text{mh}} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i1} a_{j2} d_{ij} + C_{\text{hm}} \sum_{i=1}^{n} \sum_{j=1}^{n} a_{i2} a_{j1} d_{ij} \right\} \tag{37.9}$$

$$M_{\text{c}} = \sum_{i=1}^{n} a_{i1} t_i \tag{37.10}$$

M_{i} values are either negative or zero. It implies that if M_{i} is zero, then it will not have any effect on MIQ. If M_{i} is negative, it means the interface overhead deteriorates the machine intelligence quotient.
Step 11: To compute MIQ, subtract H from C;

$$M = C - H \tag{37.11}$$

37.4.2 Estimation of MIQ by Geometric Polytope Method

This method is used to estimate a relative means of determining how to contrast across different machines for comparative intelligence or autonomy. A stepwise procedure for MIQ measurement or estimation is as follows:

Step 1: To define the axes of the geometric polytope to be relevant to that performance of the attribute (goal) under consideration.
Step 2: To define the scales of each axis of the geometric polytope to be relevant to the attribute or goal of interest.
Step 3: To plot the attribute values (productivity, cost, cycle time, etc.) for different machines (die casting) on the same axis.
Step 4: To consider polytope area resulting will provide a relative comparison of the efficiency of a particular machine to perform certain attributes.
Step 5: To use the relative measures of machine intelligence, the effectiveness to perform specific attributes or goal under various conditions can be determined. It is difficult to get an absolute value of MIQ but the comparison is made based on the area of polytope. Higher the area of polytope, then higher is the MIQ of machine.

37.4.3 Estimation of MIQ by Multiple Perspective Analysis

In this method, three filters are used, i.e., technical filter (T), organizational filter (O), personal filter (P). Technical filter includes

Step 1: To define linguistic variables to take values on the interval between zero and one for measuring machine intelligence.
Step 2: To consider three filters, i.e., technical, organizational and personal and their attributes.
Step 3: To normalize the micro and macroeconomic measures (attributes) like productivity, cost and cycle time, etc., for three die casting machines.
Step 4: To correlate the fuzzified values {T-O-P} to MIQ.

For example, three die machines have {0.4, 0.3, 0.2} for gravity die casting, {0.5, 0.4, 0.3} for pressure die casting and {0.7, 0.6, 0.5} for automatic die casting. Then, automatic die casting machine has got the higher MIQ of {0.7, 0.6, 0.5} on linguistics scale which reflects 0.7 for technical, 0.6 for organizational and 0.5 for personal filter.

37.5 Computation of MIQ by Machine–Human Cooperative System for Die Casting Machines

Computation of MIQ for three die casting machines is given below considering intelligent tasks and costs.

Step 1: To find out the input–output variables of the die casting machines and decide the performance performed by either machine or human.
Step 2: To define die casting machine intelligent task set.

$$T = \{T_1, T_2, T_3, T_4, T_5, T_6, T_7, T_8\}$$

Step 3: To determine intelligent task cost,

$$t_i = \{6, 10, 15, 19, 22, 18, 16, 14\}$$

Step 4: To prepare data transfer matrix,

$$\begin{aligned}
D &= d_{ij} = d_{8\times 8} \\
D_1 &= \{d_{14}, d_{16}, d_{24}, d_{26}, d_{34}, d_{36}, d_{45}, d_{46}, d_{47}, d_{56}, d_{67}, d_{68}\} \\
&= \{25, 25, 15, 17, 19, 15, 30, 12, 20, 20, 20, 22\} \\
D_2 &= \{d_{13}, d_{14}, d_{16}, d_{17}, d_{23}, d_{24}, d_{26}, d_{27}, d_{34}, d_{35}, d_{38}, d_{45}, d_{48}, d_{56}, d_{57}, d_{67}, d_{68}, d_{78}\} \\
&= \{20, 22, 12, 14, 15, 24, 17, 15, 12, 30, 12, 32, 14, 28, 24, 15, 19, 22\} \\
D_3 &= \{d_{12}, d_{13}, d_{14}, d_{15}, d_{17}, d_{23}, d_{28}, d_{34}, d_{38}, d_{45}, d_{48}, d_{56}, d_{58}, d_{67}, d_{68}, d_{78}\} \\
&= \{14, 18, 28, 30, 26, 19, 12, 20, 16, 12, 24, 22, 22, 30, 20, 28, 26\}
\end{aligned}$$

Step 5: Task allocation matrix

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Gravity die casting machine

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

Pressure die casting machine

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Automatic die casting machine

Step 6: To prepare task allocation table considering the task performed by either human or machine, let us consider the machine–human interface (C_{mh}) and human–machine interface complexity (C_{hm}). The values considered are from 0.02 to 0.05.
Step 7: Task allocation shows the task performed by human, machine and events of three die casting machines. Tasks allocations are shown in task allocation matrices (step 5).
Step 8: To determine the control intelligence quotient for three die casting machines,

$$C_1 = 120,\ C_2 = 120, C_3 = 120$$

(Eq. 37.6)
Step 9: To compute the human intelligent quotient (Eq. 37.7) for three die casting machines,

$$H_1 = 83 + (0.04 \times 94) + (0.05 \times 136) = 93.56;$$
$$H_2 = 52 + (0.03 \times 129) + (0.04 \times 111) = 60.31;$$
$$H_3 = 6 + (0.025 \times 138) + (0.02 \times 62) = 10.69$$

Step 10: To compute die casting machine control intelligence (Eq. 37.10) and interface intelligence (Eq. 37.9),

$$M_{c1} = 37,\ M_{c2} = 68,\ M_{c3} = 114;$$
$$M_{i1} = -10.56,\ M_{i2} = -8.31,\ M_{i3} = -4.69$$

Step 11: To compute MIQ (using Eq. 37.11),

$$M_1 = C_1 - H_1 = 120 - 93.56 = 26.44;$$
$$M_2 = C_2 - H_2 = 120 - 60.31 = 59.69;$$
$$M_3 = C_3 - H_3 = 120 - 10.69 = 109.31$$

The total intelligent costs have been considered 120 for all eight tasks for die casting machines and considered the task costs from 6 to 22 according to the task complexity, and it is shown in Table 37.4. The interface intelligence for all of the casting processes is negative due to the fact that the interface overhead deteriorates the MIQ. The control intelligence stayed top for the automatic die casting machine as this machine has controlled the more intelligence task than other two methods. The computed MIQ values are presented in Table 37.5.

Table 37.4 Task description, tasks intelligent costs

Task element	Task description	Intelligent task description	Task intelligent cost	Task intelligent cost
T_1	Machine on/off	Detect	t_1	6
T_2	Die lubrication	Identify	t_2	10
T_3	Core insertion	Observe	t_3	15
T_4	Closing mold	Interpret	t_4	19
T_5	Pouring the metal	Define tasks	t_5	22
T_6	Opening the mold	Select actions	t_6	18
T_7	Ejection of the casting	Execute	t_7	16
T_8	Extraction of the casting	Evaluate	t_8	14
			Total	120

Table 37.5 Estimation of die casting machines intelligence quotient

Description	Gravity die casting machine	Pressure die casting machine	Automatic die casting machine
Total intelligent cost ($C = H + M$)	120	120	120
Human intelligence quotient (H)	93.56	60.31	10.69
Control intelligence (M_c)	37	68	114
Interface intelligence (M_i)	−10.56	−8.31	−4.69
Machine intelligence quotient (M)	26.44	59.69	109.31

37.6 Conclusions

In the present paper, MIQ and its sub-systems are defined and formulated for computation of machine intelligence. Three methods for measuring MIQ were formulated, and MIQ was calculated by human–machine cooperative process for gravity die casting, pressure die casting and automatic die casting machines. Human–machine cooperative system is more suitable for estimation of die casting machines for the manufacture of pistons. MIQ of automatic die casting machine has the highest and gravity die casting machine lowest. It is also useful for comparison of different machines with different intelligent attributes manufactured by original equipment manufacturers. From the MIQ computation, it was found that automatic die casting machine has the highest MIQ of 109.31 followed by pressure die casting machine (56.69) and gravity die casting machine (26.44).

Acknowledgements The authors gratefully acknowledge the support extended by North Eastern Regional Institute of Science and Technology (NERIST), Itanagar, Arunachal Pradesh-791109, India in the present work.

References

1. Bien, Z., Bang, W.C., Kim, D.Y., Han, J.S.: Machine intelligence quotient: its measurements and applications. Fuzzy Sets Syst. **127**, 3–16 (2002)
2. Muralidhar, M.: Introduction to Methods for Calculation of Machine Intelligent Quotient, Short Term Training Program on Soft Computing in Engineering Application, pp. 87–103, NERIST, Arunachal Pradesh, India (2010)
3. Muralidhar, M.: Development of intelligent induction furnace and measurement of furnace intelligence quotient: a study. In: 69th World Foundry Congress, pp. 0754–0761, Hangzou, PR China (2010)

4. Anthony, A., Jannett, T.C.: Measuring machine intelligence of an agent-based distributed sensor network system. In: Advances and Innovations in Systems, Computing Sciences and Software Engineering, pp 531–535 (2007)
5. Wang, Y.: On Abstract Intelligence: Toward a Unifying Theory of Natural, Artificial, Machinable, and Computational Intelligence, Int. J. Softw. Sci. Comput. Intell. 1(1), 1–17 (2009)
6. Muralidhar, M.: Mathematical computational methods for determination of MIQ for machines. In: International Conference on Mathematics and Applied Mechanics, Bangalore (2007)
7. Mourtzis, D., Milas, N., Vlachou, A.: An internet of things-based monitoring system for shop-floor control. J. Comput. Inf. Sci. Eng. **18**, 0210051-8, (2018)
8. Park, H.J., Kim, B.K., Lim, K.Y.: Measuring the machine intelligence quotient (MIQ) of Human–machine cooperative systems. In: IEEE Trans Syst, Man, and Cybern—Part A: Syst Hum **31**(2), 89–96 (2001)
9. The Machine Intelligence Continuum: https://www.topbots.com/topbots-ai-machine-intelligence-continuum. Last accessed on 2018/11/02
10. Kumar, S., Muralidhar, M.: CTS Estimation in Shock Absorber Assembly Workforce in Indian Automobile Industry: A Study. HWWE, AMU, Aligarh, India (2017)
11. Reddy, A.C., Rajanna, Ch.: Design of gravity die casting process parameters of Al-Si-Mg alloys. J. Mach. Form. Technol. 1–24 (2009)
12. Ashwini, A., Avinash, K.S.: Rejection analysis in Piston manufacturing unit. Int. J. Innov. Res. Sci., Eng. Technol. **4**(3), 1157–1163 (2015)

Chapter 38
Optimization of Porosity in Cast Aluminium Foams

R. Kumar, G. Avinash and Neha Gupta

Abstract Closed-cell ultra-light aluminium foams are often used as structural materials in aerospace, railways and automobile sectors as they provide good strength to weight ratio with high impact strength and corrosion resistance. Melt - based metal foams can be casted using indirect foaming techniques, wherein gas bubbles are created inside the melt which has been pre-treated in a suitable way. The properties of these metal foams are directly affected by nature of porosity, which is difficult to measure and control experimentally. Therefore, theoretical solutions have to be developed to control this porous structure. In the present work, a aluminium foam, prepared using foaming method using blowing agents with porosity $\sim$86%, is used, where the effect of melt viscosity and solidification time on gas bubble size and its rising velocity inside melt have been theoretically studied. The calculations showed that rising velocity of hydrogen gas bubbles, aluminium melt viscosity and solidification time are responsible for motion of bubbles in the aluminium melt. Control of this movement of bubbles can result in optimum porosity in the solidified foam. A relationship has been established for viscosity enhancement due to the addition of additives like Ca, Al_2O_3, SiC, TiB_2, SiO_2, BN and their influence on bubble size and its rising velocity during solidification of this foam. Overall, the study suggested amount of additives that are required to be added in a given volume of aluminium melt, to increase melt viscosity and to maximize bubble entrapment. The results also showed that a cooling time of $\sim$8 s leads to efficient bubble entrapment with uniform pore structure and hence properties of aluminium foam.

Keywords Aluminium foam · Porosity control · Blowing agents · Solidification

R. Kumar · G. Avinash · N. Gupta (✉)
Department of Metallurgical and Materials Engineering,
NIT Raipur, Raipur (C.G.) 492010, India
e-mail: ngupta.met@nitrr.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_38

38.1 Introduction

In recent years, light-weight metal foams have gained a lot of attraction both for the scientific research and industrial applications. Amongst metal foams, aluminium alloys are commercially the most exploited ones due to their high strength to density ratio, good corrosion resistance and competitive cost [1]. The application potential of these materials has already being explored for different fields, namely construction, energy absorption, thermal or sound insulation, etc. where improvement in safety is needed. Several processing methods can be used to produce these structures as casting, powder metallurgy, metal deposition, additive manufacturing and many more [2]. Closed-cell metallic foams can be economically obtained by solidifying melt containing well-dispersed gas bubbles, also known as indirect gas foaming technique. Such foams are commonly used for structural applications as they exhibit properties as enhanced energy absorption with high stiffness to weight ratio [3], thereby offering significant performance gains.

Most of the properties, such as strength, sound insulation, impact energy absorption, etc. of these foams are extremely sensitive and are directly affected by its final pore size and its distribution [1–3]. Hence, a well-controlled closed-cell pore structure is often desired in order to obtain high-quality metal foams, so as to meet the requirements of specific applications. Therefore, it is very much required to have good pore distribution and its control during processing. However, it is a tedious task to experimentally control porosity using foaming techniques due to the complex nature of foaming process which often leads to non-homogeneous distribution of gas bubbles. Pore size, shape and its distribution depend on melt viscosity, bubble size and rising velocity, solidification time, surface tension, fluid buoyancy and drag velocity. Therefore, it is extremely important to control these parameters during foaming process. Only few studies have been reported on theoretical optimization studies related to pores formation in metal foams [4–8]. It is known that low melt viscosity and high bubble rise velocity tends to drain the melt faster leading to rapid escape of gas bubbles before foaming [5]. This leads to fabrication of imperfect foam having lower porosity. Therefore, melt with sufficient viscosity and optimal bubble rise velocity has to be prepared to efficiently stabilize the liquid foam and fabricate structurally sound aluminium foams [6, 7]. Talaria discussed and predicted terminal bubble velocity in a given liquid [5]. His viscosity expression can be easily used to relate with a given bubble and it's rising velocity. Studies by Banhart showed the importance of studies, to understand stabilization mechanisms for metal foams [6]. Akiyama et al. have prepared closed-cell aluminium foams through indirect gas foaming technique which used titanium hydride (TiH_2) as foaming agent and calcium as additive to enhance the viscosity of the melt so as to adjust its foaming properties [7].

In this work, a closed-cell aluminium foam was prepared using similar technique as by Akiyama et al. by adding TiH_2 as foaming agent to create hydrogen bubbles [7]. These bubbles resulted in inhomogeneous pore development in aluminium foam upon solidification. Here, the effect of melt viscosity and cooling conditions

on pore formation in aluminium foam has been investigated theoretically. Further, the effect of different additives (Ca, Al_2O_3, TiB_2, SiO_2, SiC and BN) on melt viscosity and its influence on bubble diameter and hence its rising velocity during solidification of this foam was analysed. A relationship has been developed for melt viscosity, solidification time, bubble size and bubble rising velocity for efficient pore entrapment and thereby uniform distribution in aluminium foam.

38.2 Materials and Methods

Aluminium alloy was heated at 720 °C in a Heat Treat Electric furnace for 1.5 h, to which 1.8 wt% additive (calcium) was added as thickening agent under continuous stirring to adjust the melt viscosity. Thereafter, 1.5 wt% TiH_2 was added as blowing agent to this melt to form hydrogen gas bubbles. Continuous stirring (total time = 150 s, speed = 1000 rpm) of the melt resulted in a homogeneous distribution of additive, and blowing agent inside melt till cooling helped to obtain a foam with porosity of $\sim$86%.

Under these well-controlled experimental conditions, a porous structure with uniform pore size and spherical shape was expected. However, a non-uniform porous structure was observed as illustrated in Fig. 38.1, despite taking precautions in the process, which resulted in decrement of foam properties (strength, stiffness, etc.) when compared to theoretical values, as was also observed by Suresh et al. [9]. Efficient control of porosity in the foam is a solution to this. Also, it is known that bubble rising velocity and melt viscosity in the aluminium melt affects the porosity in foam [1–3]. For study, a mould with dimensions 15 × 15 × 15 cm^3 with 20 cm sprue length and 2 cm sprue diameter was used. Here, an effort has been done to solve and study for the in situ pore size and its distribution with respect to melt viscosity during solidification using theoretical calculations.

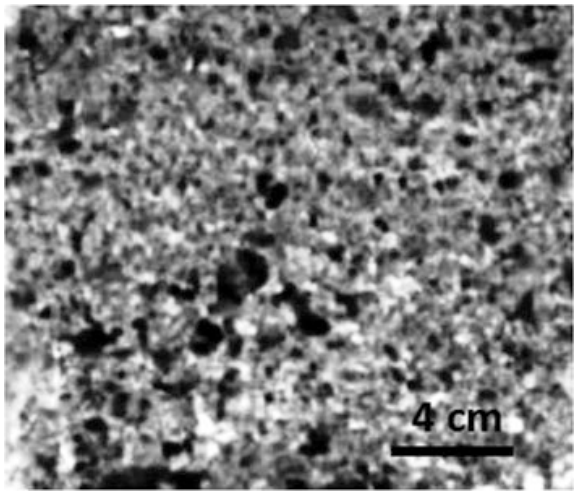


Fig. 38.1 Microstructure of foam showing non-uniform porosity

38.3 Results and Discussion

In continuation of our earlier works [4], we know that pore size can be controlled by controlling the diameter of hydrogen bubble generated inside the aluminium melt. Also, pore size distribution of foam can be controlled by controlling the rising velocity of these hydrogen bubbles. Here, a relationship between bubble size with pore size and bubble rising velocity with pore distribution inside the aluminium foam during solidification have been developed. Using Stokes equation [4, 10], hydrogen gas bubble rising velocity can be calculated, which indicates the movement of bubbles inside aluminium melt leading to pore distribution inside foam upon solidification. It was calculated that the small isolated gas bubbles rose in melt at 720 °C with rising velocity of 0.038 m/s. For this rising velocity, the distance covered by bubbles at solidification start (~615 °C, 30 s) [11] is 1.1241 m. This bubble displacement was too high compared to the height of the mould used; therefore, the gas bubbles get easily escaped resulting in undesired low porosity foam structure. The bubble rising velocity has to be therefore decreased by increasing the melt viscosity so that gas bubbles stay in the melt for longer time, thus resulting in a structurally sound porous structure. The same was confirmed by Akiyama et al., who used calcium as thickening agent to increase the melt viscosity [7].

This velocity is strongly dependant on the aluminium melt viscosity. The relative viscosity (μ_{rel}) due to the addition of additives is [4, 11]:

$$\mu_{rel} = \frac{\mu_{sus}}{\mu_{liq}} \tag{38.1}$$

μ_{sus} Dynamic viscosity of the suspension
μ_{liq} Dynamic viscosity of the aluminium melt (1.38 MPa) [12].

Using Eq. (38.1), it was calculated that after the addition of 1.8 wt% additive, the viscosity of aluminium melt increased from 1.38 to 7.4 MPa and was high enough to hold gas bubbles inside the melt for longer times. This increase in viscosity decreased the bubble rise velocity from 0.038 to 0.007 m/s, which helped in efficient entrapment of gas bubbles in the melt to achieve desired pore structure.

Table 38.1 Volume percentage of different additives that can be used

Additives	Volume %
Ca	2.79
Al_2O_3	7.11
TiB_2	8.14
SiO_2	4.77
SiC	5.78
BN	3.16

Different additives can be used to obtain such increment in melt viscosity. Table 38.1 reports the calculated volume percentages of some of these additives.

However, when TiH_2 was added to the melt, it decomposed and resulted in formation of gas bubbles. These bubbles further expanded as they rose in the melt. The increase in bubble diameter w.r.t. solidification time can be given as [13]:

$$d(t) = 2\left[\frac{K}{\sqrt{D} \pm \sqrt{\frac{K^2}{D} + 2(0.9253K)}}\right] t^{\frac{1}{2}} \tag{38.2}$$

where,

$d(t)$	Bubble diameter (mm)
$K = \sqrt{\frac{3}{\pi}} \times \frac{D(C-C_w)}{\rho_g}$	1.22×10^{-6} m^2/s
C	Gas concentration in the liquid (0.14×10^{-3} kg/m^3)
C_w	Gas concentration within the pore (0.055×10^{-3} kg/m^3)
ρ_g	Bubble density (0.17×10^{-3} kg/m^3)
D	Diffusion coefficient of hydrogen in the Al-melt (2.5×10^{-6} m^2/s)
t	Solidification time (s).

Equation (38.2) showed that gas expansion leads to bubble growth, which further influences its rising velocity. These changes were calculated using Stokes equation and Eq. (38.2) and are reported in Table 38.2. Table 38.2 shows that with increase in solidification time, bubble growth occurs, which also enhances their rising velocities due to expansion. This increase in bubble growth during solidification can lead to increased pore size.

Overall, the study concludes that for a foam size of 15 cm × 15 cm × 15 cm, to have hydrogen gas bubble being effectively entrapped inside aluminium melt, a reduction in solidification time has to be done. As per calculations, the solidification time should be less than 8 s so as to get uniform porosity in foam. Therefore, for

Table 38.2 Influence of solidification time and bubble size on bubble rising velocity

t (sec)	d (mm)	U_∞ (mm/s)
1	0.74	2.33
2	1.05	4.71
3	1.29	7.05
4	1.49	9.40
5	1.67	11.72
6	1.83	14.11
7	1.97	16.46
8	2.11	18.8
9	2.24	21.15
10	2.36	23.52

parameters chosen, along with improvement in melt viscosity by adding 1.8 wt% additives, fast cooling of the melt are required so that an optimal uniform pore size and distribution can be achieved.

38.4 Conclusion

This work was performed to have a better understanding and control over pore size and its distribution in aluminium foam to have better property control. It was observed that the melt viscosity increased on the addition of 1.8 wt% additives like Ca, Al_2O_3, TiB_2, SiO_2, SiC and BN, which helped in holding bubbles in the melt. The viscosity enhanced from 1.38 to 7.4 MPa. This increased melt viscosity helped in entrapping the bubbles in the melt by decreasing bubble rise velocity from 0.038 to 0.007 m/s.

It was also observed that with increment in solidification time, bubble growth occurs due to gas expansion, which affects their rising velocity. This leads to increase in pore size, which is sometimes undesirable. Calculations suggested that reducing solidification time from 30 s to less than 8 s, improved the entrapment of gas bubbles inside the aluminium melt, thereby creating effective porosity in the foam.

Overall, using calculations discussed here, it is expected that a researcher can optimize his/her parameters used to fabricate foams as mould size, melt used, melt viscosity, amount of additives and solidification time to obtain uniform porosity in the metal foam and hence it's properties. This study can help in better designing of the experiments used for development of metal foams by indirect foaming method.

References

1. Banhart, J.: Europhys News (30), 17–20 (1999)
2. Banhart, J.: Manufacture, characterization and application of cellular metals and metal foams Prog. Mater. Sci. **46**, 559–632 (2001)
3. Gibson, L.J., Ashby, M.F.: Cellular Solids: Structure and Properties, Cambridge. Cambridge University Press, UK (1997)
4. Avinash, G., Harika, V., Sandeepika, Ch., Gupta, N.: Pore size control in aluminium foam by standardizing bubble rise velocity and melt viscosity. IOP Conf. Ser.: Mater. Sci. Eng. **338**, 012010 (2018)
5. Talaia, M.A.R.: Terminal velocity of a bubble rise in a liquid column. World Acad. Sci. Eng. Technol. (28), 264–68 (2007)
6. Banhart, J.: Metal foam: production and stability. Adv. Eng. Mater. **8**(9), 781–794 (2006)
7. Akiyama, S., Imagawa, K., Kitahara, A., Nagata, S., Morimoto, K., Nishikawa, T., Itoh, M.: Foamed metal and method of producing same. US Patent 4713277 A (1987)
8. Suresh, C., Vidyashankar, S.: Investigation of the effect of pore size on mechanical properties of open cell aluminium foam structure. International J. Adv. Eng. Manag. Sci. **2**(9), 1641–1644 (2016)

9. Stokes, G.G.: Mathematical and Physical Papers, Sommerfeld. Cambridge University Press, London (1880)
10. Bala, K.C., Khan, R.H.: Rate of solidification of aluminium casting in varying wall thickness of cylindrical metallic moulds. Leonardo J. Sci. (25), 19–30 (2014)
11. Konijn, B.J., Sanderink, O.B.J., Kruyt, N.P.: Experimental study of the viscosity of suspensions: effect of solid fraction, particle size and suspending liquid. Powder Technol. (266), 61–69 (2014)
12. Assael, M.J., Kakosimos, K., Banish, R.M., Brillo, J., Egry, I., Brooks, R., Quested, P.N., Mills, K.C., Nagashima, A., Sato, Y., Wakeham, W.A.: Reference data for the density and viscosity of liquid aluminum and liquid iron. J. Phys. Chem. Ref. Data **1**(35), 285–300 (2006)
13. Yang, C.C., Nakae, H.: The effects of viscosity and cooling conditions on the formability of aluminum alloy. J. Mater. Process. Technol. (141), 202–206 (2003)

Chapter 39
To Predict Surface Roughness and Linear Shrinkage of Die Casting Process by Using of Fuzzy Algorithm Model

Narendra Krishania and Anil Kumar Birru

Abstract This research paper narrates a manually constructed Mamdani based on fuzzy algorithm model for envisaging surface roughness and linear shrinkage of die casting, so that defects can be refrained from the casting in terms of surface finish and dimensions. A set of rules established by the help of mathematical model have been used to derive two fuzzy controllers which are being used in this process. With the help of this fuzzy algorithm, high production rate and high quality of products can be obtained by controlling process parameters. Confirmation experiments reveal that these fuzzy logics are able to attain optimum grouping of the process parameters. Hence, the quality of casted products in die casting process can be improved to a greater extent by this approach. The predicted surface roughness and linear shrinkage by this model had an error of only 3.55 and 6.02%, respectively, which was confirmed by checking the validity of the model developed by performing confirmation experiments. This proposed model can be reasonably utilized by the industries involved in die casting around the world to increase the overall effectiveness of the process and product.

Keywords Fuzzy logic · Orthogonal array · Surface roughness · Linear shrinkage

39.1 Introduction

Metal casting is a vital part of industrial production. It is generally used to develop near net shape component. Nowadays, the theory of direct/net shape manufacturing is attaining great importance in terms of lead time and cost reduction [1].

The die casting is an economical and efficient method offering a wide range of shapes than any other process. During manufacturing a metal part, the shapes of the

N. Krishania (✉) · A. K. Birru
National Institute of Technology Manipur, Imphal 795004, India
e-mail: anilbirru@nitmanipur.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_39

metals depend upon the injection temperature and liquid level. Die casting parameters like injection temperature, moisture level, injection pressure, and liquid level by using embedded controller. In the conventional system, the die casting parameters are measured by manual. The process of die casting involves five main stages which are clamping, injection, cooling, ejection, trimming. Basically, non-ferrous alloys are die casted like alloys of aluminium, alloys of copper, alloys of magnesium, and alloys of zinc. Die casting machines used are of two types—one is hot chamber machine (alloys with low melting temperature are die casted like zinc) and cold chamber machine (alloys with high melting temperature are die casted, like aluminium). Several engine components and pump components are manufactured using die casting.

Accuracy of die casting is affected by many factors like injection pressure too high, injection temperature too high, cooling time too short, low pouring temperature, clamp force too low, ejection force too high, non-uniform cooling rate.

To recognize the significant controllable process parameters which influence the quality of the casted products, it is essential to understand and find out the relationship among these controllable process parameters and quality of the casted products. The selection of the suitable process parameters for making qualitative casted products is one of the primary challenges in casting industry. It is very difficult to set appropriate process parameters with consideration of multiple performance characteristics and to depend on large amount of experimental operations. Therefore, it is important to effectively obtain the optimal process parameters to reduce trial and error time and consuming cost in the casting process [2].

Fuzzy reasoning is one of the prevailing tools which is used for modelling and governing undefined multi-criterion conclusion-making problems. It is grounded on the notion to obtain a set of association between inputs and outputs describing a procedure. So, the technique of fuzzy modelling conveys a nonlinear course better than several existing process [3]. Fuzzy logic techniques have been implemented by a number of researchers in their conclusion-making glitches. A fuzzy logic controller for forecasting the level regulator of molten steel in strip casting methods was created by Park and Cho [4]. With the aid of orthogonal array fuzzy linguistic approach, a general optimization system was implemented in the CNC turning operation for surface by Lan [5]. With the help of fuzzy-based Taguchi technique, Hsiang et al. [6] discovered the optimal process parameters that were maximizing the multiple performance characteristic index (MPCI) for hot extrusion of AZ31 and AZ61 magnesium alloy bicycle carriers [7, 8]. So, in the present study an effort has been made to create fuzzy logic controllers for estimation of quality of casting products with some provided set of process parameters like injection temperature, injection pressure, and time. It was originated that the predicted and measured values are in close proximity [9–12]. The closeness of results implies that the fuzzy modelling technique is efficient to estimate the quality of the patterns used for die casting.

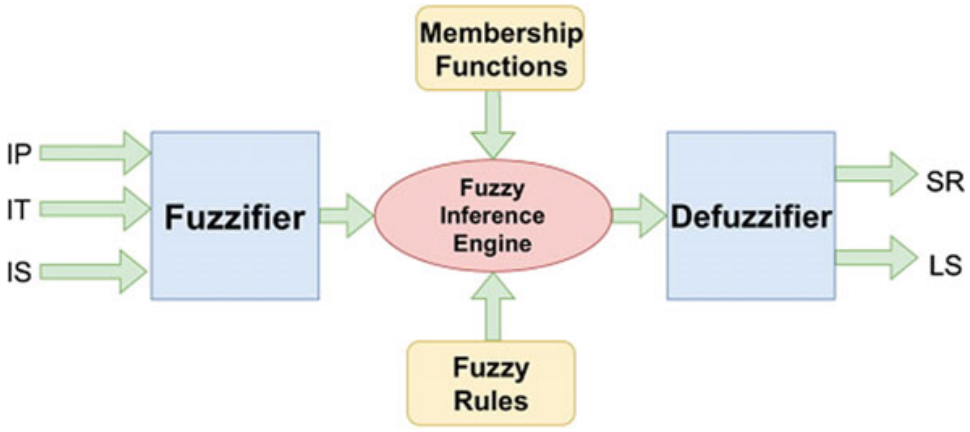


Fig. 39.1 Structure of three inputs two outputs' fuzzy logic

39.1.1 Fuzzy Logic Overview

Mamdani method is the most used fuzzy inference. Initially, one of the first fuzzy processes to regulate a steam engine and boiler was developed by Mamdani and Assilia in [13–15]. A fuzzy logic arrangement consisting of three inputs and two outputs is developed for this study which is shown in Fig. 39.1 [16]. Following steps are followed to execute Mamdani fuzzy inference:

(a) fuzzification of the input variables,
(b) rule output are evaluated and aggregated, and
(c) defuzzification.

Mamdani fuzzy rule is described as:

$$\text{Rule 1: If } y_1 \text{ is } M_1,\ y_2 \text{ is } N_1 \text{ and } y_3 \text{ is } O_1 \text{ then } z_1 \text{ is } P_1 \text{ and } z_2 \text{ is } Q_1 \text{ else}$$
$$\text{Rule 2: If } y_1 \text{ is } M_2,\ y_2 \text{ is } N_2 \text{ and } y_3 \text{ is } O_2 \text{ then } z_1 \text{ is } P_2 \text{ and } z_2 \text{ is } Q_2 \text{ else}$$
$$\text{Rule } n\text{: If } y_1 \text{ is } M_n,\ y_2 \text{ is } N_n \text{ and } y_3 \text{ is } O_n \text{ then } z_1 \text{ is } P_n \text{ and } z_2 \text{ is } Q_n \text{ else}$$

where M_i, N_i, O_i, P_i, and Q_i are termed as fuzzy subsets which are described by the analogous membership function [17–20].

39.2 Experimental Work

By performing the experiment with single variable at a time approach, the span of the particular process parameters is determined. The designed symbols and ranges of the process parameters are provided in Table 39.1. In the present study, the response parameters considered are surface roughness (SR) in µm and linear shrinkage (LS) in percentage.

Table 39.1 Factors and levels used in the experiments

Factor	Description	Unit	Range	Level 1	Level 2	Level 3
A	Pouring temperature	°C	650–750	650	700	750
B	Injector pressure	Bar	120–240	120	180	240
C	Plunger velocity	m/s	1.2–3.8	1.2	2.5	3.8

Table 39.2 L_9 orthogonal array used in the primary experiments

Experiment no.	A	B	C	SR	LS
1	650	120	1.2	48.5541	1.412
2	650	180	2.5	55.8354	1.384
3	650	240	3.8	63.0352	1.204
4	700	120	1.2	58.2841	2.974
5	700	180	2.5	65.3251	2.575
6	700	240	3.8	73.7327	2.045
7	750	120	1.2	61.0512	3.715
8	750	180	2.5	66.0263	3.315
9	750	240	3.8	75.1121	3.012

The primary experiments were conducted using nine test trials; these nine test trials were designed on the basis of L_9 orthogonal array. The reason of choosing L_9 orthogonal array is that shorter arrays are not capable of producing ample data to appropriately analyse the method since they are too simple and longer arrays would create complexity for the trial method. The response of assigned L_9 orthogonal array is shown in Table 39.2. In order to consider the three unlike performance characteristics in Taguchi method, corresponding values of the surface roughness (SR) and linear shrinkage (LS) are processed by the fuzzy logic unit.

39.2.1 Fuzzy Logic Modelling

In this work, three inputs and two outputs are used to evolve a fuzzy controller. Pouring temperature (PT), injector pressure (IP), and plunger velocity (PV) are three input variables, and surface roughness (SR) and linear shrinkage (LS) are two output variables. To make the fuzzy rule base, Mamdani fuzzy "If-Then" declarations for 27 rules has been considered, and to execute inference of max–min operation, Mamdani implication method has been adopted. To change fuzzy data to crisp response values, the centre of gravity method has been used by the defuzzifier. By acquiring triangular membership function, fuzzy division of input and output variables is carried out. Block diagram of fuzzy controller is shown in Fig. 39.2. This has been developed by MATLAB.

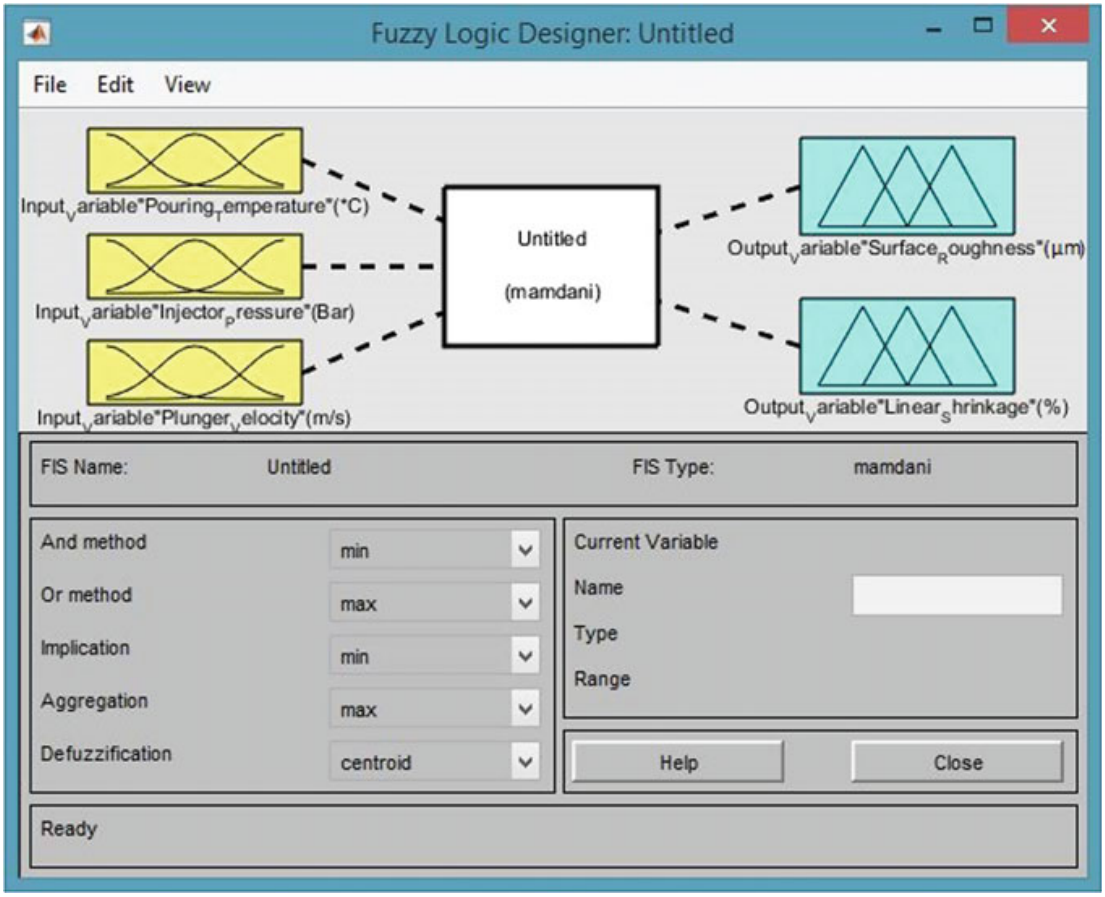


Fig. 39.2 Block diagram of fuzzy controller

Membership functions of the three inputs are small (S), medium (M), and large (L), and membership functions of the two outputs are very small (VS), small (S), medium (M), large (L), and very large (VL). Input and output functions are shown in Figs. 39.3 and 39.4. Table 39.3 shows the fuzzy intervals for input and output for forecasting the response, namely surface roughness (SR) and linear shrinkage (LS) of the casted products in die casting process. By taking into account the max–min combinational operation and with provided set of inputs, experimental results are originated on which these fuzzy rules are established. A non-fuzzy output is provided by the fuzzy reasoning of these rules as shown in Fig. 39.5.

Validation of Fuzzy Model Prediction

Three experiments were performed with different combination of the input process parameters which were not incorporated in the training set (randomly chosen). In Table 39.4, the corresponding experimental responses are measured and listed. From the fuzzy controller, the response values are again acquired for the same set of input process parameters. Then, to find out the extent by which the fuzzy controller is working, the percentage of error was calculated. The experimental and the predicted result obtained from the Mamdani fuzzy model has been plotted on graph as shown in Fig. 39.6. Also, bar graph has been drawn between experimental and fuzzy predicted data as shown in Fig. 39.7.

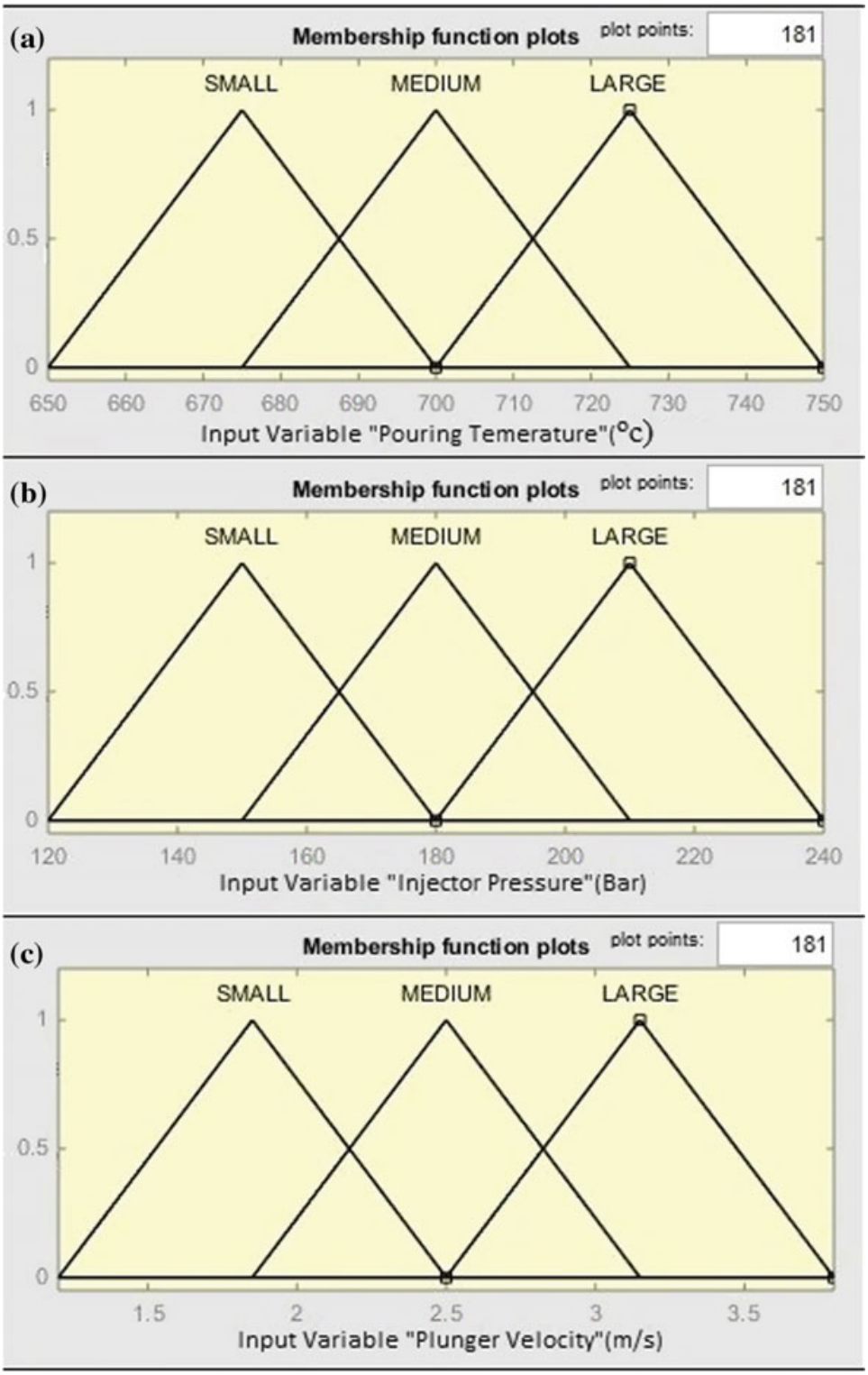


Fig. 39.3 Membership function of input variables: **a** pouring temperature, **b** injector pressure, and **c** plunger velocity

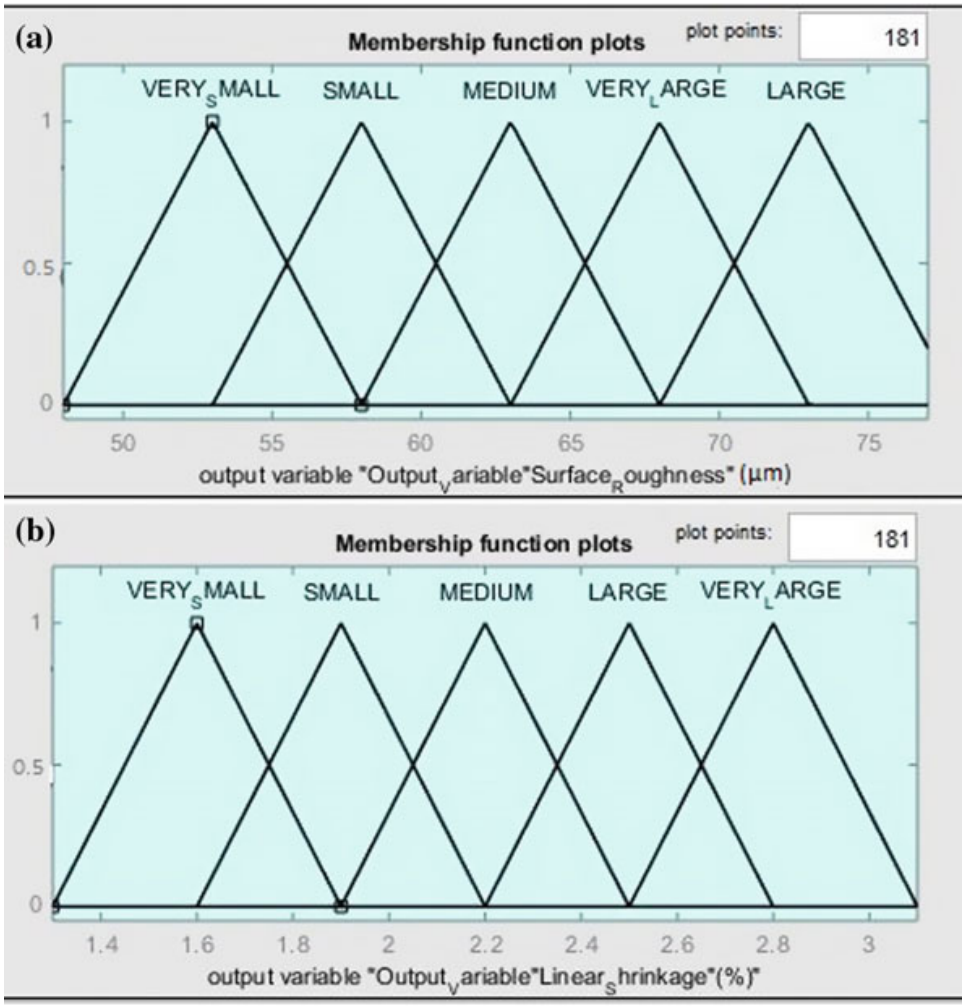


Fig. 39.4 Membership function of output variables: **a** surface roughness and **b** linear shrinkage

Table 39.3 Input output and fuzzy intervals

S. no.	System's linguistic variable		Linguistic values	Fuzzy interval (min–average–max)
1	Inputs	Injection speed	Small	650–675–700
			Medium	675–700–725
			Large	700–725–750
		Injection pressure	Small	120–150–180
			Medium	150–180–210
			Large	180–210–240
		Injection time	Small	1.2–1.85–2.5
			Medium	1.85–2.5–3.15
			Large	2.5–3.15–3.8

(continued)

Table 39.3 (continued)

S. no.	System's linguistic variable		Linguistic values	Fuzzy interval (min–average–max)
2	Outputs	surface roughness (μm)	Very small	48–53–58
			Small	53–58–63
			Medium	58–63–68
			Large	63–68–73
			Very large	68–73–78
		linear shrinkage (%)	Very small	1.3-1.6–1.9
			Small	1.6–1.9–2.2
			Medium	1.9–2.2–2.5
			Large	2.2–2.5–2.8
			Very large	2.5–2.8–3.1

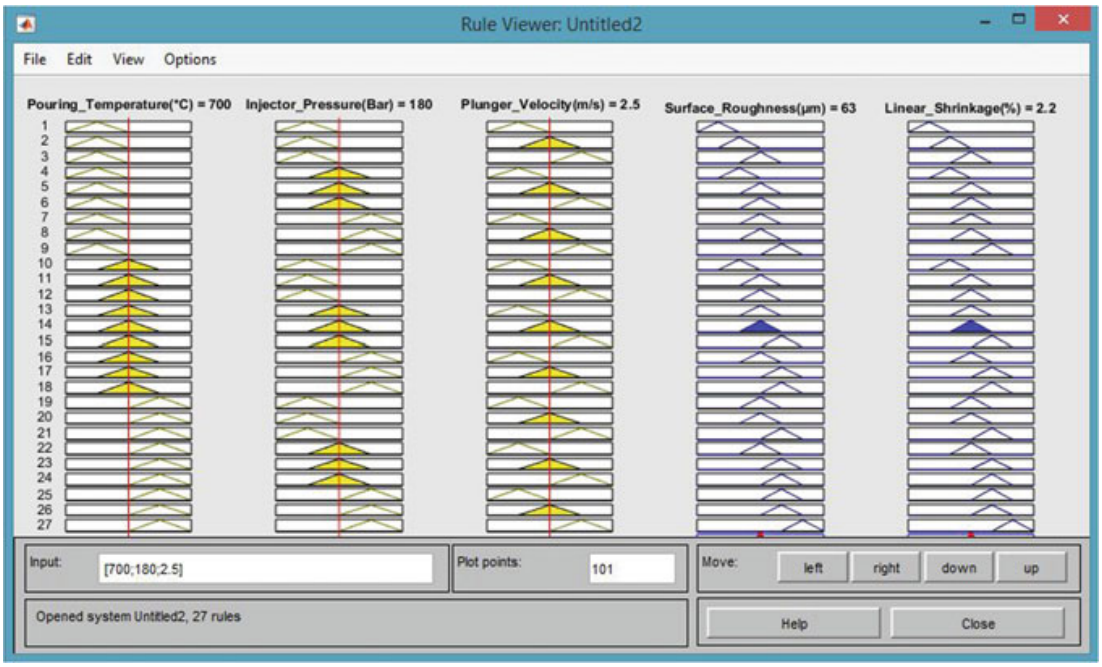


Fig. 39.5 Fuzzy inputs and prediction of outputs

Table 39.4 Comparison of results between the actual and the predicted values

S. no.	Injection process parameter			Experimental values		Fuzzy predicted values		% error	
	PT	IP	PV	SR	LS	SR	LS	SR	LS
1	650	120	1.2	48.554	1.412	52	1.5	7.09	6.23
2	700	180	2.5	65.325	2.075	63	2.2	3.55	6.02
3	750	240	3.8	75.112	3.012	79	3.3	5.17	9.56

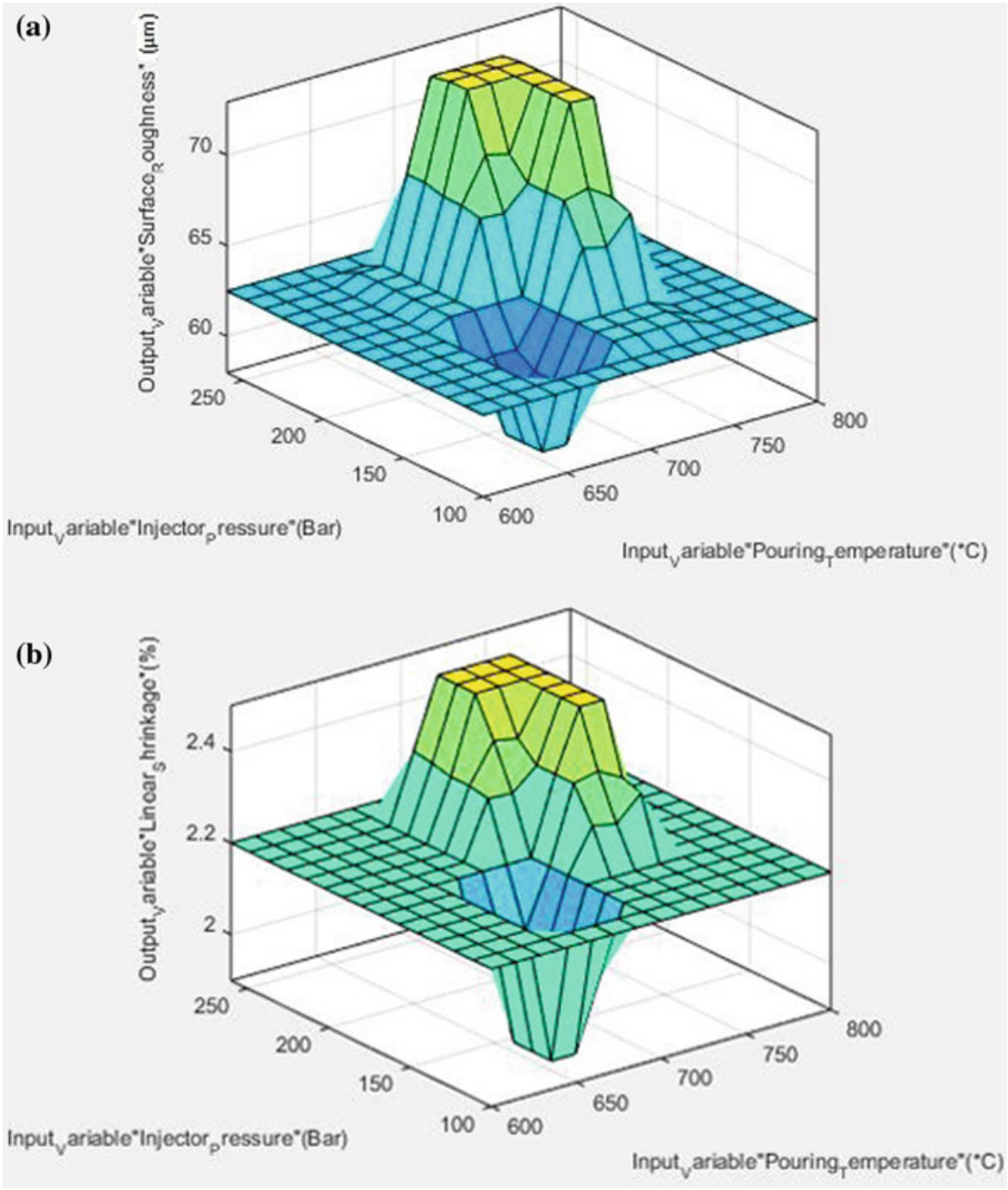


Fig. 39.6 Predicted surface roughness and linear shrinkage by fuzzy logic in relation to parameters change: **a** pouring temperature and injector pressure and **b** pouring temperature and injector pressure

39.3 Discussion

In this work, to predict the quality characteristics, namely surface roughness and linear shrinkage of casted product in die casting process, a fuzzy logic controller using Taguchi orthogonal array design has been developed. Prediction results show that Mamdani inference was on the basis of maximum error calculated. So, for

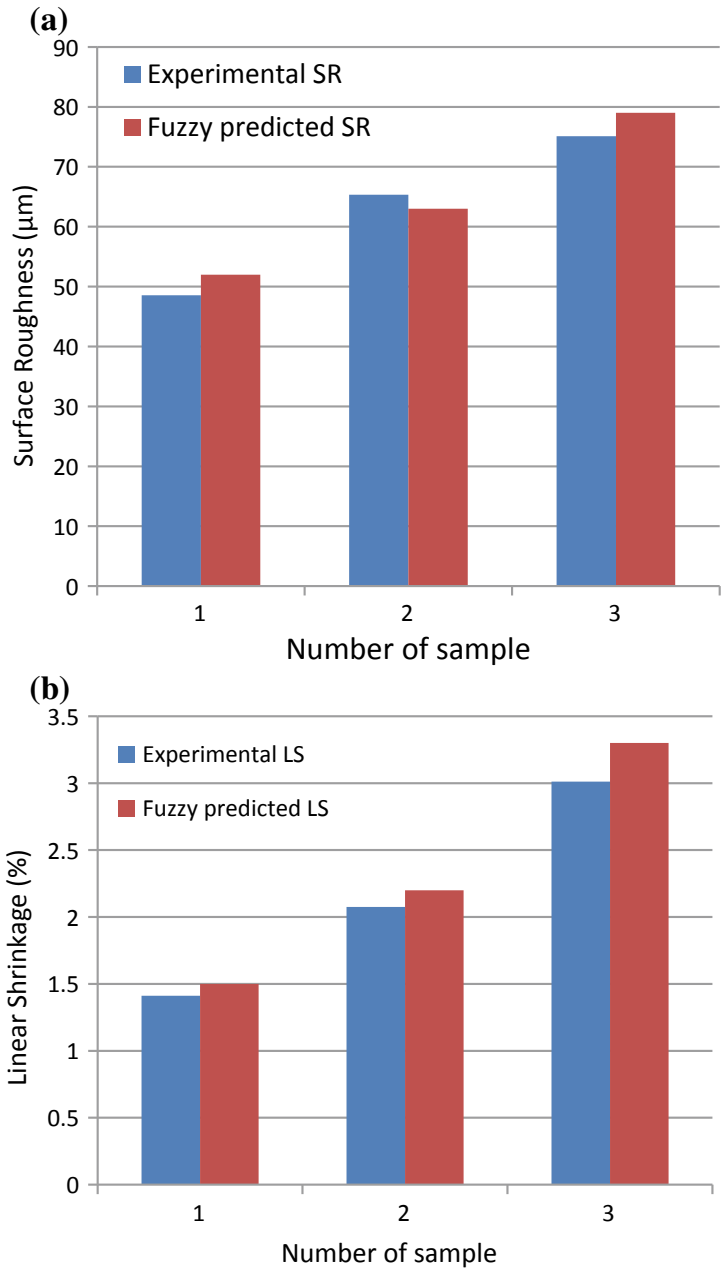


Fig. 39.7 Prediction of **a** surface roughness and **b** linear shrinkage

improving the quality and reliability of the entire process, it has proved to be an effective and powerful tool. By enhancing the number of experiments, the anticipated values of fuzzy output can be further precisely predicted. It is also very useful and can also be used for the optimization of multiple performance quality process parameters and casted product material, to test the ability of the expert systems in prediction of the outputs.

39.4 Conclusion

From the present study, resulting conclusions are summarized as follows:

- The capability of generalization and prediction of pattern characteristics within the range of experimental data such as surface roughness and linear shrinkage of casted product in die casting process have been shown with the help of Mamdani fuzzy inference.
- Industries which are using the experience of skilled workers can use a set of fuzzy rules which further result in improvement in forecasting ability of the process.
- For automation of the process, this methodology seems to be beneficial.

References

1. Campbell, J.: The concept of net shape for casting. Mater. Des. **21**, 373–380 (2000)
2. Mahapatra, S.S., Nanda, S.K., Panigrahy, B.K.: A Cascaded fuzzy inference system for Indian river water quality prediction. Adv. Eng. Softw. **42**, 787–796 (2011)
3. Khrais, S., Al-Hawari, T., Al-Araidah, O.: A Fuzzy logic application for selecting layered manufacturing techniques. Expert Syst. Appl. **38**, 10286–10291 (2011)
4. Park, Y., Cho, H.: A fuzzy logic controller for the molten steel level control of strip casting processes. Control Eng. Pract. **13**, 821–834 (2005)
5. Lan T.S.: Fuzzy linguistic optimization on surface roughness for CNC turning. Math Prob Eng, pp. 1–10, Hindawi Publishing Corporation (2010)
6. Hsiang, S.H., Lin, Y.W., Lai, J.W.: Application of fuzzy-based Taguchi method to the optimization of extrusion of magnesium alloy bicycle carriers. J. Intell. Manuf. (2010)
7. Zimmermann, H.: Fuzzy Set Theory and Its Applications Second. Revised Edition, Springer (1992)
8. Xu, D.: Applications of Fuzzy Logic in Bioinformatics. Imperial College Press, London, UK (2008)
9. Ramesh, S., Karunamoorthy, L., Palanikumar, K.: Fuzzy modeling and analysis of machining parameters in machining titanium alloy. Mater. Manuf. Processes **23**, 439–447 (2008)
10. Lo, S.-P.: An adaptive-network based fuzzy inference system for prediction of workpiece surface roughness in endmilling. J. Mater. Process. Technol. **142**, 665–675 (2003)
11. Jiao, Y., Lei, S., Pei, Z.J., Lee, E.S.: Fuzzy adaptive networks in machining process modeling: surface roughness prediction for turning operations. Int. J. Mach. Tools Manuf. **44**, 1643–1651 (2004)
12. Gupta, A., Singh, H., Aggarwal, A.: Taguchi-fuzzy multi output optimization (MOO) in high speed CNC turning ofAISI P-20 tool steel. Expert Syst. Appl. **38**, 6822–6828 (2011)
13. Mamdani, E.H., Assilia, S.: An experiment in linguistic synthesis with a fuzzy logic controller. Int. J. Man Mach. Stud. **7**, 1–13 (1975)
14. Dussud, M., Galichet, S., Foulloy, L., Simonnin, P.: Fuzzy logic control of continuous casting installations. In: Proceedings of the13th IFAC World Congr., San Fransisco, CA, vol. M, pp. 469–474 (1996)
15. Foulloy, L., Galichet, S.: Typology of fuzzy controllers. In: Nguyen, H.T., Sugeno, M., Tong, R., Yager, R. (eds.) Theoretical Aspects of Fuzzy Control, pp. 65–90. Wiley, New York (1995)
16. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 338–353 (1965)

17. Saarlainen, E.J., Lautala, P., Inkinnen, M., Johansson, J.: Steel caster mold level control using fuzzy logic. In: 13th PTD Conference Proceedings 1995, pp. 397–402
18. Klir, G.J., Yuan, B.: Fuzzy Sets and Fuzzy Logic: Theory and Applications. Prentice-Hall, New Delhi (1997)
19. Bouhouch, S., Lahreche, M., Moussaoui, M., Bast, J.: Quality monitoring using principal component analysis and fuzzy logic. Appl. Continuous Casting Process, Am. J Appl. Sci. **4**(9), 637–644 (2007)
20. Bartolini, G., et al.: Development of performance adaptive fuzzy controllers with application to continuous casting plants. In: Sugeno, M. (ed.) In Industrial Applications of Fuzzy Control. Amsterdam, Elsevier, The Netherlands (1985)

Chapter 40
Prediction of Hot Spot and Hot Tear of the Al–Cu Cast Alloy by Casting Simulation Software

Kumara Swamy Pulisheru and Anil Kumar Birru

Abstract The process of casting in foundry is complex and involves large process to get the required part, and such process needs simulation before manufacturing. Casting simulation may save material and gives optimal process of casting in foundry. The defect minimization or prevention is challenging and plays significantly in enhancing the productivity of the foundry. The aim of the present study is to compute hot spots and hot tearing in Al–Cu cast alloy. The originating points of the feed paths represent the hot spots that lead to a shrinkage defect. Some studies have been carried out on hot tear. It was observed that the location of hot tear is at inter-junction of mold cavity to the gate. The developed model is used as bench mark to predict the location of hot spot and hot tear of Al–Cu cast alloy by green sand casting method with casting simulation software.

Keywords Hot spot · Hot tear · Sand casting · Al–Cu cast alloy · Simulation

40.1 Introduction

Poor quality and low yield are major issues for foundries which will increase the rejection rate. Additional melting of molten metal, which directly reflects the quality of the casting and casting yield, highly depend upon solidification of molten metal in the cavity. The direction of molten metal referred to as a feed path which is automatically generated using the vector element method (VEM). The volumetric contraction during solidification compensated from adjacent hotter regions of molten metal. Solidification of casting progressively takes place from the mold boundary which isolates last solidifying points in the mold cavity said to be a 'hot spot' which are the most probable locations for shrinkage related defects. Computer-aided design and simulation have been playing vital role from the last two decades. Casting simulation is to improve the quality and casting yield in

K. S. Pulisheru · A. K. Birru (✉)
National Institute of Technology Manipur, Imphal 795004, India
e-mail: anilbirru@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_40

shorter time in shop floor trial-and-error approach. The effect of feeder location, shape, size, and feed aids might give tangible information by casting simulation software. The minimization of defects may be performed by an appropriate feeding system to ensure directional solidification from thin to thick sections of the casted produced. Sutaria et al. [1] observed that computed feed path and hot spot by combining level set method were the last solidified region in casting by vector notation method. Computed feed paths during solidification represent the mass-less particles along the liquid–solid interface. Sutaria et al. [2] investigated with the application of gradient vector method (GVM) and found that directional solidification and hot spot in the casting observed by pouring three layouts of a multi-junction casting. Sutaria et al. [3] investigated the optimization of casting feeders by vector element method (VEM) using feed paths.VEM resulted in terms of hot spot locations correlated with FEM.

Choudhari et al. [4] it was observed that the optimizing feeder dimensions and increasing feeding efficiency of casting by minimizing of the casting defect. Casting simulation enables visualization of progressive freezing regions inside the casting to locate the hot spot. Choudhari et al. [5] studied that the defect minimization in casting to improve the casting yield was compared with its experimental validation. Similarly, applied simulation-based gradient vector method is used to identify the location of shrinkage porosity.

Jain et al. [6] comparative study was made on sand casting and investment casting to minimize the tooling cost for reduction of production cost. It was found that the tooling cost was more in case of sand casting due to additional allowances to get defect-free casting which gives more errors in dimensional allowances, whereas in case of investment casting get more surface finishing. Hodbe et al. [7] studied the manual design, fabrication of tools, and trial-and-error in process to minimize the defects, said process was repeated to get defect-free casting in foundry industries. The defect-free casting increases the casting yield. Simulation-based trials save the time, loss of material, and other industrial expenditure. Rabindra et al. [8] investigated the solidification behavior and detection of hotspots in aluminum alloy castings: Computer-aided analysis by AutoCAST and compared with experimental validation. Birru and Benny Karunakar [9] investigated minimization of hot tear by adding the grain refiners with different levels of composition Al–2.5, Ti–0.5C and Al–3.5, Ti–1.5C with the variation of pouring temperatures as 700, 750, and 780 °C. Gajbhiye et al. [10] investigated the reducing shrinkage porosity in a sand casting process by simulation in AutoCAST-X software with experimental validation by destructive testing. Hence, much work is not emphasized to compute hot spots and hot tearing in Al–Cu cast alloy, and especially in sand casting, authors made an attempt to predict the hot tearing and hot spots.

40.2 Methodology

40.2.1 Problem Definition

The objective of the study is to identify the defects in Al–Cu cast alloy casting such as hot spot and hot tear by AutoCAST simulation software. By optimizing the defects in casting enables to improve the casting yield and quality of the casting part.

40.2.2 Process Design

The entire study has been divided as section, viz. design of model, numerical simulation using AutoCAST software, and analyzing the results. The entire process of simulation is as shown in Fig. 40.1. Casting material is used as Al–Cu alloy and mold is taken as green sand as per aimed work for the sand casting foundry requirements. The CAD model is developed in Creo software. Design parameters are taken from the design calculation as depicted in Table 40.1. The material properties are mentioned in Table 40.2, restricted to simulation software which is considered as boundary conditions to casting simulation.

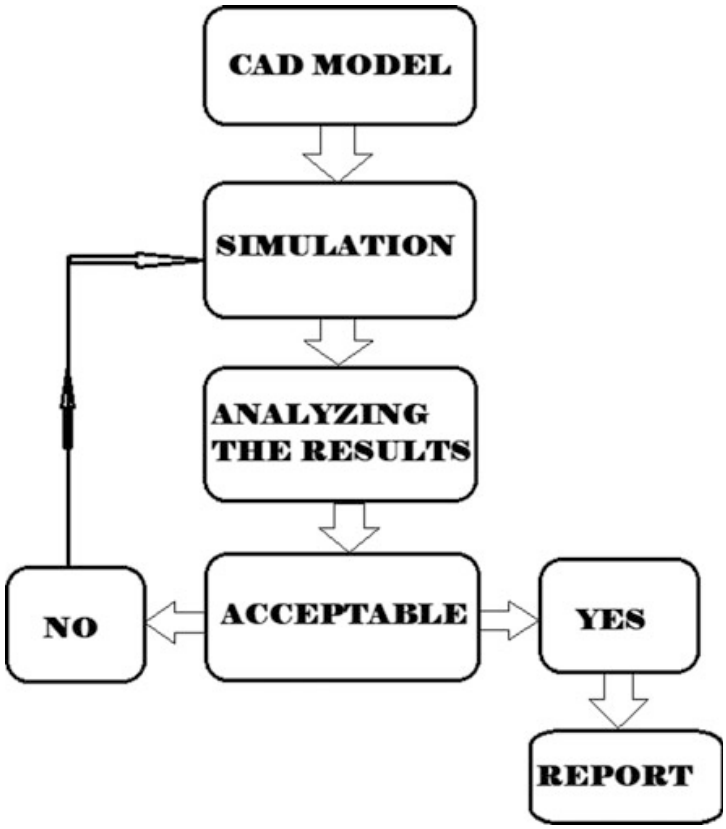


Fig. 40.1 Flow chart for the casting simulation in foundry

Table 40.1 Design parameters

S. N.	Design parameters	Value (mm)	Cross section
1	Pouring basin, height, length, and width	8, 20, 15	Rectangular
2	Diameter of sprue at top, bottom, and sprue length	10, 8, 23.5	Cylindrical
3	Sprue well diameter and height	10, 5	Cylindrical
4	Ingate dimensions, width, height, and length	8, 4, 52	Rectangular

Table 40.2 Material properties

S. N.	Description	Designated value	Units
1	Density	2800	Kg/m^3
2	Thermal conductivity	121	W/m-K
3	Specific heat	920	J/Kg-K
4	Solidus temperature	570	°C
5	Liquidus temperature	650	°C

The design of pattern helps to obtain the acceptable design mold which leads to get good quality casting in foundry industries. The design of part consists of various elements such as pouring basin, sprue, sprue well, and ingates, and they act as a channel to molten metal to flow into the mold cavity to get the desired shape of the mold cavity.

40.2.3 Numerical Simulation

Numerical simulation describes the applications of AutoCAST software for method design and optimization of casting process to enable to get defect-free casting. In the present investigation, simulation process of the selected CAD part was modeled in (.stl file), imported into AutoCAST for simulation. The mold box dimensions have been considered as 200 × 100 × 80 mm as shown in Fig. 40.2. The simulation was carried out by choosing the green sand casting process with aluminum-based alloy. The mold is subdivided into cubic elements with the element size of 1 mm for internal computation for uniform solidification and mold filling. The pouring temperature is considered 700 °C in casting simulation. It was observed from Fig. 40.3 that the Al–Cu alloys were much prone at TC-1 region and TC-2 region.

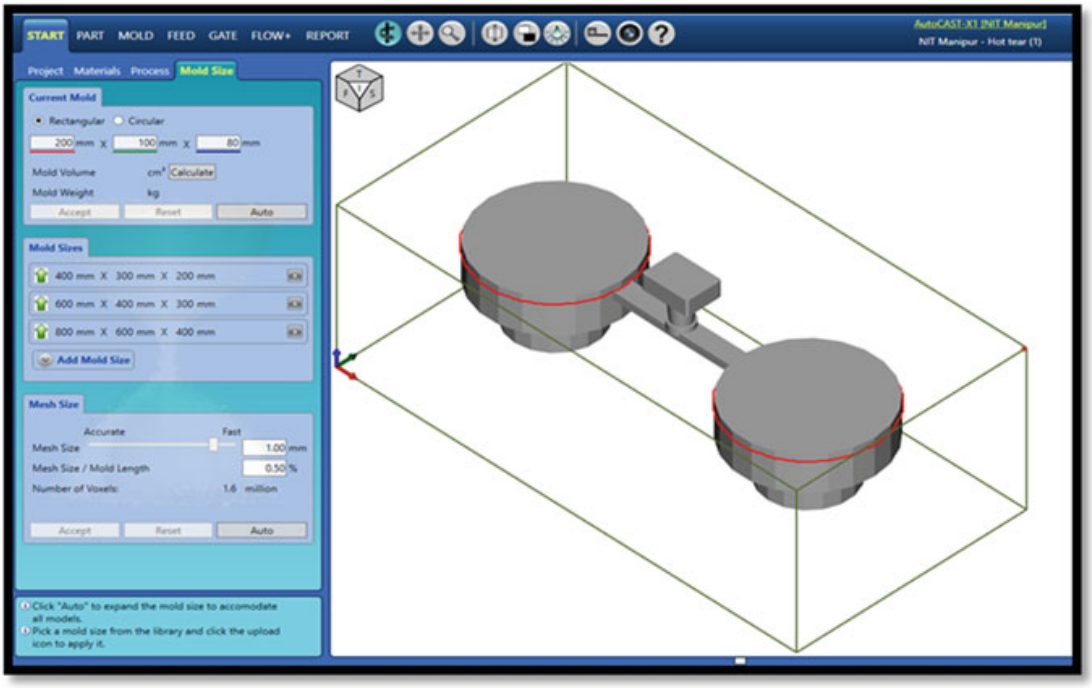


Fig. 40.2 Mold box dimensions for the present investigations

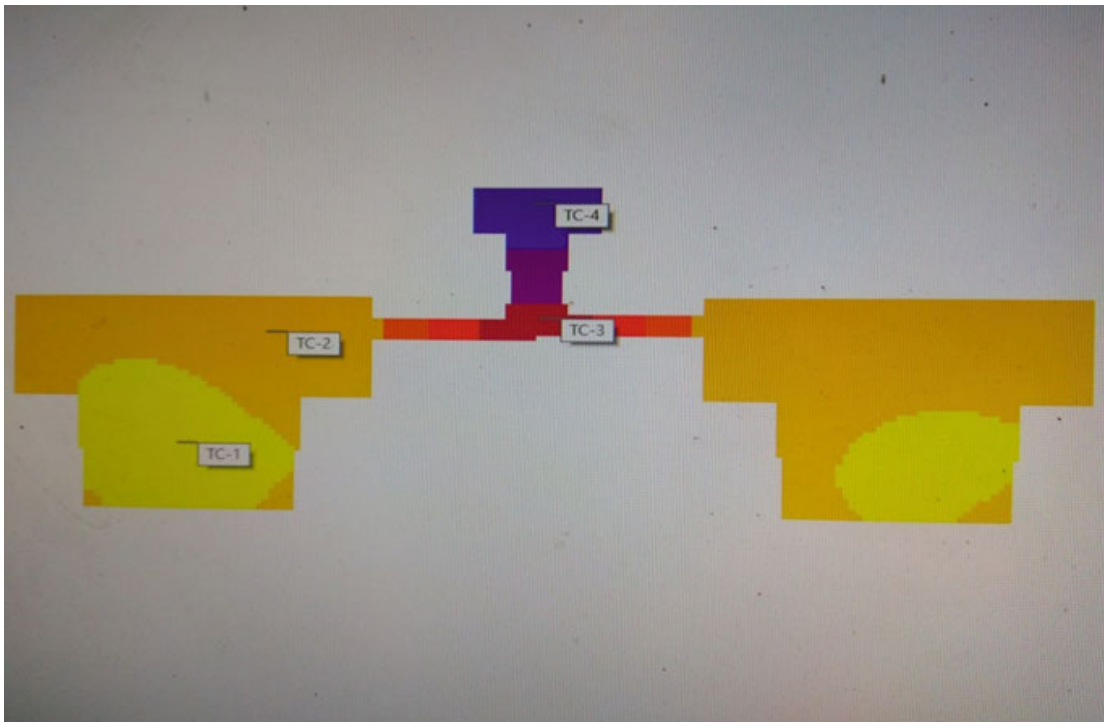


Fig. 40.3 Temperature at TC-1 to TC-4 regions in the casted part

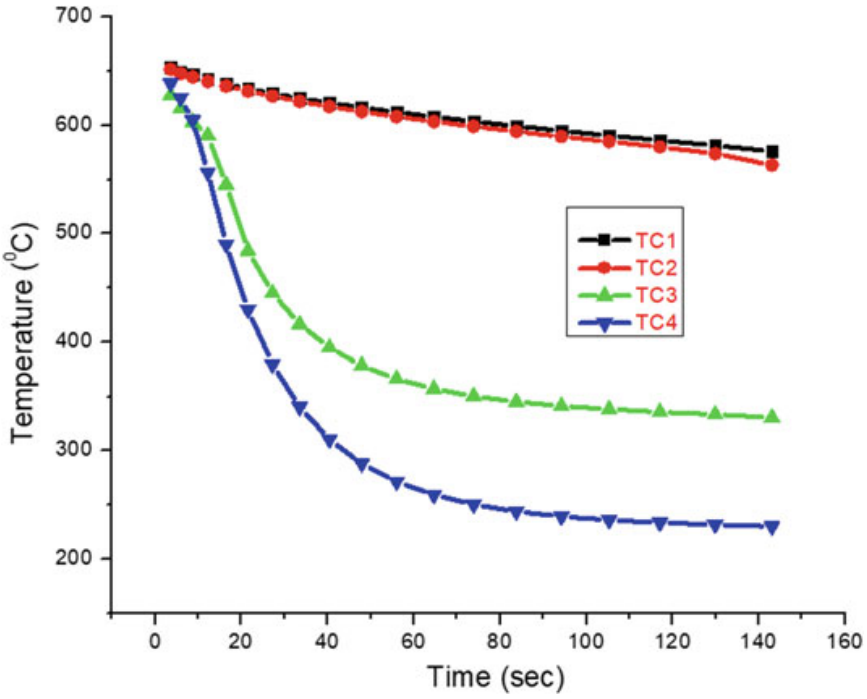


Fig. 40.4 Effect of pouring temperature on different zones in Al–Cu cast alloy

Hot tear was shown at the left end of the casted product, similarity index at the right side of the casted product was observed, TC-3 region the molten metal was distributed at both the sides uniformly, and TC-4 was observed as hard zone of the Al–Cu alloy casted product.

The solidification of molten metal has been compared with the time Vs temperature as shown in Fig. 40.4. An average of four K type thermo-couples at different regions in casting solidification process was observed. TC-1, TC-2 was hot sport and hot tear zones, respectively TC-3 is junction zone were the molten Al–Cu alloy flows, and TC-4 is the hard zone.

40.3 Results and Discussion

In Fig. 40.5, it was observed that the feed metal paths show the solidification direction of Al–Cu alloy casting in 2D and 3D sectioned casting with feed path vector notation as mentioned below

- Progressive Solidification-Cooling
- Directional Solidification-Feed metal paths.

The progressive solidification is represented by isothermal maps (at equal temperature). The directional solidification is represented by feed paths (with temperature gradients).

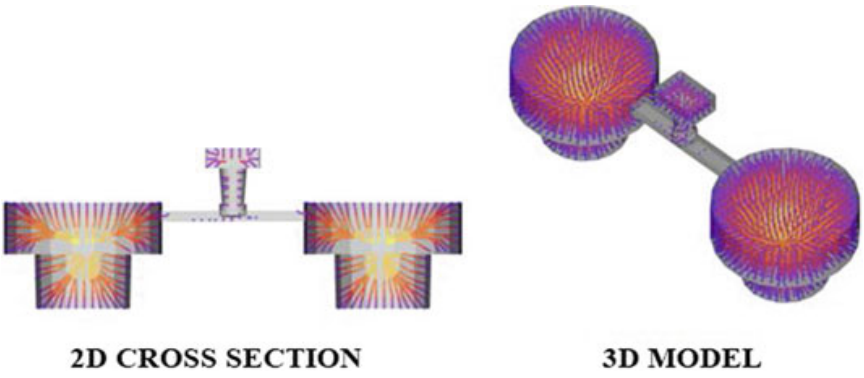


Fig. 40.5 Feed paths by vector notation method of 2D cross section and 3D model

The last solidified region in casting called the hot spot which leads to volumetric shrinkage of the molten metal in the process of solidification. In casting process to compensate the volumetric shrinkage of the molten metal in solidification process the feeder has facilitate to overcome the shrinkage porosity in the casting part. The casting simulation enables to identify the hot spot region to facilitating the feeder for getting the quality casting in foundry.

Figures 40.6 and 40.7 it shows the hot tear occurs where the weakest zone in cavity while in solidification process due to sudden cooling of the molten metal from high temperature to low temperature. Casting simulation enables to identify the weaker zone in mold cavity for the redesign of mold cavity to overcome the hot tears in casting hence, to get the good quality casting part in foundry.

The hot spot and hot tear zones were correlated with Figs. 40.6 and 40.7 for TC-1 to TC-4, the significance of pouring temperature with respect to the solidification time has effect on the Al–Cu cast alloy. In the process of solidification noticed that the TC-3 conditions as the temperature decrease with respect to the time, the flow of material decreased as the time increased and it was also noted that drastically drop down in the temperature was observed at 20–40 s. Similarly, hard zone at TC-4 condition, the temperature drastically drops down at 15–35 s. Hence, it leads to the hot tearing defect in casting solidification. Similarly, for TC1-TC2, the hot spot was observed at the verge of solidification, i.e., TC-1, respectively,

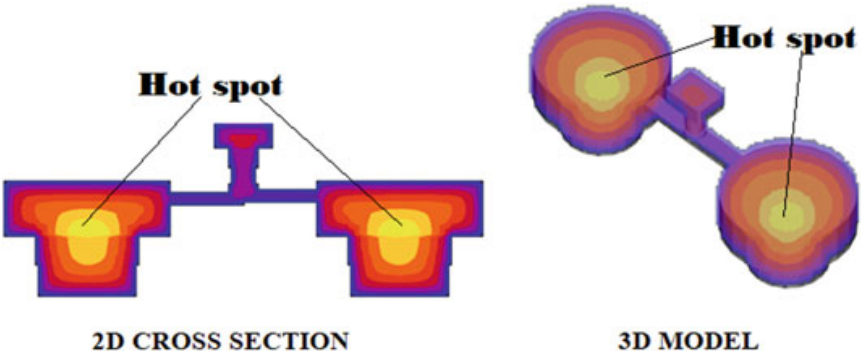


Fig. 40.6 Hot spot region of Al–Cu cast alloy: 2D-cross-sectional view and 3D-isometric view

Fig. 40.7 Hot tear region of Al–Cu cast alloy: 2D-top view and 3D-model

before the formation of the hot spot, the ending of the hot tear starts in Al–Cu cast alloy, and it was further scope to the authors that the grain refinement may also consider for aforesaid alloy.

40.4 Conclusion

Casting simulation was implemented for Al–Cu alloy with green sand mold which enables to predict the problems in casting process in foundry industries. With a motive to minimize the defects in casting especially for green sand molds to acquire the quality of the castings

1. The present study enabled to identify the hot spot and hot tears in casting part by casting simulation.
2. It facilitated optimization of design and casting yield for the preceding process to overcome the hot spot and hot tears by providing feed aids and redesign of the mold cavity.
3. It was concluded that casting simulation enables for trial-and-error to get the accuracy in casting by optimizing the defects in casting, which saves the time and loss of material in foundry which ultimately impacts to enhance the productivity of the industry.
4. It was observed that the Al–Cu alloy in the green sand mold system, the hot tearing and hot spot occurred in the range of 600–700 °C as the solidification time increases at 120–140 s with a minor variation in the hot tearing and hot spot hence, it was concluded that the hot spot increases, respectively, as hot tearing increases vice versa.

Acknowledgements Authors are much thankful for the TEQUIP-III, NPIU-New Delhi, for financial support for purchase of AutoCAST software to carry out the simulation of designed casting.

References

1. Sutaria, M., Gada, Vinesh H., Sharma, A., Ravi, B.: Computation of feed-paths for casting solidification using level-set-method. J. Mater. Process. Technol. **212**, 1236–1249 (2012)
2. Sutaria, M., Ravi, B.: Computation of casting solidification feed-paths using gradient vector method with various boundary conditions. Int. J. Adv. Manuf. Technol. **75**, 209–223 (2014)
3. Sutaria, M., Joshi, D., Jagdishwar, M., Ravi, B.: Automatic optimization of casting feeders using feed-paths generated by VEM. In: Proceedings of the ASME 2011 International Mechanical Engineering Congress & Exposition IMECE-2011, Nov 11–17, 2011, Denver, Colorado, USA
4. Choudhari, C.M., Narkhede, B.E., Mahajan, S.K.: Casting design and simulation of cover plate using AutoCAST software for defect minimization with experimental validation. In: 3rd International Conference on Materials Processing and Characterisation (ICMPC 2014), Procedia Materials Science 6, pp. 786–797 (2014)
5. Choudhari, C.M., Narkhede, B.E., Mahajan, S.K.: Methoding and simulation of LM 6 sand casting for defect minimization with its experimental validation. In: 12th Global Congress on Manufacturing and Management, GCMM 2014, Procedia Engineering 97, pp. 1145–1154 (2014)
6. Jain, P., Kuthe, A.M.: Feasibility study of manufacturing using rapid prototyping: FDM approach. The Manufacturing Engineering Society International Conference, MESIC 2013, Procedia Engineering 63, pp. 4–11 (2013)
7. Hodbe, G.A., Shinde, B.R.: Design and Simulation of Lm 25 sand casting for defect minimization. In: ICMPC 2017, Materials Today: Proceedings 5, pp. 4489–4497 (2018)
8. Rabindra, B., Kayal, S., Sutradhar, G.: Solidification behavior and detection of hot spots in aluminium alloy castings: computer aided analysis and experimental validation. Int. J. Appl. Eng. Res., Dindigul, 1(4), (2011)
9. Birru, A.K., Benny Karunakar, D.: Effects of grain refinement and residual elements on hot tearing of A713 aluminium cast alloy. Trans. Non ferrous Met. Soc. China **26**, 1783–1790 (2016)
10. Gajbhiye, A.K., Choudhari, C.M., Raut, D.N., Narkhede, B.E., Bhandarkar, B.M.: Minimization of shrinkage porosity in A sand casting process by simulation in AUTOCAST software with experimental validation by destructive testing. Int. J. Modern Eng. Res. (IJMER), **4**(5), 18 (2014)

Chapter 41
Influence of Pouring Temperature on Fluidity of Al–Cu Cast Alloy: Analyses Using Casting Simulation Software

Kumara Swamy Pulisheru and Anil Kumar Birru

Abstract The fluidity of molten metal plays a vital role in the casting process when we are confined in manufacturing thinner sections. The aim of the present study is to determine the effect of pouring temperature on the fluidity of Al–Cu cast alloy. The simulation was performed by AutoCAST-X software with altered pouring temperatures of 700, 750, and 800 °C. It was observed that the fluidity was enhanced when the pouring temperature at 800 °C for the aforesaid cast alloy. Similarly, it was found that the volume of fill increased with the increasing pouring temperature and also strong agreement was observed with the solidification time and fill time.

Keywords Sand casting · Fluidity · Al–Cu cast alloy · Pouring temperature · Simulation

41.1 Introduction

In the manufacturing process, casting is one of the most economical processes among all the available process even in the twenty-first century with respect to the volume of the production. Aluminum alloys are one of the most widely used materials from the past few decades and versatile raw material found in the common foundry metal. The ratio of the cast to wrought aluminum alloy products is increasing in a number of castings being used in automotive, aerospace industries, and domestic goods. The advantage of aluminum alloys mainly from their low density, moderate castability, and excellent mechanical properties.

K. S. Pulisheru · A. K. Birru (✉)
National Institute of Technology Manipur, Imphal 795004, India
e-mail: anilbirru@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_41

In manufacturing industry, computer-aided design and simulation play a vital role in forecasting the problems in the manufacturing process for improving of production and quality of the product. The application of casting simulation is increasing in the foundry as it essentially replaces or minimizes the shop-floor trials to achieve the desired internal quality at the highest possible time. Casting solidification simulation is now widely used to visualize the effect of feeder location, shape, size, and feedaids (such as feeder sleeves and mold chills) on the solidification sequence, temperatures, and cooling rates. Based on the earlier researchers, the feeding system design is iteratively modified and simulated until an optimal design the smallest number and volume of feeders to prevent casting shrinkage defects is achieved. In a casting process, the term "fluidity" indicates the distance of the molten metal can flow in a mold at a constant cross-sectional area before it solidifies [1]. The fluidity in casting enhances the casting yield. Over the decades, a considerable amount of experimental and numerical work has been carried out to understand casting solidification and feeding mechanism. Mohan Krishna et al. [2] investigated understanding of the fluidity behavior of narrow and wide freezing range aluminum alloys, respectively (47,100, and 51,300), then extended to cast superalloy CMSX4 and ductile cast iron (GJS-400-18) product. It was found that spiral and complex shaped sand and mullite molds with molten alloys having varied superheats of 25, 50, and 100 °C over the liquid temperatures were simulated using finite element method-based software. Birru et al. [1] investigated the fluidity of Al–Zn alloys, such as the standard A713 alloy with and without scrap addition. Fluidity measurements were performed with a double spiral fluidity test consisting of gravity casting of double spirals in green sand molds with good reproducibility. The influence of recycled alloy on fluidity has been compared with that of the virgin alloy, and the results showed that the fluidity decreased with the increase in recycled alloy at minimum pouring temperatures. Joshi et al. [3] investigated the analytical hierarchy process to the determination of relative weights of various criteria for castability evaluation. The criteria are categorized in a three-level hierarchy, with the middle level comprising mold, feeding, and gating groups. The methodology is illustrated by modifying the product and tooling design of a gray iron bracket casting and estimating the improvement in manufacturability. Manoj et al. [4] studied on casting simulation of a virtual process which ensures casting without defects. It involves computer-aided modeling, pattern design, simulation, and optimization. Fluidity of molten metal plays an important role in producing sound complex castings. The objective of the analysis study is to identify the factors affecting casting fluidity by changing sprue height, pouring temperature, and flow rate on the fluidity of aluminum alloy (AlSi-13).

Ravi et al. [5] studied the desired quality at the least cost without shop-floor trials, with design, process simulation, and quality evaluation and cost estimation carried out. It was found that simulation results are evaluated using three quality indices, moldability, fillability, and feedability. Sutaria et al. [6] investigated the volumetric contraction accompanying solidification of molten metal manifests in

defects like shrinkage cavity, porosity, centerline shrinkage, corner shrinkage, and sink. These defects may be minimized by designing an appropriate feeding system to ensure directional solidification from thin to thick sections in the casting, leading to feeders. Ravi et al. [7] identified the location of a feeder, calculating to minimum size creating its solid model, attaching the feeder model to the casting, carrying out solidification simulation, predicting quality and estimating the yield. Naveen et al. [8] studied the castings that serve these specific applications have to achieve the quality requirements of superior mechanical properties and zero porosity. To achieve these objectives within a limited time frame in a product development process, CAD technologies combined with process simulation tools are increased to optimize the form filling and solidification of the cast parts. Hence, the newly developed simulation of flywheel component that was prototyped via sand casting route. Results of casting trials showed a high level of confidence in the simulation tools.

41.2 Methodology

41.2.1 Problem Definition

The objective of the present study on the fluidity of Al–Cu cast alloy with altered pouring temperature by choosing the method of green sand mould process. In which effect of pouring temperature on the casted part is being analyzed to optimize the defects in the casting process by choosing an appropriate method design. The optimization of method design is carried out by casting simulation.

41.2.2 Process Design

The entire study has been divided into sections, viz. design of the model, numerical simulation using AutoCAST software, and lastly analyzing the results. The entire process is as shown in flowchart as in Fig. 41.1. The entire process shows the method design in the foundry.

To study the fluidity of Al–Cu cast alloy with altered pouring temperatures such as 700, 750, and 800 °C considered the model as a rectangular plate with the size of 220 × 140 × 5 mm by Cre-O software as shown in Fig. 41.2 and converted it into **.stl** file to import it into an AutoCAST software for simulation.

Casting material is used as A206 [9], the chemical composition is as shown in Table 41.1, and mold material is taken as green sand with the dimensions of 300 × 100 × 160 as shown in Fig. 41.3, as per the foundry requirement. The design of the gating system consists of various elements like pouring basin, sprue,

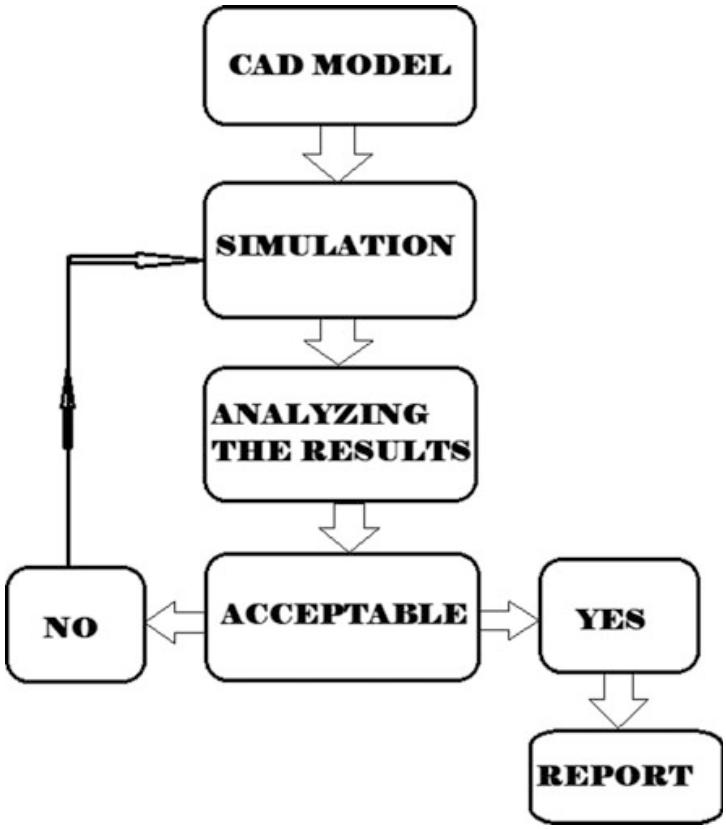


Fig. 41.1 Flowchart for the casting simulation

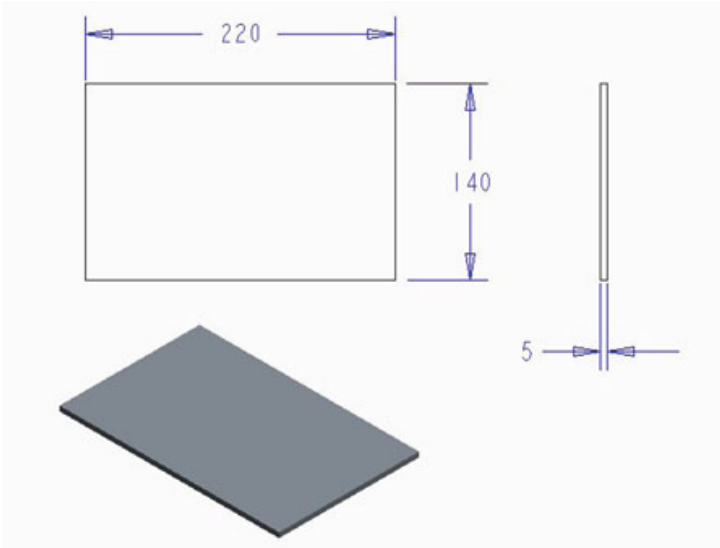


Fig. 41.2 Part model by Cre-O for simulation

Table 41.1 Chemical composition of A206 alloy [9]

Element	Cu	Mn	Mg	Si	Ti	Sr	Al
% of composition	5.0	0.5	0.35	0.05	0.25	0.20	Balance

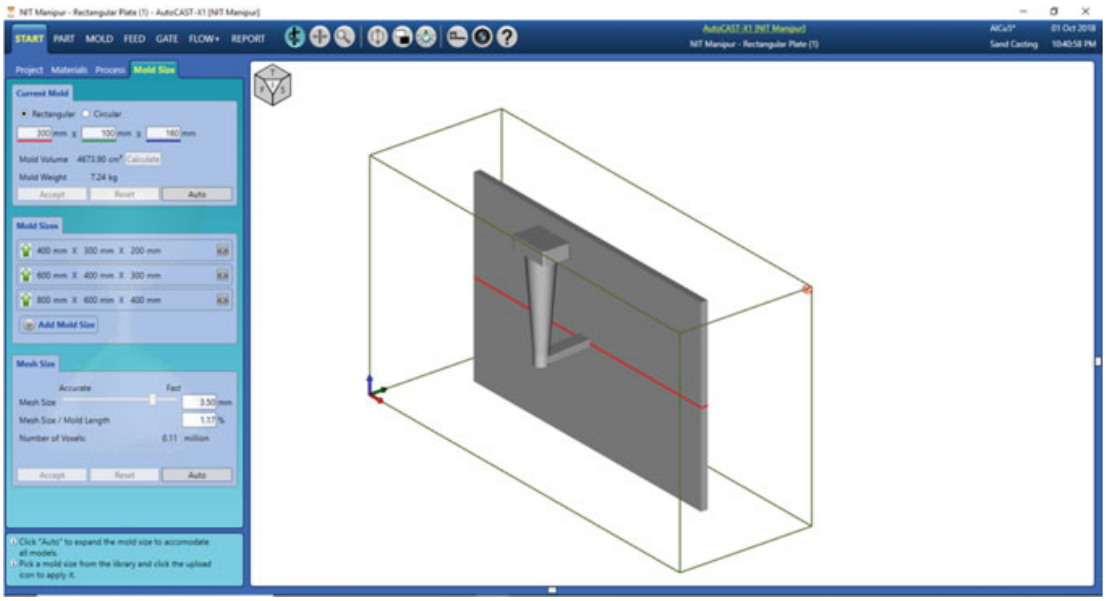


Fig. 41.3 Casting part with mold box

Table 41.2 Design parameters

Sr. no.	Design parameters	Designed value (mm)	Section shape
1	Pouring basin height, length, width	10, 30, 20	Rectangle
2	Diameter of sprue at top, bottom, and sprue length	20, 10	Cylindrical
3	Sprue well: diameter and height	10, 10	Cylindrical
4	Ingrate dimensions: width, height, and length	10, 5, 35	Rectangle

Table 41.3 Material properties

S. no.	Description	Designated value	Units
1	Density	2800	Kg/m^3
2	Thermal conductivity	121	W/m-K
3	Specific heat	920	J/Kg-K
4	Solidus temperature	570	°C
5	Liquidus temperature	650	°C

and sprue well, runner, and ingates. They act as passageways for the flow of molten metal from the ladle to various portions of the mold cavity. Design dimensions are obtained for gating system, and the feeder is mentioned in Table 41.2. The material properties, as mentioned in Table 41.3, restricted to simulation software which is considered as boundary conditions to casting simulation.

41.2.3 Numerical Simulation Using AutoCAST Software

Simulation has been performed in AutoCAST environment according to the design dimensions obtained for a pattern with allowances and gating system. Casting simulation enables to check the part-process compatibility, parting plane and moldability index, and feed metal paths–directional solidification (thin to thicker regions) and to predict the location of shrinkage defects such as porosity and cracks.

Simulation-based trials do not involve wastage of material, energy, and labor and do not hold up regular production. Computer simulation provides a clear understanding of the casting phenomena to identify the location and extent of internal defects, ensuring defect-free castings. Thus, numerical simulation of casting can be considered as an important method to make casting technique change from experience test to science guidance. It requires a standard **.stl** file of the methods layout. However, the application of proper boundary conditions leads to accurate results of the simulation.

Part and Mold box

The prerequisite of this software is to create the part model in CAD software and save it as a standard **.stl** file for importing in AutoCAST. The mold box dimensions have been taken as 300 × 100 × 160 mm as shown in Fig. 41.3. The entire mold containing the casting is automatically subdivided into cubic elements for internal computations such as thickness, solidification, and mold filling. The element size mesh was considered as 3.5 mm.

Simulation with different pouring temperatures

The simulation at different pouring temperatures of 700, 750, and 800 °C was performed. It was observed that the fill time and solidification time increased with respect to the increasing temperature, and the volume of fill is high at higher pouring temperature. Casting simulation between Figs. 41.4 and 41.12 indicates the fill time, solidification time, and percentage of the volume of fill with respect to the altered pouring temperatures, i.e., 700, 750, and 800 °C. In this process, molten metal flow may be visualized at different cross sections of the mold cavity. It enables to visualize the flow of molten metal regions and solidification regions in the casting process (Figs. 41.4, 41.5, 41.6, 41.7, 41.8, 41.9, 41.10, 41.11, and 41.12).

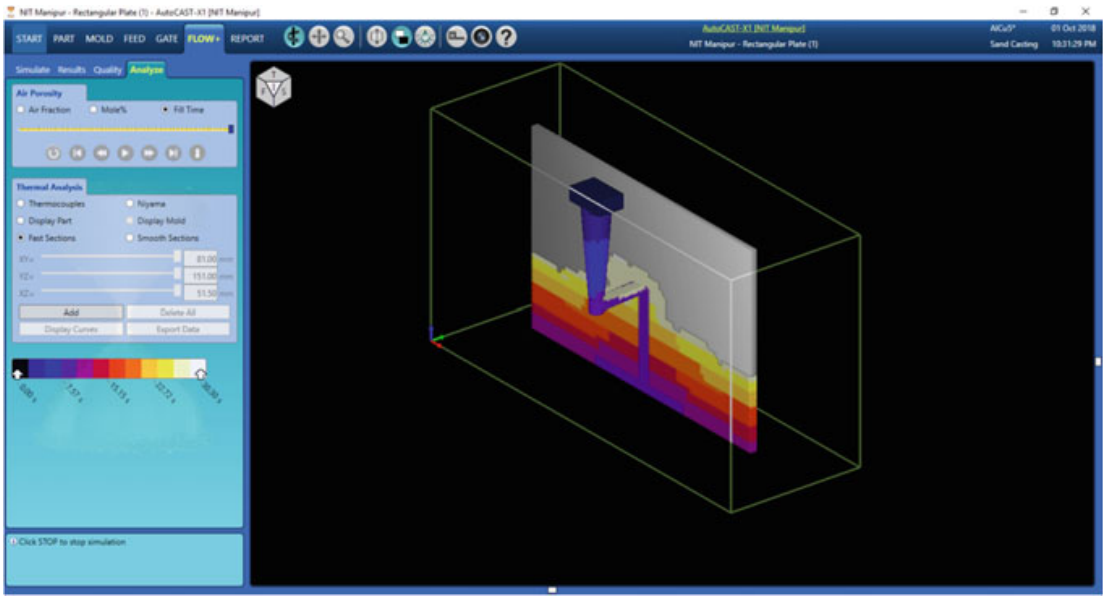


Fig. 41.4 Fill time with the pouring temperature of 700 °C

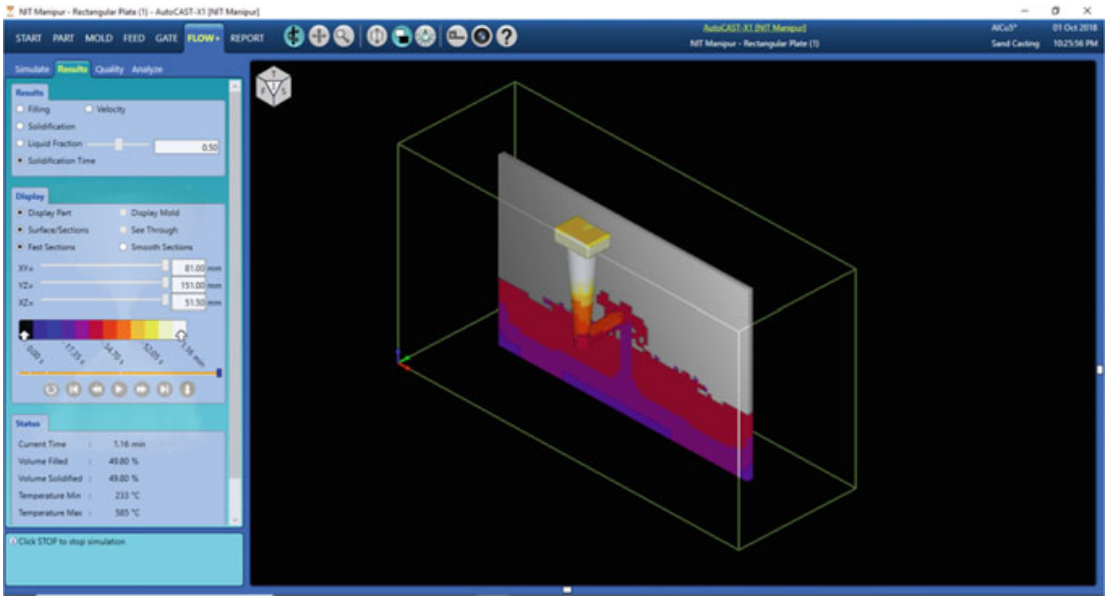


Fig. 41.5 Solidification time with the pouring temperature of 700 °C

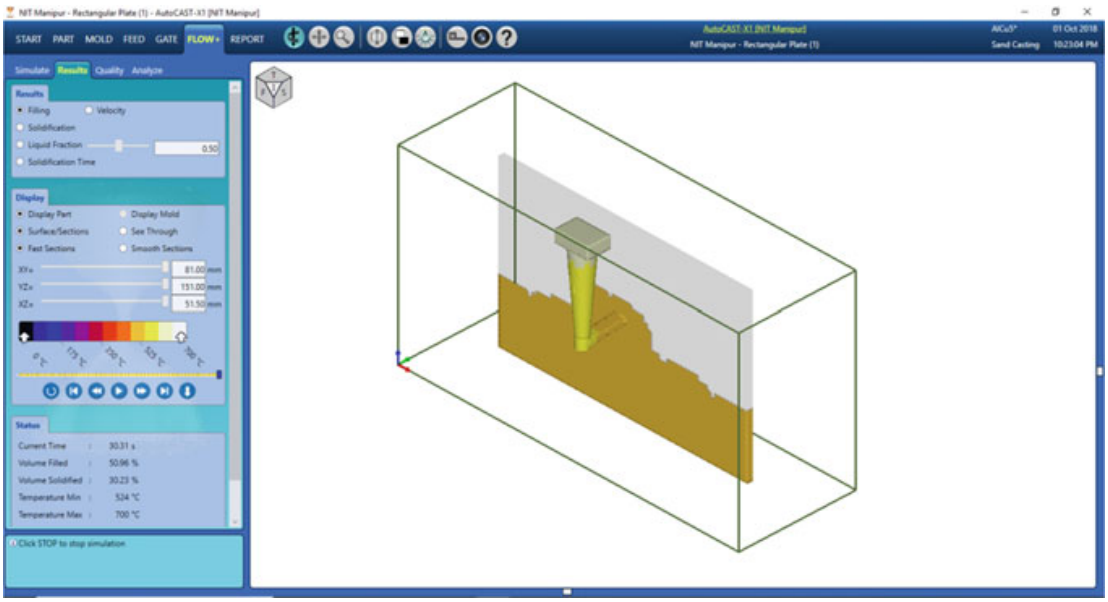


Fig. 41.6 Volume of fill with the pouring temperature of 700 °C

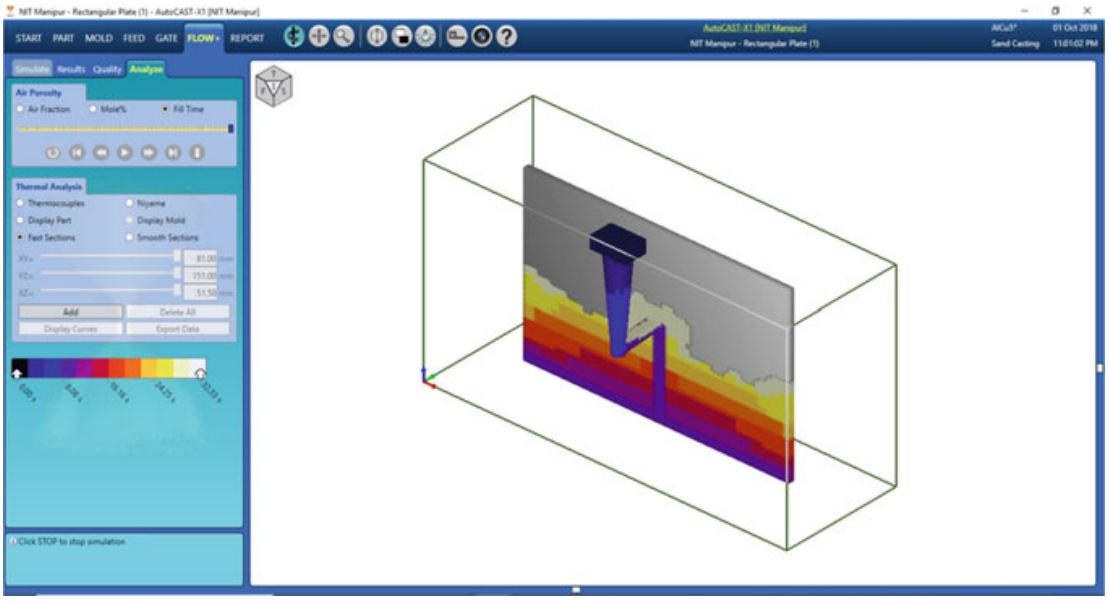


Fig. 41.7 Fill time with the pouring temperature of 750 °C

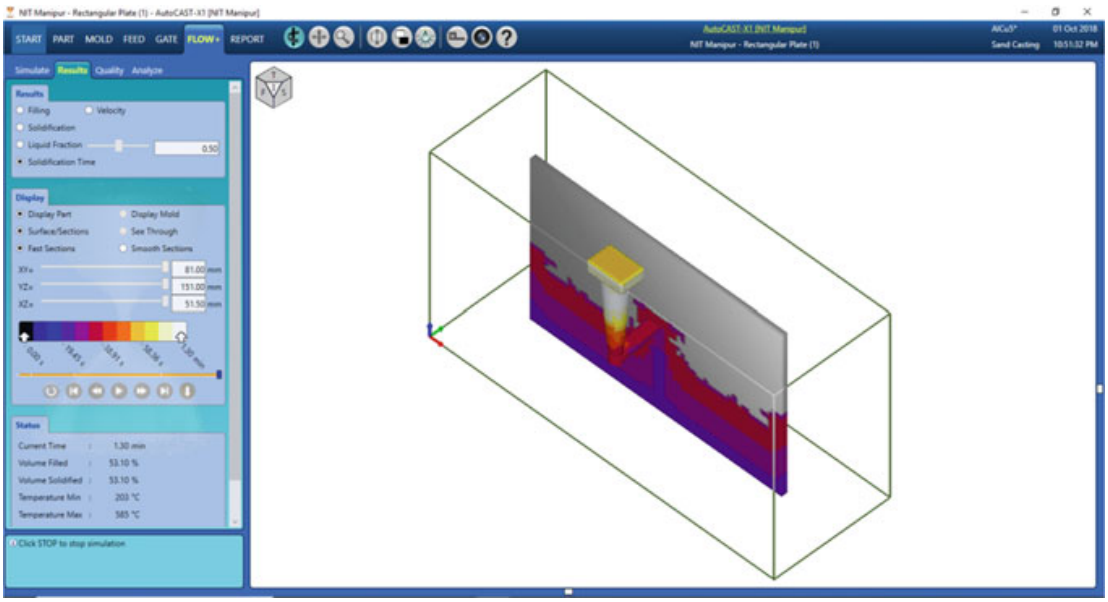


Fig. 41.8 Solidification time with the pouring temperature of 750 °C

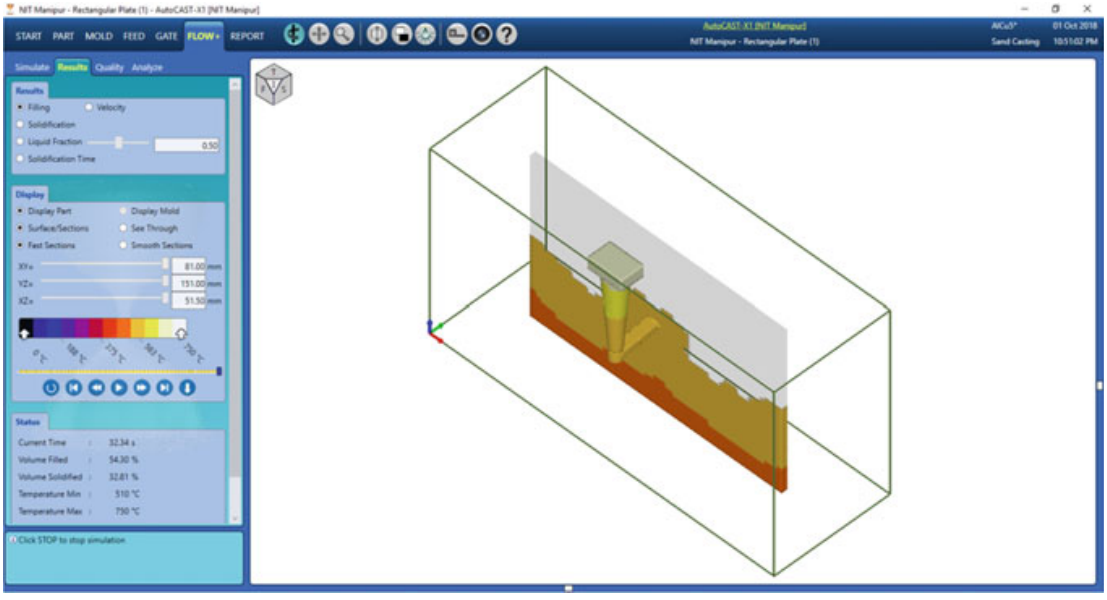


Fig. 41.9 Volume of fill with the pouring temperature of 750 °C

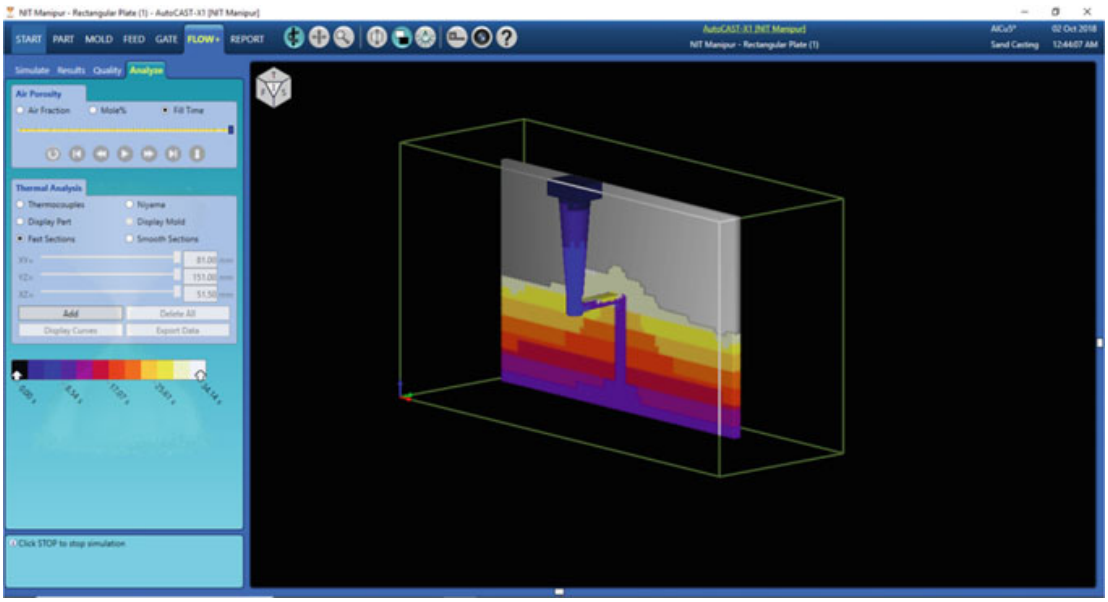


Fig. 41.10 Fill time with the pouring temperature of 800 °C

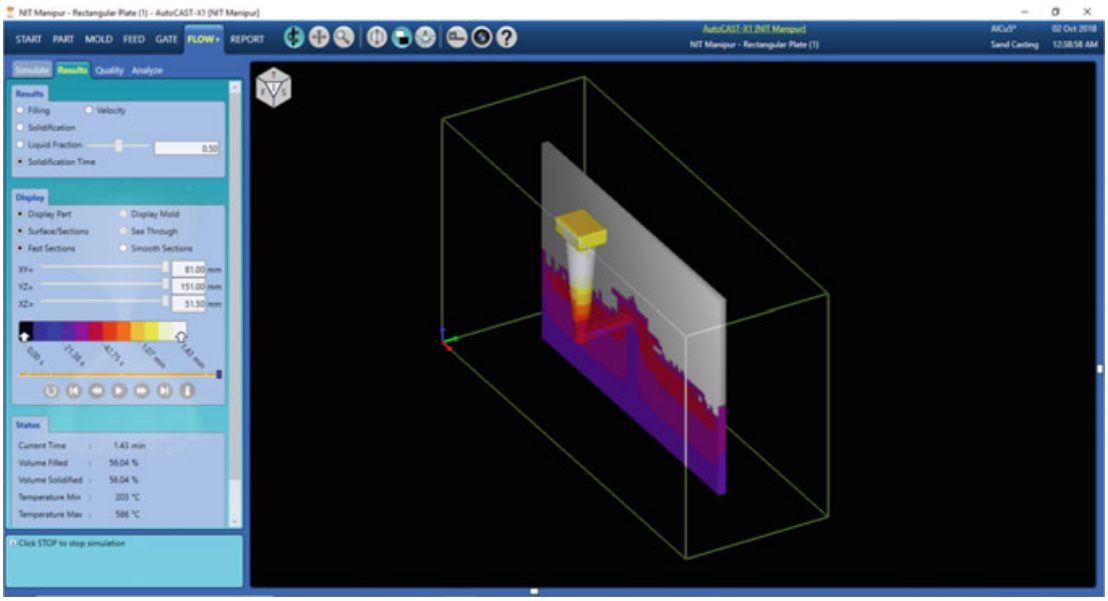


Fig. 41.11 Solidification time with the pouring temperature of 800 °C

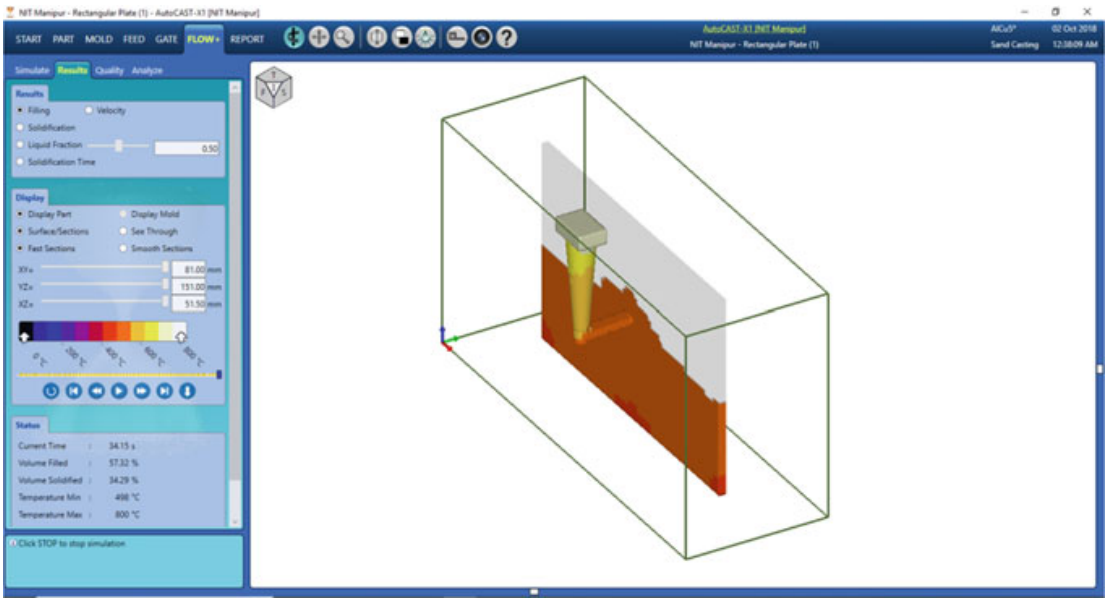


Fig. 41.12 Volume of fill with the pouring temperature of 800 °C

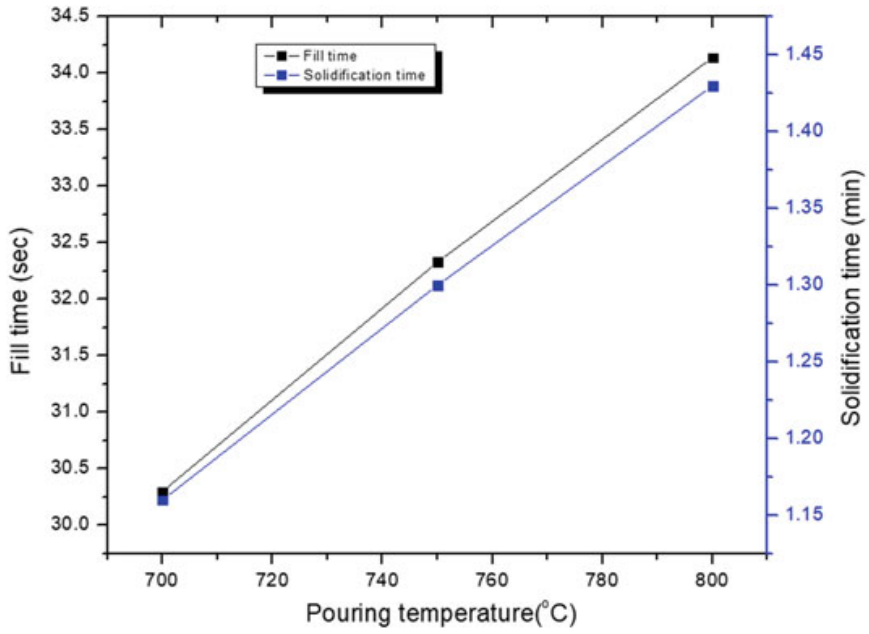


Fig. 41.13 Fill time and solidification time with altered pouring temperatures

41.3 Results and Discussion

Casting simulation enhanced the fluidity of the molten metal with altered pouring temperatures. In Fig. 41.13, it was observed that with the increasing pouring temperature, the fill time and solidification time were increased. In Fig. 41.14, it was observed that when the fill time increased, then the percentage of the volume of fill in the cavity increased. In Fig. 41.15, it was observed that the percentage of the

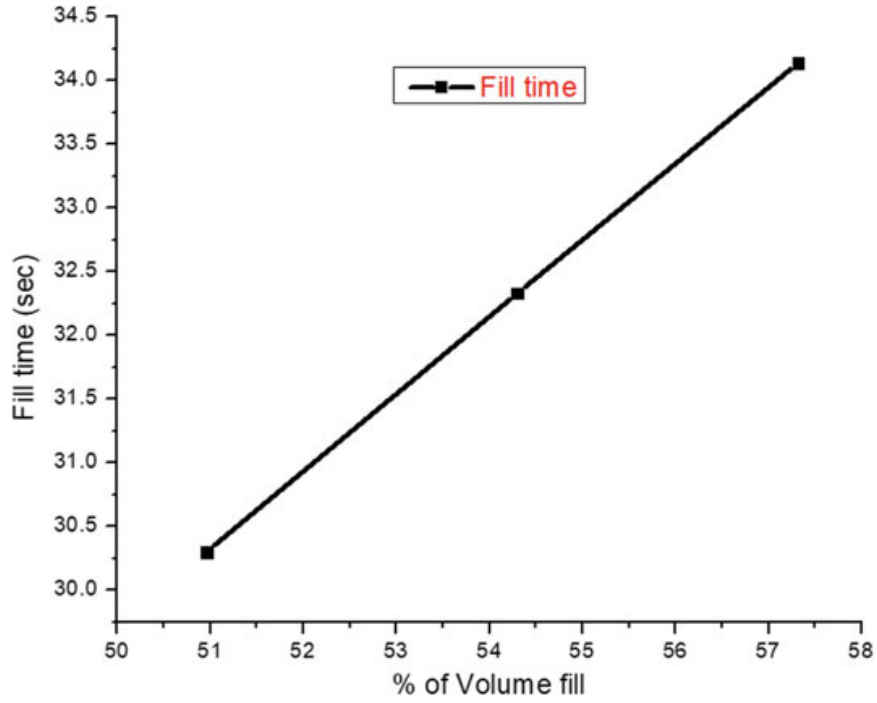


Fig. 41.14 Percentage of volume of fill with respect to fill time

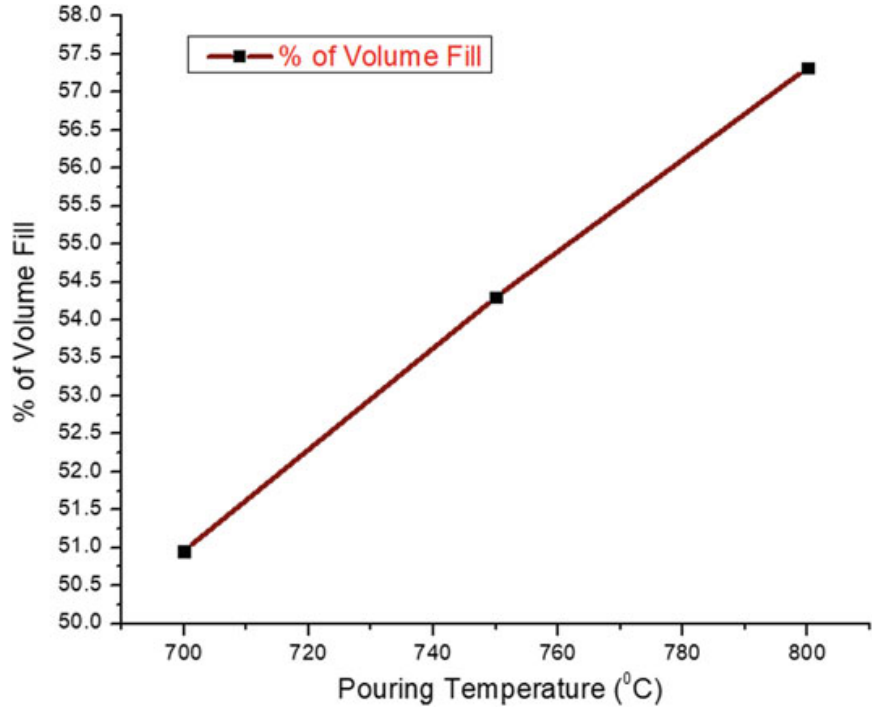


Fig. 41.15 Percentage of volume fill with the altered pouring temperatures

volume of fill increased with the increased pouring temperature. Hence, it was observed that the fluidity increased with the increasing temperature and fill time according as shown in Figs. 41.13, 41.14, and 41.15. Hence, it was observed that fluidity increased with increased pouring temperature.

41.4 Conclusion

In this present investigation, it was observed that solidification simulation enables visualization of freezing inside a casting and identification of the last freezing regions.

1. The effect of pouring temperature enabled the filling time and solidification time; it was observed at with the increasing pouring temperatures, i.e., 700 to 800 °C, the fill time and solidification time were increased.
2. It was observed that the volume of fill increased with increasing temperature and with the increased pouring temperature, the fill time and percentage of the volume of fill increased.
3. In casting simulation, it was observed that the effect of pouring temperature on the fluidity of the Al–Cu cast alloy was increased with the increasing temperature. It was found that the volume of fill increased with the increasing pouring temperature and also strong agreement was observed with the solidification time and fill time.
4. It was concluded that fluidity of Al–Cu cast alloy might be increased with increasing temperature. Hence, it was recommended that the industries have might implement the simulation before manufacturing to economize the product to reach all the levels of the customers.

Acknowledgements Authors are much thankful for the TEQIP-III, NPIU-New Delhi, for financial support for the purchase of AutoCAST-X software to carry out the simulation of designed casting.

References

1. Birru, A.K., Karunakar, D.B.: Fluidity of A713 Cast Alloy with and without scrap addition using double spiral fluidity test: a comparison. Int. J. Mech. Aerosp. Ind. Mechatronic Manufacturing Eng. **6**(1) (2012)
2. Mohan Krishna, D., Reddy, G.S.: Simulation of fluidity in aluminum alloys, super alloy cmsx4, and ductile cast iron. In: 5th International & 26th All India Manufacturing Technology, Design and Research Conference (AIMTDR 2014) December 12th–14th, IIT Guwahati, Assam, India (2014)
3. Joshi, D., Ravi, B.: Early castability evaluation using analytical hierarchy process Int. J. Adv. Manuf. Technol. **50**, 21–36 (2010)
4. Manoj, K., Vaibhav, J., Jitendra, P., Qureshi, F.: Filling analysis of spiral fluidity standard using ProCast simulation. Int. J. Eng. Trends Technol. (IJETT) **46**(8) (2017)
5. Ravi, B., Joshi, D., Kamalesh Singh, D.: Part, tooling and method optimisation driven by castability analysis and cost model. In: 68th WFC—World Foundry Congress 7th–10th February, pp. 261–266 (2008)
6. Sutaria, M., Joshi, D., Ravi, B.: Feeding system design and evaluation using temperature gradient (Feed Path) Maps. American Foundry Society (2014)

7. Ravi, B., Joshi, D.: Feed ability analysis and optimisation driven by casting simulation. Technical Paper submitted to the Indian Foundry Journal (2017)
8. Naveen, H., Sunil, M.: Casting simulation for sand casting of flywheel. IOSR J. Mech. Civil Eng. (IOSR-JMCE), **11**, 37–41 (2014)
9. Birru, A.K., Karunakar, D.B.: A study on hot tearing susceptibility of Al–Cu, Al–Mg, and Al–Zn alloys. Trans. Indian Inst. Met. **65**(1), 97–105 (2012, February)

Part V
Modelling and Analysis

Chapter 42
Analysis and Optimization of Tool Wear Rate in Magnetic Field-Assisted Powder-Mixed Electrical Discharge Machining of Al6061 Alloy Using TLBO

Arun Kumar Rouniyar and Pragya Shandilya

Abstract This paper presents the optimization and analysis of machining parameters on machining of Al6061 alloy using magnetic field-assisted powder-mixed electrical discharge machining (MFAPM-EDM) process. For performance analysis, peak current (I_P), spark on time (S_{ON}), spark off time (S_{OFF}), powder concentration (PC), and magnetic field (MF) are considered as machining parameters and tool wear rate (TWR) as performance measures. The experimental design based on Box Behnken Design is used for conducting the experiments and single objective optimization is performed using teaching–learning-based optimization (TLBO) algorithm. Peak current is observed as the most significant parameters followed by spark on time, magnetic field, and powder concentration. Model to predict TWR is developed in terms of machining parameters. The optimum set of machining parameters for minimum TWR using TLBO algorithm is obtained at I_P-1 *A*, S_{ON}-30 µs, S_{OFF}-51 µs, PC-10 g/l, and MF-0.45 T.

Keywords MFAPM-EDM · Al6061 alloy · Tool wear rate · TLBO

42.1 Introduction

Aluminum 6061 alloy is light and medium strength alloy with magnesium and silicon as principal element other than aluminum. Al6061 alloy is used in the field of aerospace, automobile, marine industries, and architectural and structural applications. This is mainly due to specific properties such as lightweight, hard, strong and toughness, high corrosion resistance, better machinability and

A. K. Rouniyar (✉) · P. Shandilya
Department of Mechanical Engineering, Motilal Nehru National Institute of Technology, Allahabad, Uttar Pradesh 211004, India
e-mail: rme1602@mnnit.ac.in

P. Shandilya
e-mail: pragya20@mnnit.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_42

formability [1, 2]. Conventional machining of Al6061 possesses severe difficulties such as built of edge formation, adhesion between tool and workpiece, poor surface quality, dimensional inaccuracy and burr formation [3, 4]. Thus, to mitigate the above difficulties, the researchers have opted for the non-conventional machining process.

Among non-convectional machining process, electrical discharge machining (EDM) was found suitable for machining intricate profiles on difficult to machine materials irrespective of material hardness with high degree of automation [2, 5]. Material removal mechanism in EDM mainly occurs by melting and evaporation. EDM process was mainly used for tool and die making but at the current scenario, applications are extended to aerospace, military, automobile, and surgical industries. However, the machining with EDM at large scale has certain limitations such as low machining efficiency, poor surface finish, prone to heat and residual stress generated on workpiece material during machining [6–8]. Thus to overcome such problem, hybrid machining process known as magnetic field-assisted powder-mixed EDM (MFAPM-EDM) was developed which can machine complex profiles of conductive materials with higher efficiency and better surface quality. In this process, removal of material from the surface of the workpiece is assisted by powder particle mixed in a dielectric fluid in the presence of magnetic field. The material removal occurs by multiple discharges due to bridging of powder particle under the sparking area instead of single spark combined with Lorentz force developed due to the magnetic field-coupled electric field which assists in removing material. The Lorentz force changes the path of electrolytic ions by reducing the mean free path shifting of debris from the workpiece and tool in the discharge zone as well as accounts for increasing the electron density. Thus, this Lorentz force has a significant impact on the plasma channel which effectively expelling the debris particles leading to increase erosion volume and maintaining uniform discharge waves by gap causing improvement in the MRR, TWR, and better surface integrity [9, 10]. For example, Bhatt et al. [9] machined AISI D3, H13, and AISI D3 steel with tungsten and titanium powder mixed in EDM oil under the influence of magnetic field to examine the effect of process parameters on MRR, overcut, surface roughness, TWR, and microhardness. Krishna Kant [11] used *U*-shaped copper tool to machine EN-19 steel with graphite powder in the dielectric fluid in assisted with a magnetic field to study the effect of process parameters on overcut, MRR and surface roughness. Bains et al. [10] carried out machining of Al–Si MMCs with silicon carbide powder-mixed dielectric fluid under the influence of fixed magnetic field in EDM. Under the influence of the magnetic field at high spark energy, increase in material erosion rate and decrease in recast layer thickness, surface roughness, and microhardness were observed. Hienz et al. [12] performed machining on Ti-6Al-4V alloy under the influence of magnetic field to study the influence on erosion efficiency, plasma temperature, and debris field characterization.

After epitomized scrutiny of literature review, most of the literatures are available in magnetic assisted EDM or powder-mixed EDM process. Few literatures available on machining the difficult to cut conductive materials using

MFAPM-EDM process were improvement in MRR, TWR, and surface integrity was observed. In this present experimental work, MFAPM-EDM process was used to machine Al6061 alloy with aluminum powder mixed in EDM oil under the influence of magnetic field. The effect of machining parameters, namely peak current (I_P), spark on time (S_{ON}), spark off time (S_{OFF}), powder concentration (PC), and magnetic field (MF) on tool wear rate (TWR) were studied. Response surface model to predict TWR was developed, and single objective optimization was performed using teaching–learning-based optimization (TLBO) to determine the optimum machining parameters.

42.2 Experimental Setup and Methods

The experiments were performed on modified conventional die-sinking machine (EMS 5030, Elektra) with fabricated attachment equipped to it as shown in Fig. 42.1. The machine is supplied with ELECTRONIA PRS-20 which can operate peak current up to 35 A. Proper stirring system is designed to solve the problem of powder settling, and a magnetic system is attached for the quick removal of debris for the machining area. Al6061 alloy of dimension 40 mm × 40 mm × 10 mm was machined using 12-mm-diameter electrolytic copper as a tool, aluminum powder (mesh size-200), grade 5 EDM oil as dielectric fluid, and neodymium ring magnet. The machining parameters considered for the experimental work were peak current, spark on time, spark off time, powder concentration, and magnetic field. Box Behnken design (BBD)-based response surface methodology (RSM) was used for designing the experiments [13]. The level and range of machining parameters were decided based on literature review, one parameter at a time and specifications of machine setup. Table 42.1 shows the range and level of experimental machining conditions. For TWR, measurement of the weight of tool before and after

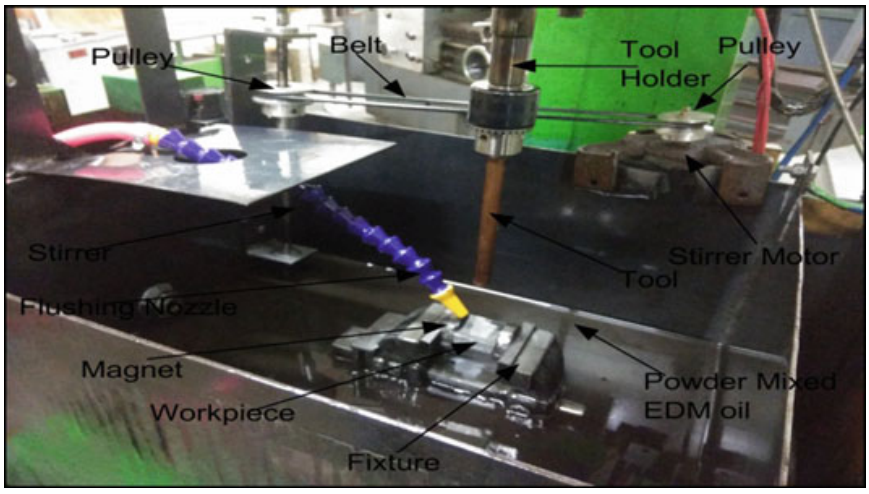


Fig. 42.1 Pictorial view of experimental setup for MFAPM-EDM

Table 42.1 Range and levels of experimental machining conditions

Sl. no.	Machining Parameters	Symbol	Levels		
			1	2	3
1	Peak current (A)	I_P	1	4	7
2	Spark on time (μs)	S_{ON}	30	50	70
3	Spark off time (μs)	S_{OFF}	35	43	51
4	Powder concentration (g/l)	PC	6	8	10
5	Magnetic field (T)	MF	0.15	0.30	0.45

machining was measured using digital precision pan balance (Model- SI-215D, Denver Instruments) with a maximum capacity of 210 grams. TWR was calculated using Eq. (42.1) [14]. The magnetic field was measured using digital gauss meter.

$$\text{TWR}\,(\text{g/min}) = \frac{T_{\text{bm}} - T_{\text{am}}}{t} \tag{42.1}$$

where

T_{bm} Tool weight before machining (g)
T_{am} Tool weight after machining (g)
t Machining time = 10 min.

42.3 Optimization Using TLBO

For this experimental work, teaching–learning-based optimization (TLBO) was employed. TLBO is a single as well as multi-objective optimizing algorithm to obtain solutions for discrete and continuous data problem. TLBO was coined by R. V. Rao for higher consistency which defines the teaching–learning capability of teachers and learners in the given classroom. Two basic modes, i.e., teacher phase and learner phase, are present in this TLBO algorithm. During teacher phase, the teacher tries to give knowledge to the learners and tries to improve the mean results of the classroom. While during learner phase, they try to interact among themselves to increase the knowledge [15]. The following steps are involved to find the optimum solution using TLBO algorithm.

(a) Design variables and its range, objective function, population size, and the number of iterations are defined.
(b) A table containing design variables value, objective function with population size are prepared and mean of all design variables are calculated.
(c) Based on objective function value, the teacher values are calculated.
(d) Other variable values are updated based on the teacher values.

(e) Comparison between new values and corresponding old values of an objective function are performed and the table is updated.
(f) Random values are selected and compared among the learners to improve the values.
(g) Comparison between new values and previous old values of the objective function are performed and the table is updated.
(h) Iterations are performed up to termination criteria.

42.4 Results and Discussion

The BBD design along with experimental results of TWR is shown in Table 42.2. Forty six different experimental combinations run are performed on machining of an Al6061 alloy by the MFAPM-EDM process.

Table 42.2 BBD design with experimental results

STD	I_P	S_{ON}	S_{OFF}	PC	MF	TWR (g/min)
1	1	30	43	8	0.3	0.0013897
2	7	30	43	8	0.3	0.0019722
3	1	70	43	8	0.3	0.0017885
4	7	70	43	8	0.3	0.0023923
5	4	50	35	6	0.3	0.0023388
6	4	50	51	6	0.3	0.0022273
7	4	50	35	10	0.3	0.0021488
8	4	50	51	10	0.3	0.0019249
9	4	30	43	8	0.15	0.0020373
10	4	70	43	8	0.15	0.0022568
11	4	30	43	8	0.45	0.0014504
12	4	70	43	8	0.45	0.0021381
13	1	50	35	8	0.3	0.0017269
14	7	50	35	8	0.3	0.0022719
15	1	50	51	8	0.3	0.0015414
16	7	50	51	8	0.3	0.0021310
17	4	50	43	6	0.15	0.0024057
18	4	50	43	10	0.15	0.0022345
19	4	50	43	6	0.45	0.0022559
20	4	50	43	10	0.45	0.0017867
21	4	30	35	8	0.3	0.0017430
22	4	70	35	8	0.3	0.0022585

(continued)

Table 42.2 (continued)

STD	I_P	S_{ON}	S_{OFF}	PC	MF	TWR (g/min)
23	4	30	51	8	0.3	0.0016038
24	4	70	51	8	0.3	0.0020801
25	1	50	43	6	0.3	0.0018884
26	7	50	43	6	0.3	0.0025449
27	1	50	43	10	0.3	0.0015708
28	7	50	43	10	0.3	0.0021372
29	4	50	35	8	0.15	0.0023692
30	4	50	51	8	0.15	0.0021542
31	4	50	35	8	0.45	0.0019972
32	4	50	51	8	0.45	0.0018848
33	1	50	43	8	0.15	0.0017885
34	7	50	43	8	0.15	0.0023174
35	1	50	43	8	0.45	0.0016011
36	7	50	43	8	0.45	0.0020561
37	4	30	43	6	0.3	0.0018072
38	4	70	43	6	0.3	0.0022951
39	4	30	43	10	0.3	0.0015976
40	4	70	43	10	0.3	0.0020364
41	4	50	43	8	0.3	0.0020480
42	4	50	43	8	0.3	0.0019392
43	4	50	43	8	0.3	0.0019365
44	4	50	43	8	0.3	0.0020703
45	4	50	43	8	0.3	0.0022817
46	4	50	43	8	0.3	0.0020436

42.4.1 Analysis of Tool Wear Rate

The calculated TWR for different experimental conditions was considered for the analysis. Table 42.3 describes the analysis of variance (ANOVA) for TWR carried out using design expert *V* 9.0.5 at 95% confidence level to determine the significant parameters. From Table 42.3, the *F*-value of a model was observed to be 64.36 implying the significance of the model. Likewise, I_P, S_{ON}, S_{OFF}, PC, and MF passed the significant test as $p < 0.05$ was observed. I_P was most influential machining parameters followed by S_{ON}, PC, and MF, while the effect of S_{OFF} was least. However, S_{ON}-PC and PC-MF among interactions and for quadratic terms I_P, S_{ON}, and PC was observed to be significant.

The effect of machining parameters depicted in main effect plot is shown in Fig. 42.2a–e. With increase in peak current from 1 to 7 A, increase in TWR was observed from Fig. 42.2a. This is mainly because of the increment in electrical power which melts the material from tool electrode due to high spark energy which is proportional to peak current. However, increase in TWR was observed with

Table 42.3 ANOVA for tool wear rate (TWR)

Source	SS	DOF	MS	*F*-value	*p*-value Prob > *F*
Model	3.348×10^{-6}	10	3.348×10^{-7}	64.36	<0.0001
A-I_P	1.281×10^{-6}	1	1.281×10^{-6}	246.3	<0.0001
B-S_{ON}	8.302×10^{-7}	1	8.302×10^{-7}	159.6	<0.0001
C-S_{OFF}	1.067×10^{-7}	1	1.067×10^{-7}	20.52	<0.0001
D-PC	3.383×10^{-7}	1	3.383×10^{-7}	65.03	<0.0001
E-MF	3.580×10^{-7}	1	3.580×10^{-7}	68.82	<0.0001
BE	5.480×10^{-8}	1	5.480×10^{-8}	10.54	0.0026
DE	2.220×10^{-8}	1	2.220×10^{-8}	4.27	0.0463
A^2	1.212×10^{-7}	1	1.212×10^{-7}	23.31	<0.0001
B^2	1.758×10^{-7}	1	1.758×10^{-7}	33.79	<0.0001
D^2	3.890×10^{-8}	1	3.890×10^{-8}	7.48	0.0097
Residual	1.820×10^{-7}	35	5.201×10^{-9}		
Lack of fit	1.028×10^{-7}	30	3.427×10^{-9}	0.22	0.9970
Pure error	7.924×10^{-8}	5	1.585×10^{-8}	*R*-Sq. = 0.9484, Adj. *R*-Sq. = 0.9337, Pred. *R*-Sq. = 0.9199	
Cor total	3.530×10^{-6}	45			

SS Sum of Square, *MS* Mean Square, *DOF* Degree of Freedom

increase in S_{ON} from 30 to 70 µs as shown in Fig. 42.2b; this is due to an increase in pulse duration for machining that produces high spark energy which causes melting and evaporation of material from tool surface. Also, decrease in TWR was observed from Fig. 42.2c with increase in spark off time from 35 to 51 µs. It is because of less time the spark gets to be in contact with the workpiece material resulting in the decrease in TWR. From Fig. 42.2d, the decrease in TWR was observed with an increase in powder concentration from 6 to 10 g/l. This was a result of the formation of a narrow channel which transfers the heat to the workpiece materials rather than tool after the addition of powder in EDM oil causing low TWR. From Fig. 42.2e, the decrease in TWR, with an increase in MF from 0.15 to 0.45 T, was attributed to the diamagnetic behavior of electrolytic copper which opposed the magnetic field when increased and the debris particles were repelled away causing low TWR.

42.4.2 Response Surface Modeling

A second-order model was developed in terms of significant machining parameters to predict the TWR. The regression coefficient of the model was obtained using experimental data at 95% confidence level. The second-order model for TWR in coded factors is stated in Eq. (42.2).

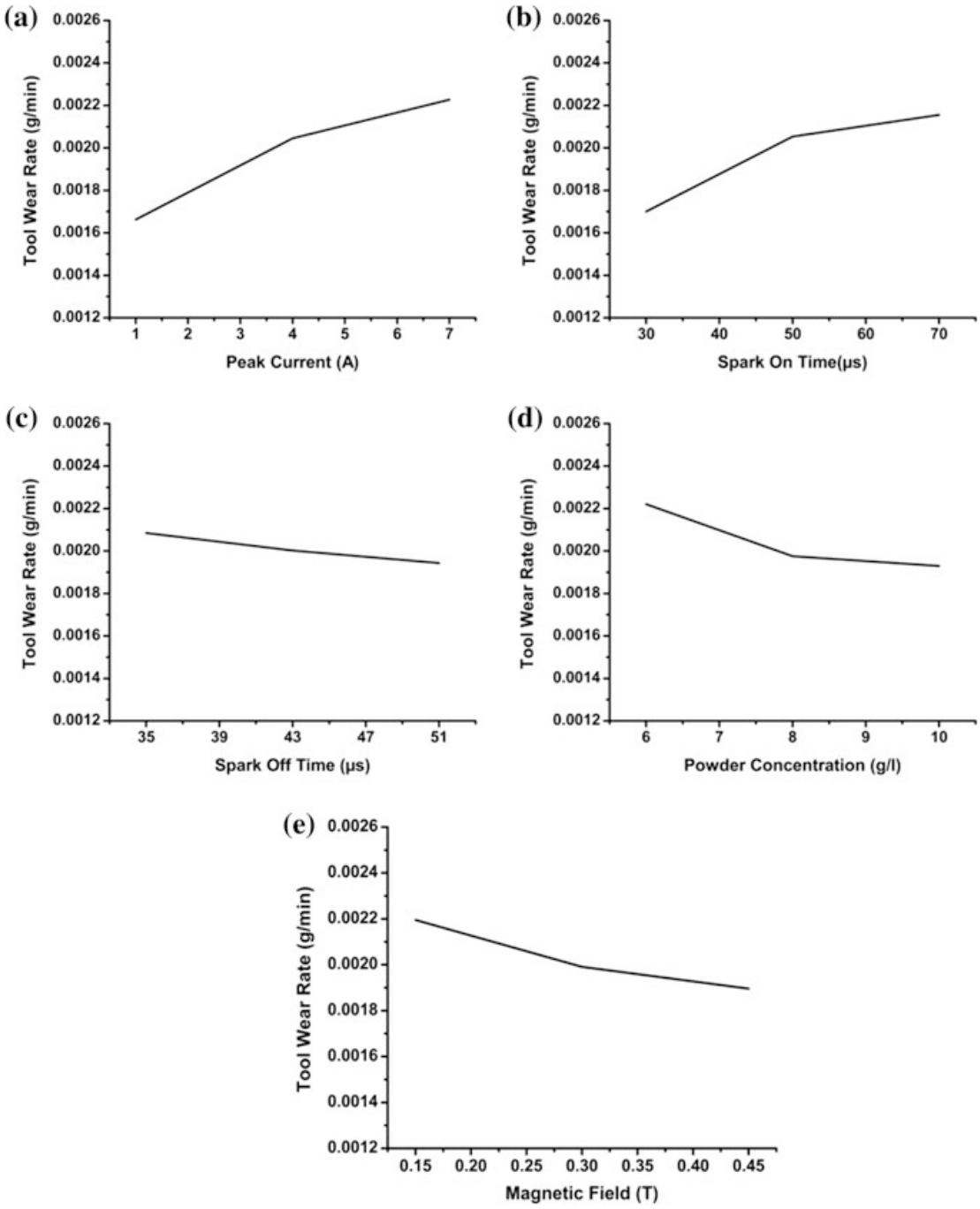


Fig. 42.2 Main effect plot of machining parameters **a** peak current **b** spark on time **c** spark off time **d** powder concentration **e** magnetic field on TWR

$$\begin{aligned}\text{TWR} = &+2.073\times10^{-3}+2.830\times10^{-4}\times A+2.278\times10^{-4}\times B-8.167\times10^{-5}\times C\\ &-1.454\times10^{-4}\times D-1.496\times10^{-4}\times E+1.171\times10^{-4}\times B\times E-7.450\times10^{-5}\times D\times E\\ &-1.108\times10^{-4}\times A^2-1.334\times10^{-4}\times B^2+6.274\times10^{-5}\times D^2\end{aligned} \tag{42.2}$$

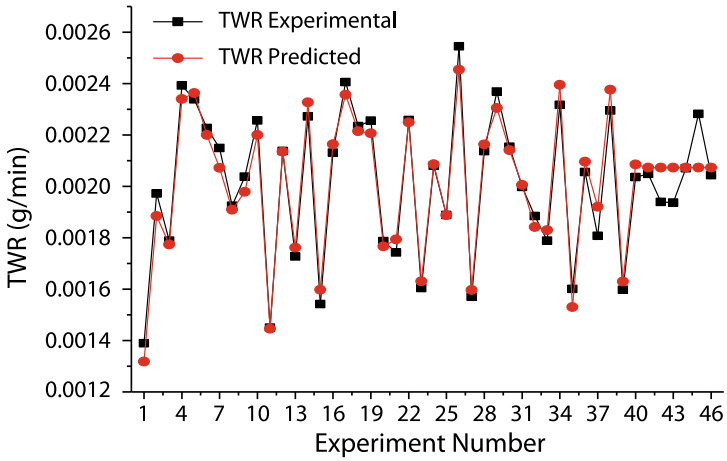


Fig. 42.3 Experimental and predicted TWR against experimental runs

Predicted *R*-Sq. value of 0.9199 for TWR was in reasonable agreement with the adjusted *R*-Sq. value of 0.9337. The second-order model developed was found to be congruent with the experimental data as shown in Fig. 42.3.

42.4.3 Optimization Using TLBO

Response surface model developed and experimental data were used in teacher–learner phase for optimizing the machining parameters. Single objective optimization was carried out to determine the optimum set of machining parameters for TWR. For optimization, the computer code was developed with population size = 46 and number of iterations = 20 using MATLAB R2016a. The optimal solution of machining parameters for TWR with the number of iterations is shown in Fig. 42.4, and convergence graph shows the decreasing curve for TWR with increasing the successive iterations. The optimized value of TWR was obtained at eighth iterations, and the optimum set of machining parameters for minimum TWR (0.000813 gm/min) was obtained at I_P-1 A, S_{ON}-30 µs, S_{OFF}-51 µs, PC-10 g/l, and MF-0.45 T, respectively.

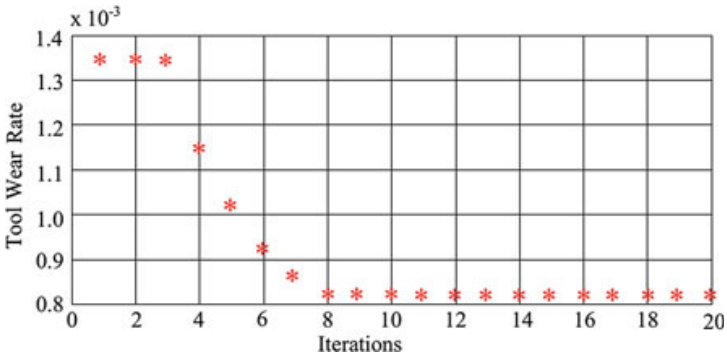


Fig. 42.4 TWR with number of iterations

Table 42.4 Confirmation results for TWR

Input machining parameters					TWR (g/min)	
I_P (A)	S_{ON} (µs)	S_{OFF} (µs)	PC (g/l)	MF (T)	Optimized	Experimental
1	30	51	10	0.45	0.000813	0.000894

42.4.4 Confirmation Test

Once the optimum machining parameters were identified using TLBO algorithm for minimum TWR, the same combination of machining parameters was used for the validation by performing the experiments in MFAPM-EDM process. The results obtained with TLBO and experiments for optimum machining parameters are shown in Table 42.4.

42.5 Conclusions

In this paper, magnetic field-assisted powder-mixed EDM was successfully used for machining Al6061 alloy with aluminum powder mixed in EDM oil. The experimental design was based on BBD, and ANOVA was carried out at 95% confidence level to determine the influential machining parameters for TWR. Modeling in terms of coded machining parameters was carried out using response surface modeling, and the optimized set of machining parameters was obtained using TLBO algorithm. The following conclusions were obtained from this work.

1. Peak current was most influential machining parameters followed by spark on time, magnetic field, and powder concentration while spark off time has least effect.
2. Second-order model to predict the TWR was developed using response surface and was found to be congruent with the experimental results.
3. Using TLBO, optimized set of machining parameters was obtained at eighth iteration for I_P-1 A, S_{ON}-30 µs, S_{OFF}-51 µs, PC-10 g/l, and MF-0.45 T for minimum TWR.

References

1. Kumar, R., Pradhan, M.K., Kumar, R.: Modeling and optimization of milling parameters on Al-60161 alloy using multi-objective genetic algorithm. In: 5th International and 26th All India Manufacturing Technology, Design and Research Conference, vol. 89, pp. 1–7 (2014)
2. Arooj, S., Shah, M., Sadiq, S., Jaffery, S.H.I., Khushnood, S.: Effect of current in the EDM machining of aluminum 6061 T6 and its effect on the surface morphology. Arabian J. Sci. Eng. **39**(5), 4187–4199 (2014)

3. Sanchez, J.M., Rubio, E., Alvarez, M., Sebastian, M.A., Marcos, M.: Microstructural characterisation of material adhered over cutting tool in the dry machining of aerospace aluminium alloys. J. Mater. Process. Technol. **164–165**, 911–918 (2005)
4. Roy, P., Sarangi, S.K., Ghosh, A., Chattopadhyay, A.K.: Machinability study of pure aluminium and Al–12% Si alloys against uncoated and coated carbide inserts. Int. J. Refract Metal Hard Mater. **27**, 535–544 (2009)
5. Bisaria, H., Shandilya, P.: The machining characteristics and surface integrity of Ni-rich NiTi shape memory alloy using wire electric discharge machining. In: Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science (2018). https://doi.org/10.1177/0954406218763477
6. Pramanik, A., Basak, A.K., Islam, M.N., Littlefair, G.: Electrical discharge machining of 6061 aluminium alloy. Trans. Nonferrous Metals Soc. China **25**(9), 2866–2874 (2015)
7. Bisaria, H., Shandilya, P.: Experimental studies on electrical discharge wire cutting performance of Ni-rich NiTi shape memory alloy. Mater. Manuf. Process. **33**(9), 977–985 (2017)
8. Bisaria, H., Shandilya, P.: Machining of metal matrix composites by EDM and its variants: a Review. Chapter 23 in DAAAM International Scientific Book, pp. 267–282, ISSN 1726-9687. Vienna, Austria (2015)
9. Bhatt, G., Batish, A., Bhattacharya, A.: Experimental investigation of magnetic field assisted powder mixed electric discharge machining. Part. Sci. Technol. **33**, 1–28 (2015)
10. Bains, P.S., Sidhu, S.S., Payal, H.S., Kaur, S.: Magnetic Field Influence on Surface Modifications in Powder Mixed EDM. Silicon (2018). https://doi.org/10.1007/s12633-018-9907-z
11. Kant, K.: Experimental investigation of magnetic field on EDM process by using Taguchi method on EN-19 tool steel. Global J. Res. Eng. Mech. Mech. **15**(3), 27–35 (2015)
12. Heinz, K., Kapoor, S.G., Devor, R.E., Surla, V.: An investigation of magnetic-field-assisted material removal in micro-EDM for nonmagnetic materials. J. Manuf. Sci. Eng. **133**, 021002-1-9 (2011)
13. Montgomery, D.C.: Design and analysis of experiments, 4th edn. Wiley, New York (1997)
14. Prasanna, J., Rajamanickam, S., Kumar, O.A., Raj, G.K., Narayanan, P.V.V.S.: MRR and TWR evaluation on electrical discharge machining of Ti-6Al-4V using tungsten: copper composite electrode. In: IOP Conference Series: Materials Science and Engineering, vol. 197, pp. 012087-1-7 (2017)
15. Rao, R.V.: Teaching learning based optimization algorithm and its engineering applications. Springer (2016). https://doi.org/10.1007/978-3-319-22732-0

Chapter 43
ANFIS-Based Subtractive Clustering Algorithm for Prediction of Response Parameters in WEDM of EN-31

Kunal Chopra, Aishwarya Payla, Guneet Kaur and Eswara Krishna Mussada

Abstract Wire electrode discharge machining (WEDM) is an accurate but an expensive and time-consuming process. In order to establish a stable connection between input and output variables, implementation of soft computing techniques can be useful. Therefore, the current study focuses on comparing adaptive neuro-fuzzy inference system (ANFIS)-based subtractive clustering algorithm with numerous input combinations as well as multivariate regression models in order to simulate and map the output variables with the process parameters used during experimentations, namely pulse-on time (T_{on}), servo voltage (S_v), wire feed (W_f), and wire tension (W_t). Results show that ANFIS models have the ability to estimate the edge roughness (E_r) and kerf width (K_w) more accurately with 96.2 and 97.3% accuracy. ANFIS model is more reliable, accurate, and productive as it uses the learning of neural networks to predict. Also, the developed model has been used to study and explain the effect of various input variables upon the quality of machining. High pulse-on time directly decreases the quality increasing the edge roughness and kerf width which are both undesirable. Low wire feed has shown to decrease both the response parameters regardless of other input parameters. Wire tension has shown much less significant effect as compared to the other three variables.

Keywords WEDM · ANFIS · Machining · EN-31 · Artificial Intelligence

K. Chopra · A. Payla · G. Kaur
SOET, BML Munjal University, Gurgaon 122413, India
e-mail: kunalchopra954@gmail.com

A. Payla
e-mail: aishwaryapayla21@gmail.com

G. Kaur
e-mail: guneet.kaur31297@gmail.com

E. K. Mussada (✉)
Amity Institute of Technology, Amity University, Noida 201313, Uttar Pradesh, India
e-mail: eswar_mussada@yahoo.com; ekmussada@amity.edu

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_43

43.1 Introduction

Manufacturing industry all around the world is focusing on improving quality even for intricate geometries but not at the expense of high cost. The issue of high forces generated in traditional machining processes has been overcome by the advent of non-conventional machining processes which removed the contact between the tool and workpiece and hence reducing contact and frictional stresses. Wire electrode discharge machining (WEDM) is one of the advanced machining process which uses electro-thermal process for machining conductive materials. A thin metal wire as the electrode removes the material by melting and vaporization of workpiece material [1]. Continuously flowing dielectric fluid ensures the flushing of debris thus formed from the workpiece surface. The spark erosion process used for material removal in WEDM makes it possible to machine even the hard materials. Since there is no contact between tool and workpiece surfaces, friction does not come into play and hence residual stresses generated at the cutting zone are low.

Mahapatra et al. [2] through the study of different input variables including pulse-on time (T_{on}), pulse frequency, discharge current, wire speed, wire tension (W_t) and dielectric flow on the surface finish, and kerf width (K_w) found discharge current, pulse duration, and dielectric flow rate to be the most influential parameters for minimization of surface roughness and kerf width in rough cutting operations. Kumar et al. [3] found I_P (peak current), T_{on}, T_{off}, and W_f to have effect on the surface finish in that order while WEDM of high-speed steel (HSS) through zinc-coated wire. Sharma et al. [4] did the experimental study followed by parameter optimization on WEDM and found out that T_{on} majorly affects the response parameters such as surface roughness (S_r) and C_r. Experimental investigation done by Garg et al. [5] focused on machining Al/10% $ZrO_{2(P)}$ metal matric composite by WEDM. They studied the effect of the input parameters on surface roughness (S_r) and cutting velocity (CV). Machining parameters such as pulse width, short time pulse and time between pulses were found to affect both the response parameters most. Servo control mean reference voltage was observed to have similar significance on affecting the response parameters. Fard et al. [6] machined AlSiC metal matrix composite to study the effect of T_{on}, T_{off}, W_f, W_t, S_v, and discharge current on S_r and cutting speed by applying ANOVA to the experimental results. In his study, T_{on} and discharge current were found to be the most significant factors influencing the studied targets while W_t had the least impact. Parmanik et al. [7] processed duplex stainless steel in WEDM and observed that increasing T_{on} did not have a linear effect on K_w and craters along with resolidified molten were noticed on the machined surface. Tosun et al. [8] analyzed the crater sizes formed due to 0.25 mm diameter brass wire under different cutting parameters. It was concluded that pulse duration and open-circuit voltage have adverse effect on crater [8]. Chopra et al. [9] found low T_{on} and S_v to be the important parameters to minimize edge roughness (E_r) while machining EN-31 through WEDM.

Regression is a statistical approach to establish a relationship between two sets of variables, dependent variables and the independent variables. Regression can further be multivariate regression or multiple regression. Multivariate regression refers to estimating more than one output variable using single regression model and when multivariate regression consists of more than one predictor variable it refers to multivariate multiple regression. Majumder et al. [10] studied the impact of pulse-on time, pulse-off time and wire feed on kerf width (K_w), S_r, and material removal rate (MRR). They developed general regression neural network (GRNN) and multiple regression models for prediction and comparison of studied key machinability aspects. Maher et al. [11] performed experiments to study the impact on CNC WEDM response parameters like cutting speed, surface roughness, and white layer thickness while machining titanium alloy grade 5(Ti6Al4V). Their study involves development of a regression model to study the effect of performance index on the machinability aspects studied. In order to increase the productivity and accuracy of the machining of super alloy, Udimet-L605 done by WEDM, Nain et al. [12] developed linear regression model to understand the relation between input and output parameters followed by application of particle swarm optimization (PSO), concluding T_{on} to be the most impacting factor on C_r.

Artificial neural networks were introduced to simulate the neural structure similar to brain by establishing a relationship between I/O data during training phase which could be recalled during the verification or optimization process [13]. In ANN algorithms, each processing element or neurons can reproduce the biological NN having numerous inputs and a single output considering certain assumptions and constraints. ANN is an effective technique which can approximate with high precision even nonlinear functions [14] but the efficiency is constrained to problem under consideration [15]. Ramakrishnan et al. [16] designed an ANN model using back-propagation algorithm to predict the characteristic parameters, namely MRR and S_r, of CNC WEDM while machining Inconel 718. Optimization was done using multi-response signal-to-noise ratio (MRSN) and analysis of variance (ANOVA) was applied to determine the level of impact of the response parameters.

Adaptive neuro-fuzzy inference system (ANFIS) combines ANN with fuzzy modeling thus capturing benefits of both techniques in one framework only. It provides a means to map the input–output data relations. Yilmaz et al. [17] developed a fuzzy model that allowed precise selection of EDM parameters using if-then. Fuzzification was done by triangular membership function while defuzzification process was then carried out by centroid area method in order to obtain minimum electrode wear and better surface quality. Maher et al. [18] applied ANFIS to predict the value of white layer thickness (WLT) in WEDM and obtained an accuracy of 97.39%. Maher et al. [19] successfully designed ANFIS model for the selection of optimum machining parameters with the objective of achieving higher productivity at the highest possible surface quality and sustainable development at minimum cost.

Present study is conducted with the objective of investigating the effect of four input parameters, pulse-on time, servo voltage, wire tension, and wire feed on the machining performance of die steel EN-31 and to develop both multivariate regression and ANFIS model to predict and improve the output parameters, viz. edge roughness (E_r) and kerf width (K_w). EN-31 alloy steel has a number of applications in manufacturing industry due to its outstanding wear resistance and high load-carrying capacity. There is limited literature available on the prediction and optimization of K_w; however, no study has been carried out for prediction of E_r while machining through WEDM.

43.2 Experimental Work

The experiments were conducted on "Electronica Wire-cut Electric Discharge Machine" which can cut with speeds upto 230 mm^2/min and workpiece of heights up to 250 mm can be held in its clamp. Out of many wires available for cutting operation, brass wire having a diameter of 250 μm was used which provides better integrity of surface while conducting current. Deionized water is a proven dielectric to provide a good conducting sphere of influence while machining and hence was used in our experimentations.

Four machining parameters, namely pulse-on time (T_{on}), servo voltage (S_v), wire tension (W_t), and wire feed (W_f), were selected in this study for investigating the effect on edge roughness and kerf width. These parameters were identified on the basis of literature survey and pilot experimentation to maintain the appropriate spark gap. T_{on} and S_v were scrutinized at three levels while W_f and W_t were scrutinized at two levels on the basis of economic impact of changing the parameters. Duty factor for the machine was kept at 0.7. A general full-factorial design for experimentation was used for studying the effect of each parameter on performance. Table 43.1 shows the control parameters and their levels along with the other parameters that were kept constant throughout the machining operation.

The crests and troughs formed along the edge were the result of craters and irregularities formed due to the rapid melting and re-solidification of material. The accuracy of machined component is a direct measure of kerf width and edge roughness which were measured at several locations along the machined edge by "Olympus optical microscope" as shown in Fig. 43.1 and averaged over all the readings for each experiment. Each machined specimen was cleaned with acetone before observing under the optical microscope.

Table 43.1 Control and constant parameters

Parameters	Level-1 (low)	Level-2 (medium)	Level-3 (high)
T_{on} (µs)	100	110	120
S_v (V)	20	40	60
W_f (m/min)	5	8	–
W_t (g)	660 g	1140 g	–
Wire diameter	250 µm	–	–
Constant machining parameters			
Wire material		Standard brass	
Peak current		12 (fine pulse and 2 power pulses ON)	
Peak voltage		11	
Cutting speed override		25%	
Duty cycle		0.70	
Servo feed		10 mm/min	
Dielectric fluid		De-ionized water	

43.3 Modeling Approach

Regression analysis has been used to map the input and output variables in order to obtain a function which could predict the response measures on the basis of given input values. Since the ranges of our parameters are not consistent and hence weights were required, instead, techniques like feature scaling, normalization, and polynomial feature scaling were applied as a part of data preprocessing step. It is usually done using Min-Max scaling method, which is the simplest method of all and consists of rescaling the range of features in range of [−1, 1]. Equation (43.1) shows the formula used. Platform used for this project is Jupyter Notebook and programming language was Python. Scikit-learn is a popular open-source library used to implement machine learning and perform statistical analysis in Python. It features various classification, regression, and clustering algorithms and is designed to interoperate with the Python numerical and scientific libraries NumPy, SciPy, and Pandas.

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{43.1}$$

After that, polynomial feature transformation was applied for both degree 2 polynomial and degree 3 polynomial. It helped find relationship between different dependent variables and their impact on independent variables, E_r and K_w. Cubic polynomial regression helped improve the R score from 67 to 94% for both of our response variables which is more accurate than linear regression. The function obtained for K_w and E_r are depicted in Eqs. (43.2) and (43.3).

$$\begin{aligned} k_w = {} & (-2.50e-10) + (-192.45)x1 + (-2165.67)x2 + (554.52)x3 + (-7.49)x4 \\ & + (-48.36)x1^2 + (6677.44)x1x2 + (-1715.03)x1x3 + (79.42)x1x4 \\ & + (-754.35)x2^2 + (-725.72)x2x3 + (-233.69)x2x4 + (554.52)x3^2 \\ & + (-159.07)x3x4 + (-7.49)x4^2 + (94.92)x1^3 + (-3708.53)x1^2x2 \\ & + (2081.60)x1^2x3 + (-222.51)x1^2x4 + (-125.49)x1x2^2 + (575.63)x1x2x3 \\ & + (-854.05)x1x2x4 + (-1950.18)x1x3^2 + (895.14)x1x3x4 \\ & + (115.820)x1x4^2 + (439.84)x2^3 + (118.36)x2^2x3 + (-11.80)x2^2x4 \\ & + (-35.93)x2x3^2 + (213.71)x2x3x4 + (605.77)x2x4^2 \\ & + (554.52)x3^3 + (-210.62)x3^2x4 + (-355.05)x3x4^2 + (-7.49)x4^3 \end{aligned} \tag{43.2}$$

$$\begin{aligned} E_r = {} & (6.68409772e-13) + (44.15)x1 + (95.005)x2 + (47.19)x3 + (-10.25)x4 \\ & + (-2.44)x1^2 + (-310.30)x1x2 + (-165.98)x1x3 + (62.73)x1x4 + (36.996)x2^2 \\ & + (59.24)x2x3 + (8.78)x2x4 + (47.19)x3^2 + (-33.66)x3x4 + (-10.25)x4^2 \\ & + (-48.77)x1^3 + (236.87)x1^2x2 + (207.42)x1^2x3 + (-156.19)x1^2x4 \\ & + (-34.74)x1x2^2 + (-127.937)x1x2x3 + (34.075)x1x2x4 \\ & + (-198.43)x1x3^2 + (216.10)x1x3x4 + (65.22)x1x4^2 + (-12.088)x2^3 \\ & + (8.95)x2^2x3 + (13.79)x2^2x4 + (42.10)x2x3^2 + (-21.76)x2x3x4 \\ & + (-28.67)x2x4^2 + (47.19)x3^3 + (-41.62)x3^2x4 + (-55.87)x3x4^2 + (-10.25)x4^3 \end{aligned} \tag{43.3}$$

where $x1$, $x2$, $x3$, and $x4$ are the independent variables (T_{on}, S_v, $W_{f,}$ and $W_{t,}$ respectively) along with the regression coefficients.

ANFIS model exploits the capability of learning by neural network in order to establish a relationship between input–output parameters which uses fuzzy rules. Due to high variability in the input variables, subtractive clustering algorithm was used for fuzzy rule determination process which checks each and every data point for potential cluster center and the point with highest density in its vicinity is considered as the cluster center with readings outside the range of influence of that center being excluded and are again checked for potential of being a cluster center. It is an iterative process that runs until all points lie within the range of influence of a cluster center. Subtractive clustering approximates the function as a generalized bell curve and determine the number of rules and membership functions. Since the number of centers to be created for the distribution of data was not clear, automated algorithm of subtractive clustering and fuzzy rules were used. Similar to feature scaling, the data were first normalized as ranges for our variables are not same. In normalization, mean is involved when standardizing the data as shown in the equation below:

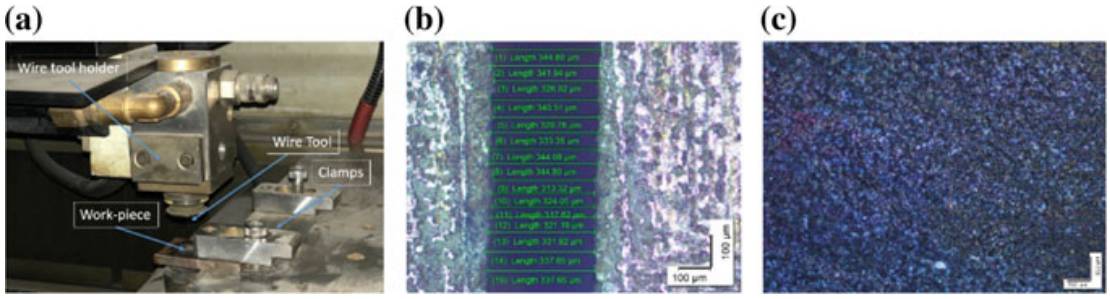


Fig. 43.1 **a** Experimental setup, **b** kerf width measurement, and **c** microscopic image of machined surface at T_{on} = 110 µs, S_v = 60 V

Table 43.2 Information of the trained ANFIS model

Number of nodes	307
Number of linear parameters	150
Number of nonlinear parameters	240
Number of training data pairs (membership functions)	30
Number of fuzzy rules	30

$$x' = \frac{x - \text{average}(x)}{\text{stdev}(x)}. \tag{43.4}$$

The data were then trained using the MATLAB's ANFIS toolkit using subtractive clustering as an FIS technique having 0.5 s the range of influence using the hybrid training algorithm. The structure with four input and single output parameters formed a structure containing 30 input membership functions and fuzzy rules as shown in Fig. 43.1. The membership functions and parameters obtained are shown in Table 43.2. Training error for E_r and K_w came out to be 2.16e–7 and 1.51e–7, respectively. The trained model was used to test the data and predict for randomly selected readings as mentioned in Sect. 43.4.

43.4 Results and Discussion

Effect of machining parameters on E_r and K_w is shown in Figs. 43.2 and 43.3. With the rise in S_v and T_{on} values roughness is found to be increasing. While wire tension at higher S_v values and wire feed has negligible effect on the roughness. Higher values of wire tension reduce the vibration in the wire thus maintaining the spark gap required for the machining process. The reduced vibrations give more interaction time to electrode with the workpiece thus utilizing the energy available for material removal and forming larger number of craters on the surface. At higher gap, voltage energy density is more resulting in more rapid vaporization and hence localized material removal takes place at the cutting edge which leads to higher

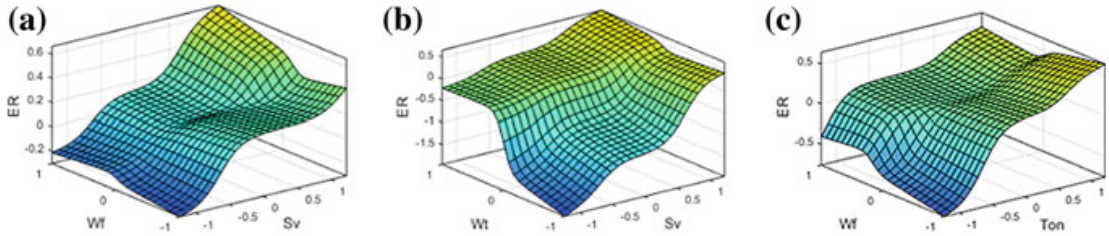


Fig. 43.2 Modeled edge roughness by ANFIS w.r.t **a** wire feed and servo voltage, **b** wire tension and servo voltage, and **c** wire feed and pulse-on time

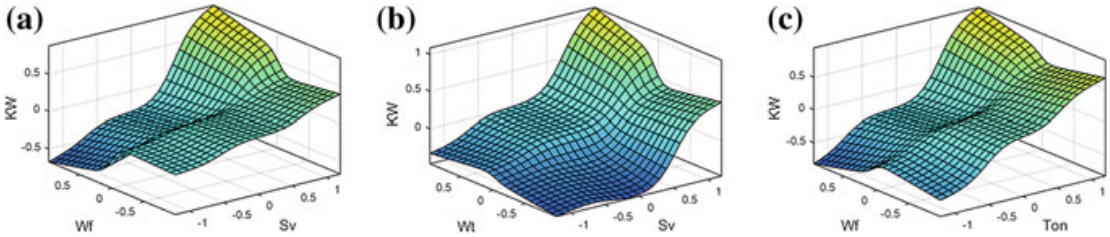


Fig. 43.3 Modeled kerf width by ANFIS in w.r.t. **a** wire feed and servo voltage, **b** wire tension and servo voltage, and **c** wire feed and pulse-on time

roughness. As the pulse duration increases, discharge energy around the wire intensifies. Due to more energy available at the spark zone, deep craters are formed which contribute to the roughness of the cutting edge.

Kerf width depends on the radius of effective spark gap. With the rise in S_v spark energy increases, this increases accounts for a wider spark gap and hence higher kerf width values. Kerf width was found to be increase with T_{on}. As at higher T_{on} values, more electrons from the wire comes in contact with the neutral particles of dielectric fluid. This interaction magnifies the ionization effect which causes more material to vaporize. While wire tension is not affecting the kerf much, wire feed at lower S_v and T_{on} values decreases the kerf. This can be accounted to the inability of wire to maintain least required spark gap for conduction when the vibrations in the wire are more at higher feed. When more discharge energy is available, increase in vibrations results in widening the gap further.

43.5 Model Verification

Six random readings were selected to test the models generated by both the methods. First, they were put into the relation obtained from multivariate regression and each output value was compared with experimental value and a prediction error was calculated for both E_r and K_w as shown in Tables 43.3 and 43.4. Average error, E_{avg} was measured by the following formula:

Table 43.3 Comparison of experimental and predicted responses of E_r for test data

Machining parameters				Response parameters			Predictive error	
T_{on}	S_v	W_f	W_t	Experimental	ANFIS	Regression	ANFIS	Regression
100	40	8	1140	8.09	8.44	6.91	0.043	0.147
100	60	8	660	11.61	10.97	15.09	0.054	0.301
110	40	5	660	10.49	10.79	11.22	0.028	0.069
110	60	5	1140	10.81	10.63	11.43	0.017	0.057
120	20	8	1140	10.12	9.82	10.98	0.030	0.084
120	20	5	1140	9.38	9.90	7.04	0.055	0.249
						E_{avg}	0.038	0.151

Table 43.4 Comparison of experimental and predicted responses of K_w for test data

Machining parameters				Response parameters			Error	
T_{on}	S_v	W_f	W_t	Experimental	ANFIS	Regression	ANFIS	Regression
100	40	8	1140	302.689	299.811	302.689	0.009	0.055
100	60	8	660	292.041	297.377	292.041	0.018	0.002
110	40	5	660	271.998	288.615	271.998	0.057	0.129
110	60	5	1140	325.289	316.444	325.289	0.027	0.033
120	20	8	1140	331.386	325.799	331.386	0.016	0.003
120	20	5	1140	311.654	322.612	311.654	0.035	0.095
						E_{avg}	0.027	0.053

$$E_{\text{avg}} = \frac{1}{m}\sum_{i=1}^{m} Ei. \tag{43.5}$$

where E_i = error of sample number i an m is the total number of samples.

$$E_i = \left|\frac{\text{Experimental} - \text{Predicted}}{\text{Experimental}}\right| \tag{43.6}$$

The errors in multivariate regression model came out to be 15.1 and 5.3% for E_r, and $K_{w,}$ respectively. The same six readings were then predicted with the ANFIS model and again a prediction error was measured. ANFIS model predicted the output with an error of 3.8 and 2.7% for E_r and $K_{w,}$ respectively. Figure 43.4 shows the comparison of measured and predicted values of test data in ANFIS model. These low values of average error show that the ANFIS model developed is adequate to predict the studied response parameters, E_r and K_w. However, the average prediction error obtained through multivariate regression model for E_r and K_w are quite high when compared with ANFIS model and is thus, not capable of predicting the machinability aspects accurately.

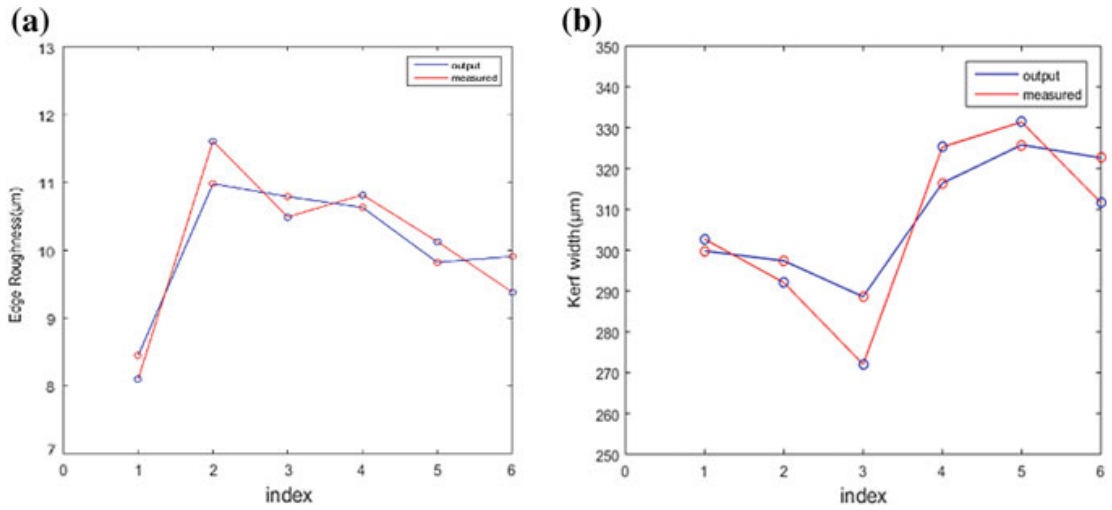


Fig. 43.4 Measured vs. predicted values for **a** edge roughness and **b** kerf width

43.5.1 Analysis of Machining Variables

Analysis of variance (ANOVA) was performed to determine significance of each machining parameters on both the responses. E_r is most influenced by servo voltage followed by W_t, W_f, and T_{on} as can be seen from Table 43.5 below, whereas Table 43.6 shows that T_{on} and W_f are the most and least important parameters, respectively, for affecting the kerf.

Main effects plots for the mean of E_r and K_w were also plotted using MiniTab software. Figure 43.5a shows the relation of E_r with the various machining parameters. Roughness is increasing with the increase in W_t, S_v, and T_{on} but is inversely dependent on the W_f, while K_w is seen to be directly dependent on all the input parameters.

Table 43.5 ANOVA table for mean w.r.t. E_r

Source	DF	Adj SS	Adj MS	*F*-value	*P*-value
T_{on}	2	2.565	1.282	0.3	0.745
S_v	2	56.17	28.087	10.39	0
W_f	1	14.83	14.827	3.86	0.058
W_t	1	1.489	1.489	0.35	0.557

Table 43.6 ANOVA table for mean w.r.t. K_w

Source	DF	Adj SS	Adj MS	*F*-value	*P*-value
T_{on}	2	7344	3671.9	16.83	0
S_v	2	2023	1011.6	2.67	0.084
W_f	1	267.3	267.3	0.64	0.43
W_t	1	729.7	729.7	1.8	0.189

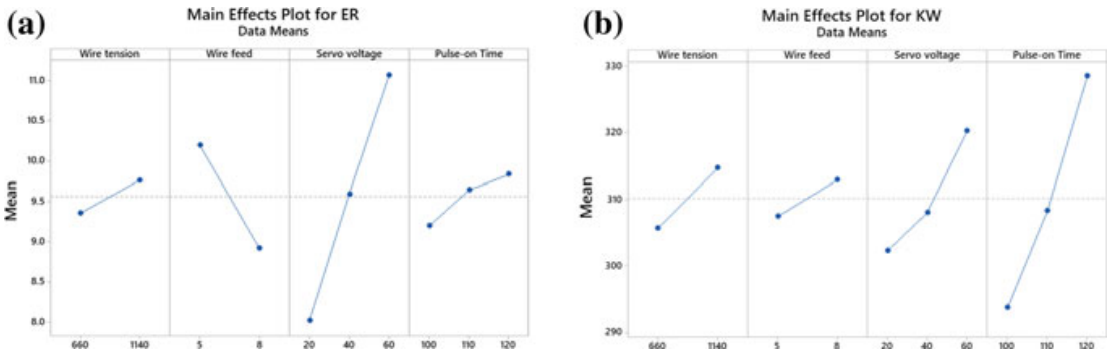


Fig. 43.5 Main effects plot for **a** edge roughness and **b** kerf width

43.6 Conclusion

Manufacturing Industry has a socio-economic responsibility in providing quality products keeping the costs incurred as low as possible. This has driven many researchers to look for best set of machining conditions. A full-factorial design has been used for experimentations upon which predictive models were built through multivariate regression and ANFIS. ANFIS model was able to predict with an accuracy of 96.2 and 97.3% for E_r and $K_{w,}$ respectively, which is higher than that obtained by regression model. Surface plots from ANFIS show that high T_{on} values increase the E_r significantly. High W_t also increased the roughness on edges. Similarly, increasing T_{on} and S_v increased the K_w. Lower wire feed value was able to decrease the K_w even at high S_v and T_{on}. These inputs cannot be kept too low as productivity is again a matter to be considered while production and hence a trade-off is necessary while deciding the input variables for any machining process. Further work can include finding the most optimum input variables which could give highly efficient process.

References

1. Mussada, E.K., Hua, C.C., Rao, A.K.P.: Surface hardenability studies of the die steel machined by WEDM. Mater. Manuf. Process. 1–6 (2018). https://doi.org/10.1080/10426914.2018.1476695
2. Mahapatra, S.S., Patnaik, A.: Optimization of wire electrical discharge machining (WEDM) process parameters using Taguchi method. Int. J. Adv. Manuf. Technol. **34**, 911–925 (2006). https://doi.org/10.1007/s00170-006-0672-6
3. Kumar, K., Agarwal, S.: Multi-objective parametric optimization on machining with wire electric discharge machining. Int. J. Adv. Manuf. Technol. **62**(5–8), 617–633 (2011)
4. Sharma, N., Raj, T., Jangra, KK.: Parameter optimization and experimental study on wire electrical discharge machining of porous Ni40Ti60 alloy. In: Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture, 1–15(2015). https://doi.org/10.1177/0954405415577710

5. Garg, S.K., Manna, A., Jai, A.: An investigation on machinability of Al/10% ZrO_2 (P)-metal matrix composite by WEDM and parametric optimization using desirability function approach. Arabian J. Sci. Eng. **39**(4), 3251–3270 (2014). https://doi.org/10.1007/s13369-013-0941-2
6. Babu, K.A., Venkataramaiah, P.: Multi-response optimization in wire electrical discharge machining (WEDM) of Al6061/SiCp composite using hybrid approach. J. Manuf. Sci. Prod. **15**(4), 327–338 (2015). https://doi.org/10.1515/jmsp-2015-0010
7. Pramanik, A., Basak, A.K., Dixit, A.R.: Processing of duplex stainless steel by WEDM. Mater. Manuf. Process. (2018). https://doi.org/10.1080/10426914.2018.1453165
8. Tosun, N., Pihtili, H.: The effect of cutting parameters on wire crater sizes in wire EDM. Int. J. Adv. Manuf. Technol. **21**(10–11), 857–865 (2003). https://doi.org/10.1007/S00170-002-1404-1
9. Chopra, K., Payla, A., Mussada, E.K.: Detailed experimental investigations on machinability of EN31 steel by WEDM. Trans. Indian Inst. Metals, 1–9(2019). https://doi.org/10.1007/s12666-018-1552-0
10. Majumder, H., Maity, K.P.: Predictive analysis on responses in WEDM of titanium grade 6 using general regression neural network (GRNN) and multiple regression analysis (MRA). Silicon **10**(4), 1763–1776 (2018). https://doi.org/10.1007/s12633-017-9667-1
11. Maher, I., Sarhan, A.A.D.: Proposing a new performance index to identify the effect of spark energy and pulse frequency simultaneously to achieve high machining performance in WEDM. Int. J. Adv. Manuf. Technol. **91**(1–4), 433–443 (2017). https://doi.org/10.1007/s00170-016-9680-3
12. Nain, S.S., Gard, D., Kumar, S.: Investigation for obtaining the optimal solution for improving the performance of WEDM of super alloy Udimet-L605 using particle swarm optimization. Eng. Sci. Technol. Int. J. **21**(2), 261–273 (2018). https://doi.org/10.1016/j.jestch.2018.03.005
13. Chojaczyk, AA.: Review and application of artificial neural networks models in reliability analysis of steel structures. Struct Saf, (2014). http://dx.doi.org/10.1016/j.strusafe.2014.09.002
14. Cardoso, J., de Almeida, J.R., Dias, J., Coelho, P.: Structural reliability analysis using Monte Carlo simulation and neural networks. Adv. Eng. Softw. **39**(6), 505–513 (2008)
15. Bucher, C.: Most T. A comparison of approximate response functions in structural reliability analysis. Probab. Eng. Mech. **23**(2–3), 154–63 (2008)
16. Ramakrishnan, R., Karunamoorthy, L.: Modeling and multi-response optimization of Inconel 718 on machining of CNC WEDM process. J. Mater. Process. Technol. **207**, 343–349 (2008). https://doi.org/10.1016/j.jmatprotec.2008.06.040
17. Yilmaz, O., Eyercioglu, O., Gindy, N.N.Z.: A user-friendly fuzzy-based system for the selection of electro discharge machining process parameters. J. Mater. Process. Technol. **172**, 363–371 (2006)
18. Maher, I., Sarhan, A.A.D., Marashi, H., Barzani, M.M., Hamdi, M.: White layer thickness prediction in WEDM-ANFIS modeling. Malaysian Int. Tribology Conf. 16–17, 240–241 (2015). Penang, Malaysia
19. Maher, I., Sarhan, A.A.D., Marashi, H., Barzani, M.M., Hamdi, M.: Increasing the productivity of the wire-cut electrical discharge machine associated with sustainable production. J. Cleaner Prod. (2015). https://doi.org/10.1016/j.jclepro.2015.06.047

Chapter 44
On Estimation of Scallop Height from CNC Part Programs

A. Kukreja and S. S. Pande

Abstract Today, complex freeform surfaces are widely used on parts in automobile, aerospace, and die– mold industries. Multi-axis computer numerical control (CNC) machines are used to manufacture such parts as they need tighter tolerances and high surface finish to meet functional and aesthetic requirements. Scallop height is primarily used to measure and assess the surface quality of such products. The inspection methods used in practice are expensive and need experienced personnel. No simulation techniques exist, as of now, to estimate the part quality apriori. This paper reports the development of an algorithm to estimate the scallop heights for freeform surface machining from the CNC part programs. The developed system takes STL file of the simulated machined component from the VERICUT and slices it along the specified planes. Points so generated are analyzed to estimate scallop heights at various regions on the surface using numerical curve-fitting techniques. The developed system has been validated with the case studies. The system provides a robust and accurate estimation of the part quality.

Keywords Scallop height estimation · Surface topography · CNC machining

44.1 Introduction

Today freeform surfaces are widely used on parts in automobile, aerospace, and die–mold industries. To suit various functional and aesthetic requirements, these parts need to be manufactured with tighter tolerances and good surface quality.

A. Kukreja
CAM Lab, Department of Mechanical Engineering, Indian Institute of Technology Bombay, Powai, Mumbai 400076, Maharashtra, India
e-mail: aman.kukreja1@gmail.com

S. S. Pande (✉)
CAM Lab, Department of Mechanical Engineering, Indian Institute of Technology Bombay, Powai, Mumbai 400076, Maharashtra, India
e-mail: s.s.pande@iitb.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_44

Multi-axes (3–5 axis) CNC machines are mainly used to manufacture such complex freeform surface parts [1]. They need a robust, error-free part program to drive the tool for machining.

The literature documents a focused research effort to develop various toolpath planning strategies (iso-planar, iso-scallop, etc.) [2–4] during the past two decades. In the commercial domain, several CAM systems exist which provide facilities for part modeling, toolpath planning, and CNC code simulation to generate and validate CNC part programs. CNC simulators provide the graphical simulation of toolpath and provide an estimation of machining time and code length. However, there is no facility in any CAM software/simulator to estimate the part scallop height which forms a primary measure of the part quality. VERICUT is one such software that provides the simulation of machining and gives the user a tool to measure scallop height at particular points [5]. The "manual" measurement system allows the measurement of scallops one point at a time which is cumbersome, error-prone, and could even leave some of the areas unobserved.

Literature reports very scant work focused on the prediction of surface roughness during milling. Ralph and Loftus [6] proposed a mathematical model for the average height, area, and volume of the scallops in freeform surface machining. However, it agreed poorly with machining trials. Bedi et al. [7] made a comparison of toroidal, ball end, and flat end mill in terms of surface roughness quantified as scallop heights and feed marks. A comprehensive review of the inspection techniques for machined freeform surfaces was presented by Li and Gu [8]. They pointed out that the inspection techniques are costly, complex, and time-consuming. The CNC part programs are thus required to be generated in a sequential design–plan–machine–inspect route which increases the cycle time. No system has been developed that can predict scallop height variation directly from part programs in the simulation. A need, thus, exists to develop an accurate procedure to estimate a priori part quality (scallop height) from the CNC part program during simulation. This paper is an attempt in this direction.

Rest of the paper is organized as follows. Section 44.2 explains the modular diagram and data flow of the system developed. The case studies are discussed in Sect. 44.3, and the conclusion is presented in Sect. 44.4.

44.2 Modular Diagram of the System

Figure 44.1 shows the modular diagram of the system. The developed system takes NC codes as input. It is simulated in VERICUT 8.0 to generate a CAD model of the machined surface in the STL format. The STL model file is exported, and is, in effect, processed by our system. The CAD model (STL file) is sliced by the inspection planes perpendicular to the lay direction of the machined surface. The inter-planar distance is definable by the user to decide a number of slicing planes. The intersection of the inspection (slicing) plane and STL machined model generates the surface profile in the form of points. A curve is fitted to these intersection

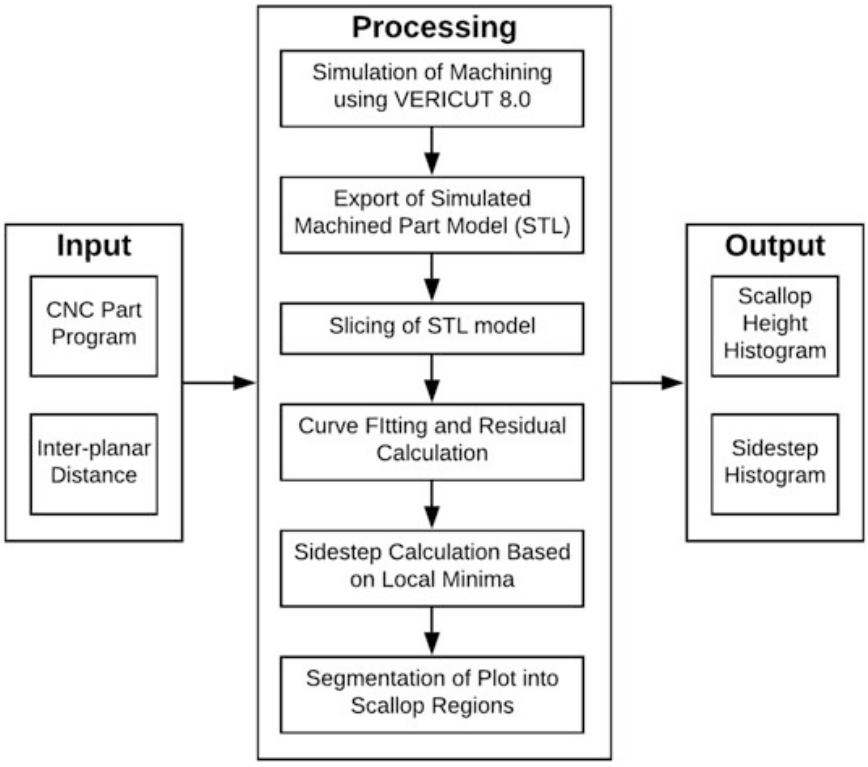


Fig. 44.1 Modular design of the system

points to generate a surface profile in the evaluation plane. Scallop height is computed from the curve by the residual calculation and segmentation at several points throughout the surface. Scallop height and step size estimated are plotted as histograms. In what follows, various system modules are explained.

44.2.1 Slicing of CAD Model

The machined part model (STL) model generated in VERICUT is sliced using parallel planes that are perpendicular to the lay direction (Fig. 44.2a). The intersection of the slicing planes and STL model would give points that form the surface topology pattern with superimposed waviness and roughness profiles (Fig. 44.2b).

Due to the direct slicing of STL facets, the scallop curve therefrom is well defined with less noise. A MATLAB function file "stl_slice_and_plot" was used [9]. The function first finds the triangles that intersect with the slicing plane. The lines created on the triangle face by the intersection plane are identified and processed further. The distance between the slicing planes is user definable, and hence, the user can control the areas to be evaluated and generate data on the desired surface topography.

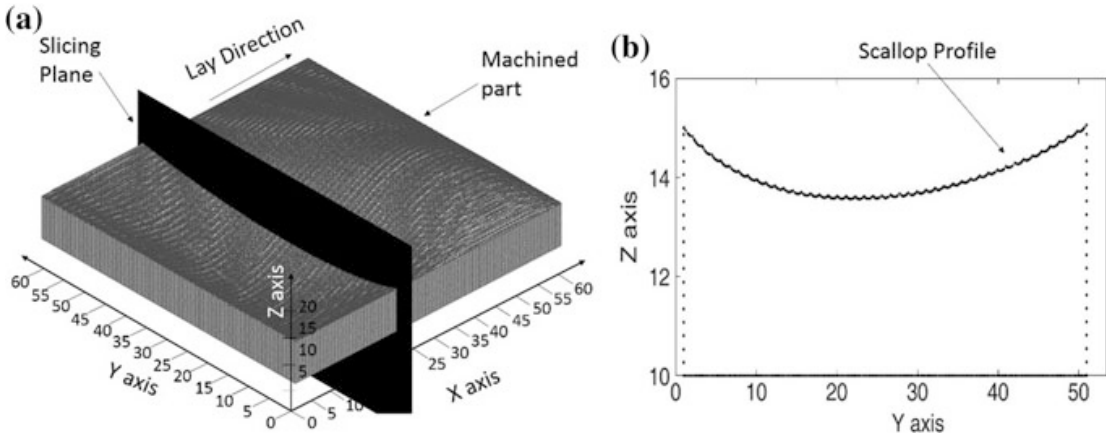


Fig. 44.2 **a** Slicing of the STL model and **b** points generated after slicing

44.2.2 Curve Fitting and Residual Calculation

After getting the point data, the next task was to remove waviness from the profile and calculate scallop heights based on the surface roughness profile. The waviness is removed by fitting a curve and then calculating the residuals of all data points. Residual gives the measure of skewness of the scallop data. Several curve-fitting methods were tried out [10]. For a large dataset, using "Gaussian fitting" and "third-degree polynomial fitting" resulted in poor fitting as the residual values were coming out to be large and were skewed on either side of the curve. After several trials, the ninth-degree polynomial fitting was found to be giving best approximation with residual variation within the desired level. Hence, in the present work, ninth-degree polynomial curve fitting approach has been used. The main advantages of polynomial fits include reasonable flexibility for data that is linear and not too complicated. This makes the fitting process fast, accurate, and computationally simple [11]. Equation (44.1) shows the polynomial model for curve fitting used in this work.

$$z = \sum_{k=0}^{n} p_i y^{n+1-i} \tag{44.1}$$

where $n + 1$ is the order and n is the degree of the polynomial ($1 \leq n \leq 9$). The order gives the number of coefficients to be fitted, and the degree gives the highest power of the predictor variable. The minimum distance (residual) of all the scallop points from the curve points are calculated based on Eq. (44.2).

$$\text{R(i)} = \min\left(\sqrt{\left(z_{1i} - z_{2j}\right)^2 + \left(y_{1i} - y_{2j}\right)^2} \quad \forall\, j\epsilon[1,m]\right) \;\forall\, i\epsilon[1,k] \tag{44.2}$$

where R = minimum distance of scallop point "i" from the fitted curve; m is the total number of points taken on the fitted curve; y_1, z_1 signifies the scallop points; k is the total number of scallop points; y_2, z_2 signifies the sample points taken on the fitted curve.

44.2.3 Segmentation and Scallop Height Calculation

Scallop pattern can be clearly seen on the profile after the residual calculation (Fig. 44.3b). The generated residual curve is then segmented based on the sidestep while generating the CNC toolpath. The sidesteps are identified by searching for consecutive local minima or local maxima points. The local minima are found by comparing the z-values of the preceding and succeeding points with the current point.

The distance of two consecutive local minima (the difference between y-values) gives the sidestep variation in the slicing plane, that, in turn, facilitate calculation of scallops. The plot shown in Fig. 44.4 is then segmented as per calculated sidesteps.

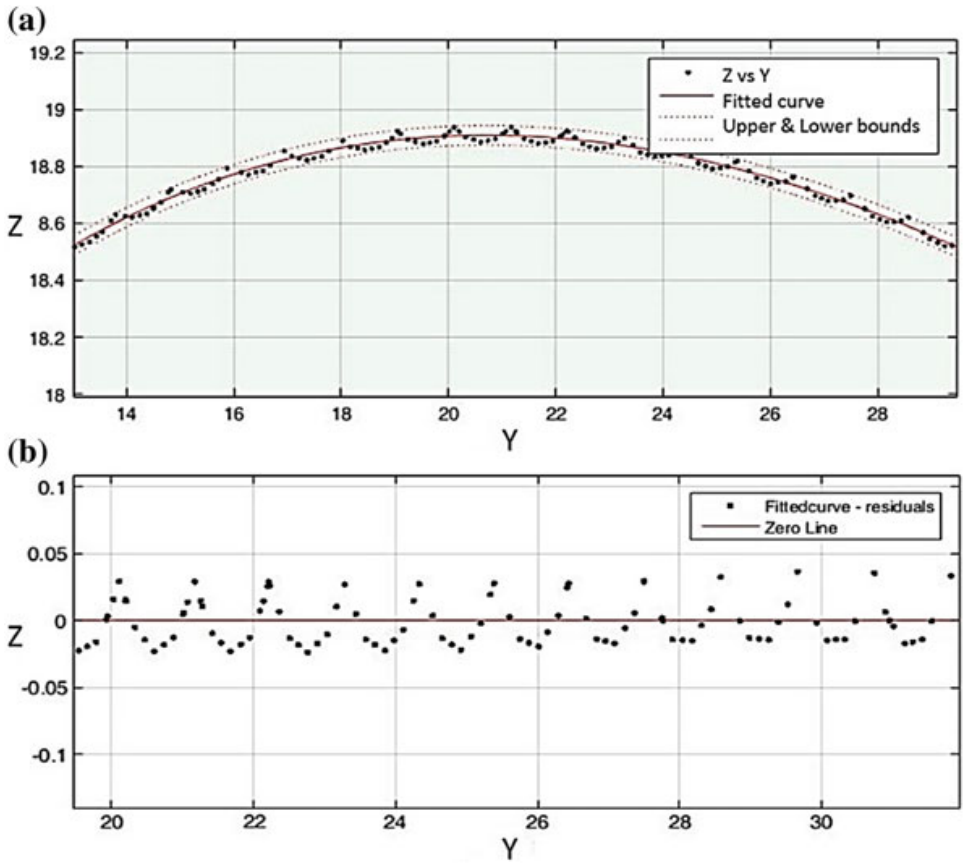


Fig. 44.3 **a** Curve fitting and **b** residual plot

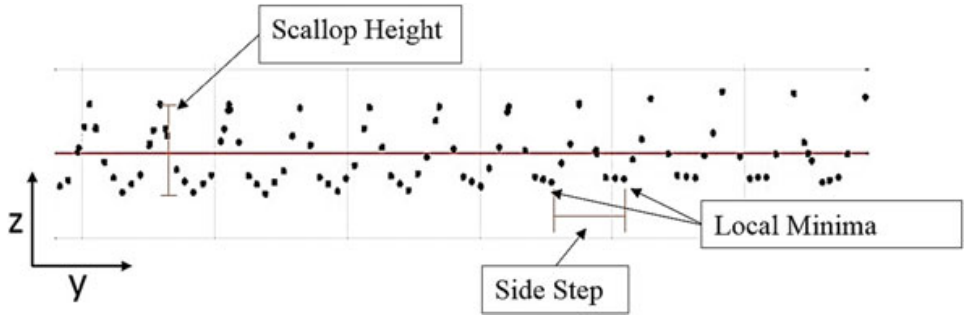


Fig. 44.4 Segmentation and scallop height calculation

The highest and lowest points (z) in the segments give scallop height for each sidestep (Eq. 44.3).

$$\check{S} = \max(z_i) - \min(z_i) \tag{44.3}$$

where $\check{S}$ is scallop height and z_i is the z-coordinate of scallop points. The case studies have been presented in the next section.

44.3 Results and Discussions

The algorithm for scallop estimation was implemented using MATLAB R2015b on Windows 10 platform, Intel Core i5, and 4 Gb RAM. The simulation software VERICUT 8.0 was used to generate a CAD model of the part (in the STL format) produced by the input CNC part program. The validation of the developed system was done using Mastercam X5-based iso-planar toolpath planning strategy. This was followed by extensive testing of the system on different kinds of surfaces produced with CNC programs with various toolpath strategies, like iso-parametric, iso-scallop, and adaptive planar [12].

The following case studies show the results and validation of the approach. The case study 1 refers to a relatively simple geometry with known values of scallops while case study 2 explores a complex bicubic Bezier surface with adaptive sidesteps generated using algorithm proposed by Dhanda and Pande [12] for three-axis CNC machining.

44.3.1 Case Study 1

In this case, the CNC part program was generated using Mastercam X5. The model has concave and convex portions (Fig. 44.5a). The scallop height was set to be 0.05 mm on a flat surface with sidestep 1.091 mm. It can go up to 0.1 mm at 45° to the lay direction.

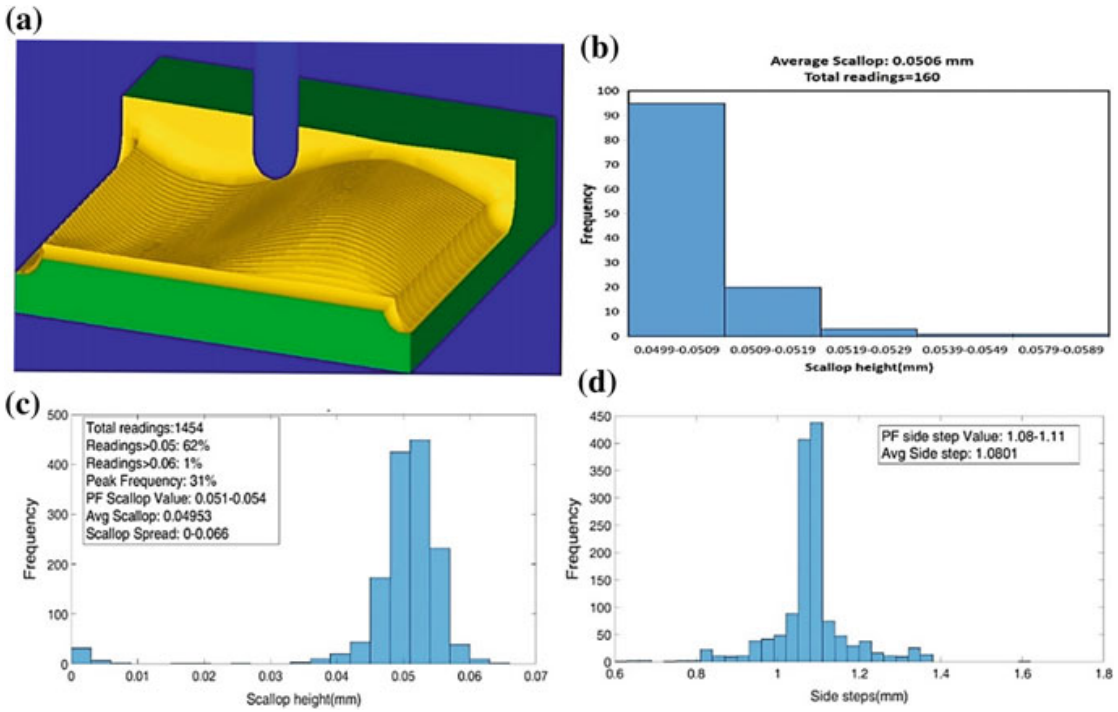


Fig. 44.5 Case 1; **a** Mastercam generated tool paths; **b** scallop estimation using VERICUT; **c** scallop estimation using developed system; and **d** sidestep variation

The scallop option in the VERICUT X-Caliper was used to measure the scallop height formed by two parallel intersecting cylinders chosen manually. It then identifies the cylinders being measured by displaying intersecting circles on the machined part model [5]. Figure 44.5b shows the generated histogram of the values estimated by the VERICUT (Range: 0.0499–0.058 mm). The scallop heights were estimated on 160 points (15 planes) throughout the surface using the X-Caliper tool of VERICUT. The average scallop was found to be 0.0506 mm. It is a known fact that the concave region has larger scallop heights and convex region has smaller scallop heights than the flatter region for the same sidesteps. The values estimated using VERICUT does not show this trend. This is because it does not consider the curvature of the surface while estimating scallops. The measurement/estimation is very approximate with limited points and sometimes unable to cover the whole machined surface.

Figure 44.5c shows the scallop height values estimated using the developed system (Range: 0.04–0.06 mm). The values below 0.05 mm indicate the convex regions and the values above 0.05 mm indicates concave regions. The scallop heights are calculated based on the sidesteps that are found using the method explained in Sect. 44.2.3. A similar trend is seen for the sidestep (Fig. 44.5d). In both the Fig. 44.5c and d, there is a spread but it is quite small. It is seen that the scallop frequency histogram from our system (Fig. 44.5c) has a well-defined symmetric peak around the expected scallop value (0.05 mm).

The spread in the sidestep values is due to the approximation in generating STL model of the machined stock, and the slicing of the STL model which may induce

errors (in calculating local minima) as the plane (facet) and curve intersection was solved using numerical methods in the developed system. This error in sidestep values directly affects the scallop height estimation and therefore some values of scallop height (Fig. 44.5c) are outside the tolerance band (0.04–0.06 mm).

44.3.2 Case Study 2

In this case, a complex bicubic Bezier surface was designed (Fig. 44.6a) and a CNC part program was generated using an adaptive planar strategy [12]. The maximum permissible sidesteps in various regions were automatically calculated by the algorithm reported in [12] based on surface curvature while maintaining the surface finish (i.e., tolerance <0.05 mm). The sidesteps were found to be in the range of 0.75–1.5 mm to maintain the required surface quality [12].

Figure 44.6(b) shows the scallop heights frequency histogram estimated using VERICUT (Range: 0.0235–0.0485 mm). The scallop height was estimated at 120 points (9 planes). The average scallop was found to be 0.0372 mm. Similar to case study 1, it is seen that the VERICUT values are not considering the effect of surface curvature. The scallop values are under tolerance but its range is much wider than desired. The developed system shows that most of the values of scallop height are in the range of 0.04–0.05 mm (Fig. 44.6c). About 83% of values were found to be

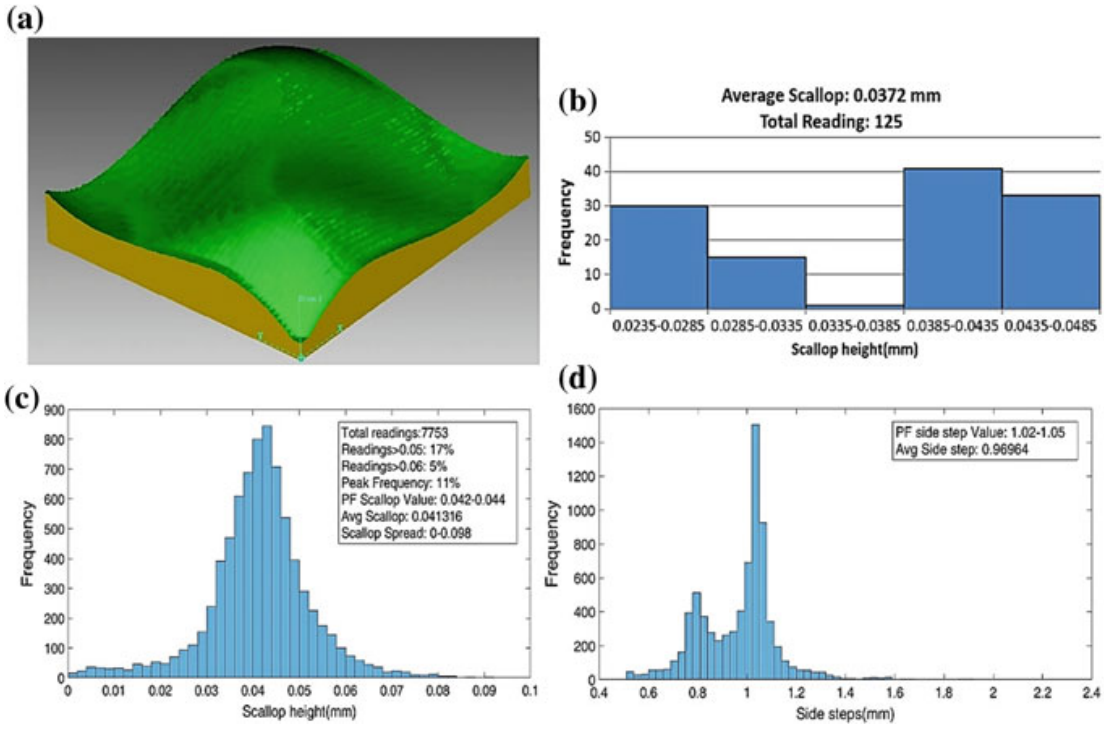


Fig. 44.6 Case 2; **a** Mastercam generated tool paths; **b** scallop estimation using VERICUT; **c** scallop estimation using developed system; **d** sidestep variation

below 0.05 mm with the average scallop of 0.041 mm. Hence, it agrees well with the desired values. According to the adaptive strategy reported in [12], the sidesteps for the concave regions were below 1.091 mm and for convex regions above 1.091 mm, a similar trend can be seen in Fig. 44.6d. The scallop heights were calculated based on these sidesteps values. The out-of-bound values seen in Fig. 44.6d may be again due to inherent error in the approximation done during STL generation and slicing algorithm.

As seen in case studies, the system is robust enough to estimate scallop for any kind of machined surface. The developed system could also be applied to five-axis machined parts. VERICUT cannot analyze these, and it fits the circle on the machined surface to find scallop heights but the surface generated in five-axis machines using flat end mill forms elliptical scallops. In essence, the developed system gives an accurate, fast, and robust estimation of scallop values from the simulation of a CNC part program without performing the actual cutting. It will thus help in significantly reducing product development lead time.

44.4 Conclusions

Despite a lot of research on CNC toolpath planning and simulation, no system exists that can estimate part quality (scallop height) directly from CNC part programs for the entire surface. Simulation software like VERICUT provides a very approximate tool. The system developed in this paper precisely caters to these requirements. The work proposes an algorithm for scallop height estimation from the simulation of machining using CNC part programs. In essence, it analyzes the simulated 3D machined surface, develops a mathematical model for the surface topography, and estimates the scallop and wavelength parameters. The system has been extensively tested for a variety of CNC part programs and surface topographies. It provides an accurate, fast, and robust estimate of part quality from the CNC part programs apriori.

References

1. Choi, B.K., Jerard, R.B.: Sculptured surface machining: theory and applications. Kluwer Academic Publishers, the University of Michigan, USA, Dordrecht (1998)
2. Elbert, G., Cohen, E.: Tool path generation for freeform surface models. Comput. Aided Des. **26**(6), 490–496 (1994)
3. Feng, H.Y., Huiwen, L.: Constant scallop-height tool path generation for three-axis sculptured surface machining. Comput. Aided Des. **34**(9), 647–654 (2002)
4. Park, S.C., Choi, B.K.: Tool-path planning for direction-parallel area milling. CAD Comput. Aided Design **32**(1), 17–25 (2000)
5. Getting Started with VERICUT Introduction. Introduction to VERICUT, https://community.ptc.com/sejnu66972/attachments/sejnu66972/installation_licensing/12339/1/vericut_help.pdf. Last accessed 09 Apr 2018

6. Ralph, W.L., Loftus, M.: Cusp geometry analysis in free-form surface machining. Int. J. Prod. Res. **30**(11), 2697–2711 (1992)
7. Bedi, S., Ismail, F., Mahjoob, M. J., C. Y.: Toroidal versus ball nose and flat bottom end mills. Int. J. Adv. Manf. Tech. 326–332 (1997)
8. Li, Y., Gu, P.: Free-form surface inspection techniques state of the art review. Comput. Aided Des. **36**(13), 1395–1417 (2004)
9. STL slice and create path, https://www.mathworks.com/matlabcentral/fileexchange/62113-slice_stl_create_path-triangles-slice_height. Last accessed 09 Apr 2018
10. MathWorks.: Polynomial curve fitting (polyfit), 1–19 (2017)
11. Mastercam.: Getting started series—basic 3D machining, http://eng.auburn.edu/∼griffgj/training/Basic_3D_Machining.pdf. Last accessed 09 Apr 2018
12. Dhanda, M., Pande, S.S.: Adaptive tool path planning strategy for freeform surface machining using point cloud. Camput. Aided Design Appl. **16**(2), 289–307 (2019)

Chapter 45
Influence of Guides on Critical Speeds of Circular Saws

S. Singhania, P. Kumar, S. K. Gupta and M. Law

Abstract This paper investigates the influence of guides on the critical speeds of circular saws. Guides constrain the out-of-plane lateral motion of these rotating saws and have a stabilizing effect by increasing the critical speeds, below which the rotating saw is stable. We present expanded analytical formulations in which guides are modelled as multiple discrete spring-damper elements approximating the distributed nature of the guide pad and saw interactions. We observe that for a given guide pad area, convergence analysis is necessary to understand how many discrete spring-damper elements are actually necessary to approximate the distributed nature of the guide pad and saw interactions. Curiously, we observe that damping in the guides has no significant influence on the critical speed and that it only changes the nature of frequency-speed characteristics of the rotating saw. We also find that critical speeds are sensitive to guides modelled with a distribution of discrete spring elements along the radial and/or circumferential directions. These observations suggest that more generalized formulations that model the guide pads as distributed spring-damper systems, rather than multiple discrete spring-damper elements, are necessary. We expect our findings to instruct and advise the placement of guides on rotating circular saws, such that a preferential increase in critical speeds can be obtained to make possible high speed and productivity circular sawing operations in the wood-working and metal-cutting industries concerned with circular sawing processes.

Keywords Circular saw · Guides · Critical speeds instability

S. Singhania · P. Kumar · S. K. Gupta · M. Law (✉)
Machine Tool Dynamics Laboratory, Department of Mechanical Engineering,
Indian Institute of Technology Kanpur, Kanpur, UP 208016, India
e-mail: mlaw@iitk.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_45

45.1 Introduction

Circular sawing is a common cutting operation across the wood-working and metal-cutting industries. It involves feeding a rotating circular saw into the work-piece to cut logs, bars, and tubes to lengths of desired sections. To meet the demands on high productivity in these industries, circular saws are desired to be operated at high speeds allowing more material to be removed per unit time. However, when the saw is rotating, forward- and backward-travelling waves are generated. The natural frequencies of the backward-travelling waves decrease with an increase in saw rotation speed, and the speed at which a natural frequency corresponding to a backward wave becomes zero is known as the critical speed [1]. When rotation speeds correspond to the critical speed, large amplitude vibrations appear, which can destroy the saw, the work material being cut, and the machine tool system. It is hence necessary to understand the factors governing critical speed instability, which is the focus of this paper.

Significant research attention across industries and academia has been paid on characterizing, predicting, and avoiding critical speeds. In some early seminal work, Mote and Nieh [2], and Szymani and Mote [3] studied the vibrations and critical speeds of rotating circular saws. Contemporarily, as a means of increasing the critical speeds of saws, Iwan and Moeller [4] presented formulations for saws constrained by guides. Guides constrain the out-of-plane lateral motion of rotating flexible saws and increase their critical speeds. Since these guides offer resistance to motion and act as a transverse load, Schajer [1], Hutton et al. [5], and Lehmann and Hutton [6], all investigated the influence of guided saws, with guides modelled as lumped spring elements. Hutton [5] observed that multiple guides were necessary to increase the critical speed. Though their [1, 5, 6] formulations allowed for the possibility of incorporating multiple guides, they ignored the influence of damping, if any, offered by guides. Damping in guides was investigated by Mote [7], and he observed that, if not favourably positioned, guides promote instability. However, his analysis was for the case of a stationary saw subjected to a moving load, and not the case we are interested in—a rotating, and constrained saw.

Since guides are actually pads distributed over a sector area constraining the lateral out-of-plane motion of the rotating saw, they need to be modelled as distributed systems. However, all the earlier reported work [1, 5–7] modelled the guide as a spring lumped at the geometric centre of the guide pad. Mohammadpanah and Hutton [8] were the first to investigate the distributed pad–saw interaction by modelling the guide as 16 linear springs distributed equally over the area of the guide pad. However, the effect of distributed damping in the guide was ignored. They also did not report on how many springs are sufficient to describe the distributed guide area, i.e. no convergence analysis of critical speeds to an increase in number of springs was presented.

This paper addresses the above two issues, i.e. understanding through systematic convergence studies, how many spring-damper systems are actually necessary to approximate the distributed pad–saw interaction, and understanding the role of

damping in guides on the critical speeds of circular saws. To do so, we systematically modify the formulations for guided rotating circular saws, to include the possibility of incorporating multiple spring-damper systems. These modified formulations are presented in Sect. 45.2. Models are first numerically verified in Sect. 45.3 with results from a commercially available circular sawing stability code, CSAW® [9]. Verification is only provided for the undamped guide modelled as a single spring. Numerically verified models are then subjected to convergence analysis in Sect. 45.4, which includes discussions on the influence of damping on critical speeds, as well as discussion about the influence of modelling discrete spring-damper elements distributed radially, and circumferentially. The paper is concluded in Sect. 45.5.

45.2 Mathematical Model for the Critical Speed of Circular Saws

A circular saw, laterally constrained by a guide pad, is modelled as a flexible disc with a clamped inner radius, r_i, outer radius, r_0, rotating with an angular velocity of Ω in the clockwise direction as shown in Fig. 45.1. Ignoring any external load, the governing equation of motion of the circular saw in the stationary frame of reference with a guide pad modelled as multiple spring-damper systems, distributed equally over the area of guide pad, is modified from [5] to become:

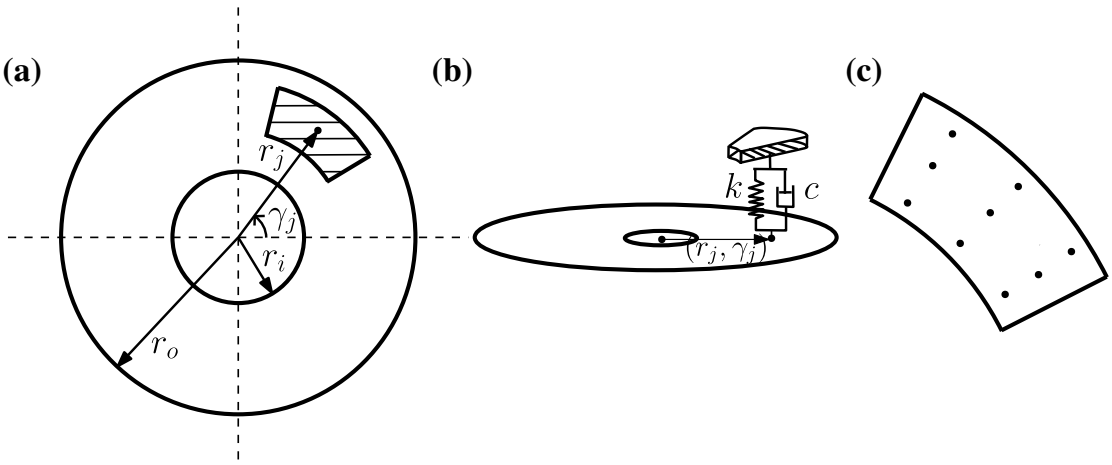


Fig. 45.1 Schematic of **a** circular saw and guide pad **b** the interaction between the guide pad and circular saw and **c** distribution of discrete spring-damper elements over the guide pad

$$D\nabla^4 u + \rho h\left(\frac{\partial^2 u}{\partial t^2} + 2\Omega\frac{\partial^2 u}{\partial t\partial\gamma} + \Omega^2\frac{\partial^2 u}{\partial\gamma^2}\right) - h\left[\left(\frac{1}{r}\right)\left(\frac{\partial}{\partial r}\right)\left(r\sigma_r\frac{\partial u}{\partial r}\right) + \left(\frac{1}{r^2}\right)\left(\frac{\partial}{\partial\gamma}\right)\left(\sigma_\gamma\frac{\partial u}{\partial\gamma}\right)\right] + \sum_{j=1}^{J} k_j u\left(\frac{1}{r}\right)\delta(r-r_j)\delta(\gamma-\gamma_j) + \sum_{j=1}^{J} c_j\frac{\partial u}{\partial t}\left(\frac{1}{r}\right)\delta(r-r_j)\delta(\gamma-\gamma_j) = 0 \tag{45.1}$$

wherein the first term represents bending stiffness, the second term represents inertial stress, the third term represents in-plane rotational stress, and the fourth and fifth terms represent stiffness and damping due to the interaction between the guide and circular saw, respectively, at the location of r_j, γ_j.

In Eq. (45.1), D is the flexural rigidity $(= Eh^3/12(1-\nu^2))$, E is the Young's modulus, h is circular saw thickness, ν is the Poisson ratio, ρ is the density of circular saw; $\nabla^4\left(= \left[\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \left(\frac{1}{r^2}\right)\left(\frac{\partial^2}{\partial\gamma^2}\right)\right]^2\right)$ is a bi-harmonic operator; $\delta(\cdot)$ is the Dirac delta function, k and c represent the stiffness and damping of the guide pad; J is the number of spring-damper elements; $u(r,\gamma,t)$ is the transverse displacement of the circular saw, and r, γ represent the radial and angular coordinates with respect to a fixed frame of reference.

Equation (45.1) is a fourth-order partial differential equation, and since it is difficult to solve it analytically, we use the Galerkin projection method to solve it [5]. To represent the deflection of circular saw due to bending, the transverse displacement $u(r,\gamma,t)$ can be written by a modal expansion. Therefore, the solution of Eq. (45.1) is assumed in the following form [5, 10]:

$$u(r,\gamma,t) = \sum_{m=0}^{m}\sum_{n=0}^{n}[C_{mn}(t)\cos(n\gamma) + S_{mn}(t)\sin(n\gamma)]R_{mn}(r) \tag{45.2}$$

wherein $C_{mn}(t)$ and $S_{mn}(t)$ are the unknown functions (modal coordinates), m and n represent the number of nodal circles and nodal diameters, respectively, and $R_{mn}(r)$ are radial functions chosen to satisfy the inner and outer boundary conditions of the circular saw. These $R_{mn}(r)$ can be approximated by polynomial functions as [5]:

$$R_{mn}(r) = \sum_{i=0}^{4} E_{mni} r^{m+i}, \tag{45.3}$$

wherein E_{mni} are unknown coefficients determined from the out-of-plane boundary conditions given by Eq. (45.4) and additional normalizing conditions, $R_{mn}(r_o) = 1$. Since the circular saw is clamped at its inner radius and is free at its outer radius, displacement and the slope at the inner radius will be zero, whereas bending

moment and shear force at the outer radius will be zero. Therefore, the boundary condition at the inner and outer radius in terms of transverse displacement can be written as:

$$\left.\begin{array}{c} u(r_i,\gamma,t)=0, \frac{\partial}{\partial r}u(r_i,\gamma,t)=0, \\ M_{rr}(r_o,\gamma,t)=0, Q_r(r_o,\gamma,t)-\frac{1}{r_o}\left(\frac{\partial}{\partial\gamma}\right)M_{r\gamma}(r_o,\gamma,t)=0, \end{array}\right\} \tag{45.4}$$

wherein M_{rr} and $M_{r\gamma}$ are bending moments, and Q_r is the shear force [5].

Equation (45.1) can be rewritten with brevity, in terms of a differential operator (L) as:

$$L(u)=0, \tag{45.5}$$

wherein

$$L=D\nabla^4+\rho h\left(\frac{\partial^2}{\partial t^2}+2\Omega\frac{\partial^2}{\partial t\partial\gamma}+\Omega^2\frac{\partial^2}{\partial\gamma^2}\right)-h\left[\left(\frac{1}{r}\right)\left(\frac{\partial}{\partial r}\right)\left(r\sigma_r\frac{\partial}{\partial r}\right)+\left(\frac{1}{r^2}\right)\left(\frac{\partial}{\partial\gamma}\right)\left(\sigma_\gamma\frac{\partial}{\partial\gamma}\right)\right]$$
$$+\sum_{j=1}^{J}k_j\left(\frac{1}{r}\right)\delta(r-r_j)\delta(\gamma-\gamma_j)+\sum_{j=1}^{J}c_j\frac{\partial}{\partial t}\left(\frac{1}{r}\right)\delta(r-r_j)\delta(\gamma-\gamma_j).$$

Since the assumed form of the solution does not satisfy the equation of motion exactly, the residue can be minimized using the Galerkin projection approach, resulting in:

$$\int_0^{2\pi}\int_{r_i}^{r_o}L(u)(R_{ql}\cos(l\gamma))rdrd\gamma=0, \tag{45.6}$$

$$\int_0^{2\pi}\int_{r_i}^{r_o}L(u)(R_{ql}\sin(l\gamma))rdrd\gamma=0, \tag{45.7}$$

wherein $q=0,1,\ldots,m$, $l=0,1,\ldots n$. Solving and simplifying Eqs. (45.6) and (45.7), we get:

$$\sum_{m=0}^{m}\alpha_1\left[\delta^{(1)}_{qml}\frac{d^2C_{ml}}{dt^2}+\delta^{(2)}_{qml}\frac{dS_{ml}}{dt}+\delta^{(3)}_{qml}C_{ml}\right]+\sum_{m=0}^{m}\left[N^{(1)}_{qml}C_{ml}+N^{(2)}_{qml}S_{ml}\right]=0, \tag{45.8}$$

$$\sum_{m=0}^{m}\beta_1\left[\delta^{(1)}_{qml}\frac{d^2S_{ml}}{dt^2}-\delta^{(2)}_{qml}\frac{dC_{ml}}{dt}+\delta^{(3)}_{qml}S_{ml}\right]+\sum_{m=0}^{m}\left[N^{(3)}_{qml}C_{ml}+N^{(4)}_{qml}S_{ml}\right]=0, \tag{45.9}$$

$$\alpha_1=\left\{\begin{array}{ll}\pi, & l\neq 0\\ 2\pi, & l=0\end{array}\right\},\beta_1=\left\{\begin{array}{ll}\pi, & l\neq 0\\ 2\pi, & l=0\end{array}\right\}.$$

For the sake of brevity, $\delta^{(1)}_{qml}$, $N^{(1)}_{qml}$, $N^{(2)}_{qml}$, $N^{(3)}_{qml}$, $N^{(4)}_{qml}\delta^{(2)}_{qml}$, $\delta^{(3)}_{qml}$ are not defined here, and the reader is directed to [5] for details. The above Eqs. (45.8) and (45.9) can be rewritten in a compact form as:

$$[A]\{\ddot{x}\} + [B]\{\dot{x}\} + [C]\{x\} = 0 \tag{45.10}$$

where the vector $\{x\}$ is an array of $\{C_{0,0}, C_{0,1},\ldots, C_{1,0}, C_{1,1},\ldots, C_{m,n}, S_{0,1}, S_{0,2},\ldots, S_{1,1}, S_{1,2}, \ldots, S_{m,,n}\}$. In Eq. (45.10), $[A]$ is the mass matrix, $[B]$ is the gyroscopic matrix, and $[C]$ is the stiffness matrix. Since $[B]$ and $[C]$ depend on the rotational speed of circular saw, dynamics of circular saw depends on rotational speed. As Eq. (45.10) is a second-order linear differential equation, the solution of the equation can be assumed in the form of $x(t) = \{X\}e^{\lambda t}$. On substituting the assumed form of the solution, Eq. (45.10) forms an eigenvalue problem which can be solved for every combination of rotational speed and other parameters. The imaginary part of the eigenvalue, λ of this eigenvalue problem, corresponds to a natural frequency of the circular saw. To reduce the effective number of parameters, we non-dimensionalize Eq. (45.10) as:

$$\Omega_0 = \left[\frac{\rho h r_o^4}{D}\right]^{\frac{1}{2}}\Omega, \zeta = k\left(\frac{r_o^2}{D}\right), \phi = \left[\frac{\rho h r_o^4}{D}\right]^{\frac{1}{2}}\omega, c_1 = \frac{c r_o^2}{\sqrt{\rho h r_o^4 D}}$$

wherein c_1 is the non-dimensional damping, ζ is the non-dimensional stiffness, ω is the natural frequency, ϕ is the non-dimensional natural frequency, and Ω_0 is the non-dimensional rotational speed.

We solve the above set of equations and, at first, verify the formulations against results from CSAW®, as discussed next, in Sect. 45.3, followed by convergence testing and investigations with damping in Sect. 45.4.

45.3 Numerical Verification of Critical Speed Model

As a preliminary check of our proposed formulations, we first verify our results with those obtained from a commercially available circular sawing stability code, CSAW®. Comparisons are made with frequency-speed diagrams, known as the Campbell diagram. We limit our numerical checks for the case of an unguided saw, and for the case of a guided saw, with the guide modelled as an undamped point spring—since CSAW® does not have the facility of including damping. We limit our comparisons for the number of nodal circles, m as 1, and the number of nodal diameters, n as 4. The saw, assumed to be made of steel, is clamped at its inner radius, $r_i = 42.5$ mm and has a free outer radius, $r_o = 142.5$ mm.

Campbell diagrams for the unguided case are compared in Fig. 45.2. As is evident, the critical speed obtained from the proposed analytical model (Fig. 45.2a) occurs at 8497 rpm, whereas the critical speed obtained from CSAW® (Fig. 45.2b)

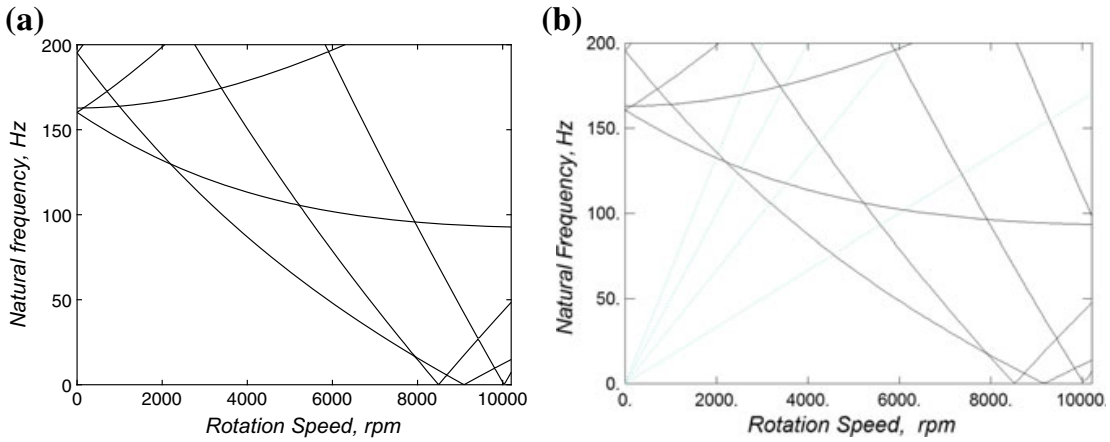


Fig. 45.2 Campbell diagram for unguided circular saw obtained **a** using analytical model and **b** using CSAW®

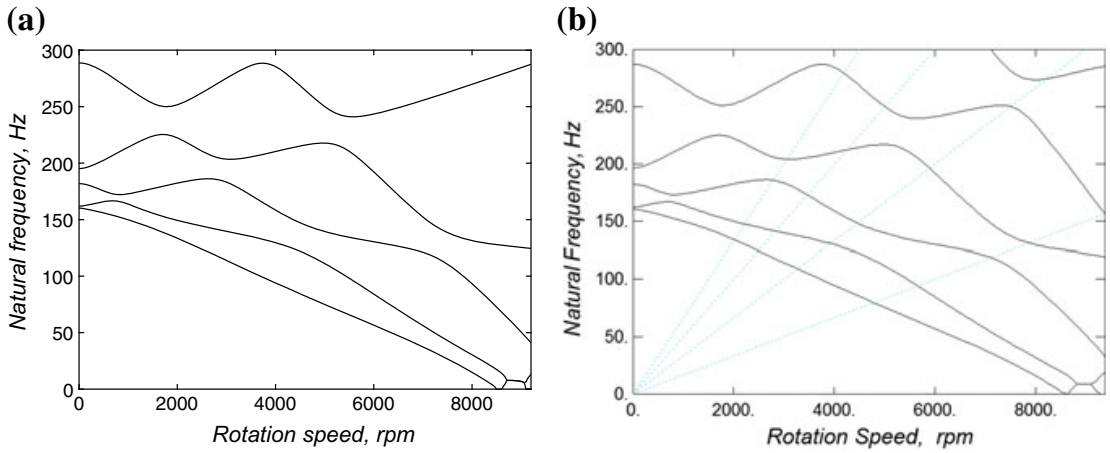


Fig. 45.3 Campbell diagram for guided circular saw obtained **a** using analytical model and **b** using CSAW®

occurs at 8519 rpm, which is approximately same as our analytical model. In the Campbell diagram obtained from CSAW®, there appear four straight lines from the origin. These represent the first four engine order (EO) lines and correspond to imperfections or unbalance in the rotating parts of the system. These faults generate harmonic and sub-harmonic levels of vibrations [8] as shown in Fig. 45.2b.

Campbell diagrams for the guided case are compared in Fig. 45.3. The guide pad has been modelled as a sector of a circular disc with orientation of inner and outer angle as 5° and 36°, respectively, whereas, inner and outer radii of the guide pad are 57 and 142.5 mm, respectively. The spring element, located at the geometric centre of the pad, has a stiffness 7.5×10^8 N/m. We observe that the analytical model predicted response (Fig. 45.3a) is consistent with the results from CSAW® (Fig. 45.3b), and that there is no change in the critical speed with a single guide. This observation is consistent with those reported earlier in [5].

Having numerically verified the analytical model, we use the analytical model to understand how many spring-damper systems are actually necessary to approximate distributed guide and saw interactions, and to understand the role of damping in guides on the critical speeds of circular saws, as is discussed next, in Sect. 45.4.

45.4 Results and Discussion

In this section, we first present convergence analysis for understanding how many spring-damper systems are actually necessary to approximate distributed guide and saw interactions, for a given guide area, followed by analysis for damped guides, followed by analysis on the influence of modelling multiple spring-damper systems radially and circumferentially. Guide pad area for all results presented herein is 4614 mm^2. Guide location is the same as described in Sect. 45.3. All combinations of the spring-damper elements approximating the guide pad are distributed uniformly over the guide pad area. All results presented herein are for the mode corresponding to the first critical speed. All results are presented in the non-dimensional form of the Campbell diagrams.

45.4.1 Convergence Analysis of Discrete Multiple Spring Elements to Model the Guide Pad

For a given guide area, and location, it is, in general, difficult to a prior establish the required number of discrete spring elements, necessary to capture the saw and guide interactions. We neglect the damping in the guide presently, influence of which is discussed separately in Sect. 4.2. We analyse sensitivity of the first critical speed with five combinations of spring elements approximating the same guide area, i.e. with J as 1, 4, 9, 16, and 25. Overall effective stiffness k_j in each case was taken to be as 7.5×10^8 N/m. Results for all cases are shown in Fig. 45.4.

As is evident from Fig. 45.4, an increase in number of springs increases the critical speed. However, beyond 16 springs, the critical speed does not appear to change, suggesting that 16 springs are sufficient to describe the size and location of the guide pad under consideration. Interestingly, we observe that with four discrete springs, the critical speed increases, which is not consistent with the findings reported in [5], in which they found that to increase the critical speed, at least eight discrete uniformly distributed guides modelled as a discrete point spring elements are necessary. This suggests that it is indeed necessary to model the pad–saw interaction as multiple discrete spring elements.

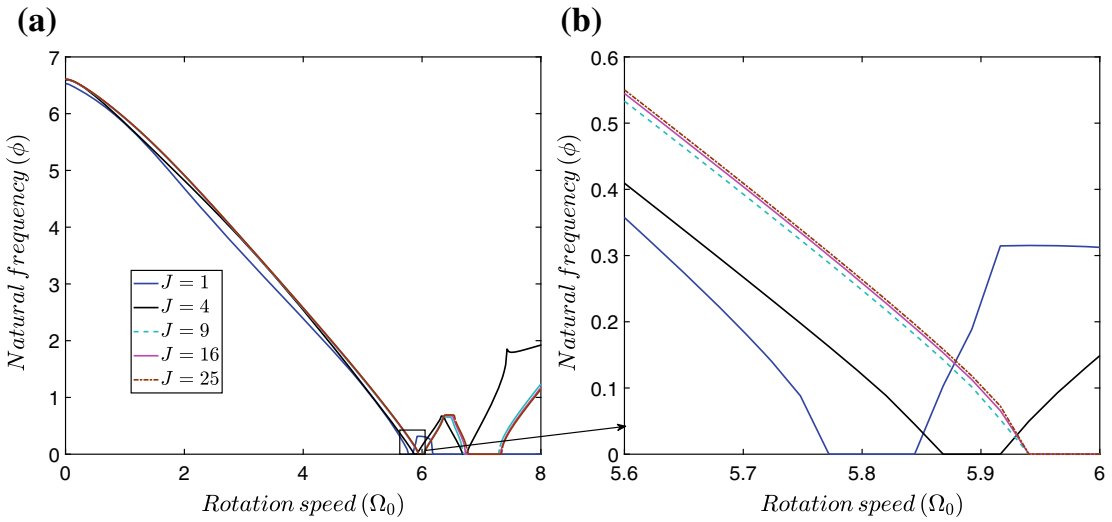


Fig. 45.4 Campbell diagram for multiple discrete spring elements distributed equally over the guide pad corresponding to the lowest mode

45.4.2 Influence of Damping on the Critical Speed

Since guides offer not just stiffness, but also damping, influence of damping on the critical speed is reported herein. To delineate the influence of damping, i.e. without stiffness, we present results for a guide with only damping in Fig. 45.5a and for a guide with stiffness and damping in Fig. 45.5b. For the case of a guide with damping only, we model the guide to have an overall effective damping of $c_j = 17.2\,\mathrm{N/m/s}$, for equivalent 16 damping elements. For the case of a guide modelled with discrete spring-damper elements, we take the number of elements to be 16, with an equivalent overall stiffness and damping of $7.5 \times 10^8\,\mathrm{N/m}$ and $17.2\,\mathrm{N/m/s}$, respectively.

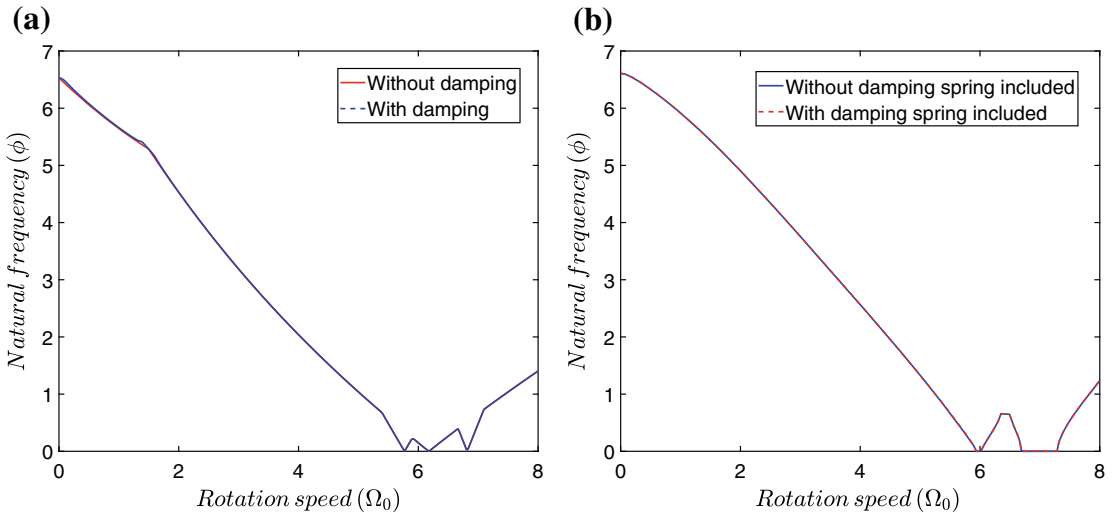


Fig. 45.5 Campbell diagram for circular saw with pad–saw interaction modelled as **a** discrete damper elements **b** discrete spring-damper elements corresponds to the lowest mode

As is evident from Fig. 45.5a, saws with guides with damping only, i.e. without springs, do not influence the critical speed and behave like unguided saws, which are curious. We also observe that damping only affects the frequency-speed characteristics. Furthermore, even for the case of guides modelled with multiple discrete spring-damper systems, the critical speed does not change when compared to the critical speed for the case of guides modelled only as springs, as is evident from Fig. 45.5b. These observations suggest that damping has no influence on the critical speeds of saws for the given guide pad configuration (mentioned in Sect. 45.3), which is unusual and unexpected and needs further investigations.

45.4.3 *Influence of Circumferential and Radial Distribution of Spring Elements on the Critical Speed*

In this section, we present the influence of orientation of the distribution of discrete spring elements that are to approximate a given guide pad area, on the critical speed. We present results for the guide pad approximated by different number of discrete spring elements, distributed uniformly in the circumferential and radial directions for a given area. Results for a uniform distribution of discrete spring elements along the circumferential direction (with fixed radial location at 99.75 mm) are shown in Fig. 45.6, and results for uniform distribution of discrete spring elements along the radial direction (with fixed circumferential location at 20.5°) are shown in Fig. 45.7. Since damping was shown to not influence the critical speed, we neglect it herein. As before, the overall effective stiffness in each case is taken to be 7.5×10^8 N/m. As is evident from Fig. 45.6, as the number of discrete spring elements increases along the circumferential direction, the critical speed increases and becomes constant after $J = 8$. However, as is evident from Fig. 45.7, for the case of different discrete spring elements along the radial direction

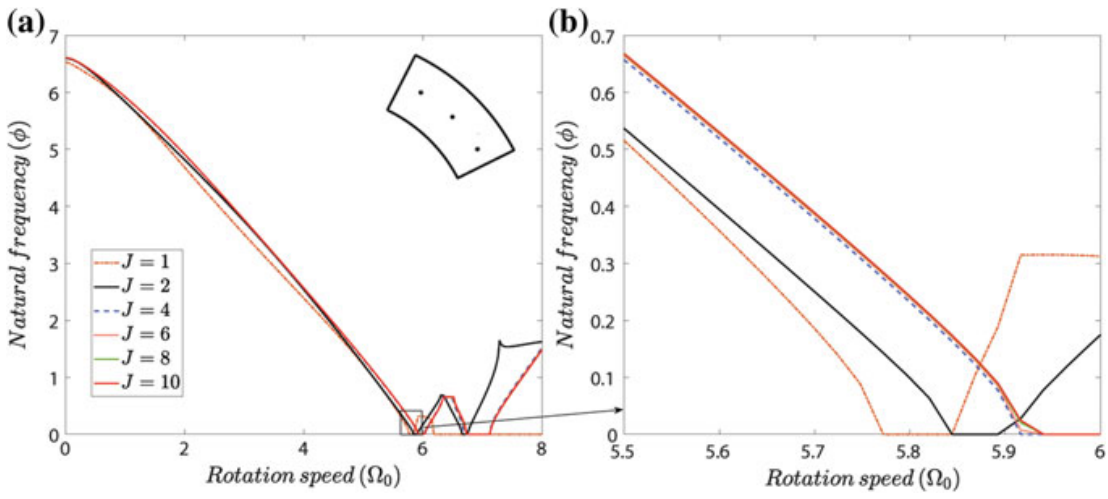


Fig. 45.6 Influence of multiple spring element over the guide pad in circumferential direction

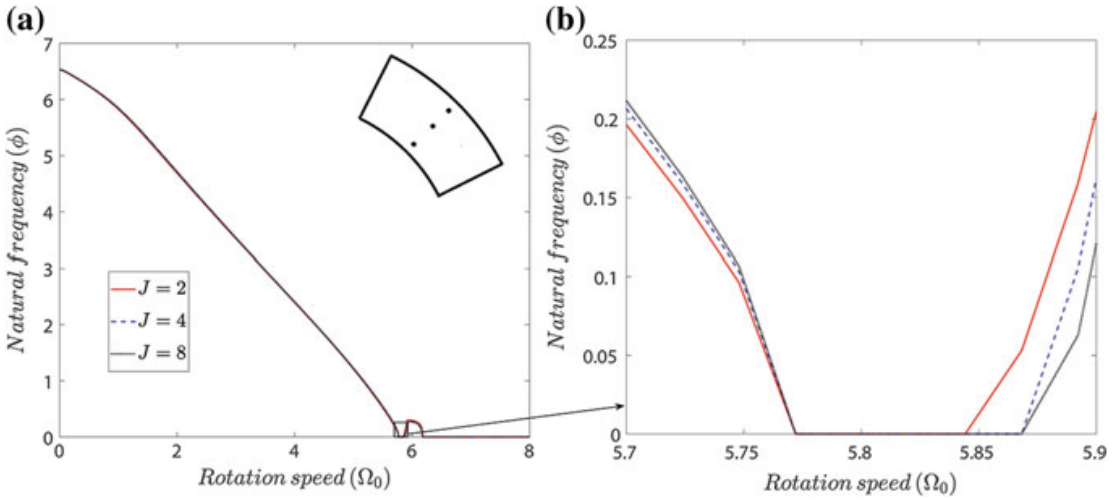


Fig. 45.7 Influence of multiple spring element over the guide pad in radial direction

at a fixed orientation, there is no observable change (increase) in critical speed with an increase in the discrete spring elements. We can hence conclude that for a given guide pad area modelled as discrete springs, an increase in discrete spring elements distributed along the circumferential direction leads to an increase in critical speed, whereas an increase in discrete spring elements distributed along the radial direction does not have any influence on critical speed.

These observations, which are different than those observed in the convergence analysis in Sect. 44.4.1, where we observed the critical speed to not change beyond 16 discrete spring elements approximating the same guide area, suggests that the approach to distributing discrete springs to approximate the guide area has a large bearing on the critical speed. Furthermore, we suspect that these results are not generalizable but need to be investigated separately for each guide pad area, and location, pointing to the need for a more generalized formulation that models the guide pads as distributed spring-damper systems, rather than multiple discrete spring-damper elements. Further investigations are also necessary to understand how these findings might change for the case of a saw with multiple guides.

45.5 Conclusion

This paper reports on the influence of guides on the critical speed of rotating saws. Since guides are known to improve critical speeds, understanding how they are modelled is essential to advise guide placement. Using numerically verified analytical models that include guides modelled as multiple discrete spring-damper elements, we observe that damping has no influence on critical speed, which is confounding, and needs further investigations. We also observe that for guides approximated with multiple discrete spring elements, the number, orientation, and location of the springs govern the increase in the critical speed, if any. We observe

that for a given guide area, it is necessary to carry out convergence analysis to understand how many discrete spring elements are actually necessary to approximate distributed guide and saw interactions. We also observe that critical speeds are very sensitive to guides modelled with a distribution of discrete spring elements along the radial direction and/or circumferential direction. These observations suggest that a more generalized formulation that models the guide pads as distributed spring-damper systems, rather than multiple discrete spring-damper elements are necessary. Experimental validation is also necessary and is planned.

Acknowledgements This research was supported by the Government of India's Science and Engineering Research Board's Early Career Research Award—SERB/ECR/2016/000619.

References

1. Schajer, G.S.: Guided circular saw critical speed theory." Proceedings SawTech 1989, Oakland, CA, USA
2. Mote, C.D., Nieh, L.T.: On the foundation of circular-saw stability theory. Wood Fiber Sci. **5**(2), 160–169 (1973)
3. Szymani, R., Mote, C.D.: Principal developments in thin circular saw vibration and control research. Holz als Roh-und Werkst. **35**, 219–225 (1977)
4. Iwan, W.D., Moeller, T.L.: The stability of a spinning elastic disk with a transverse load system. J. Appl. Mech. **43**(3), 485–490 (1976)
5. Hutton, S.G., Chonan, S., Lehmann, B.F.: Dynamic response of a guided circular saw. J. Sound Vib. **112**(3), 527–539 (1987)
6. Lehmann, B.F., Hutton, S.G.: Self-excitation in guided circular saws. J. Vib. Acoust. Stress. Reliab. Des. **110**(3), 338–344 (1988)
7. Mote Jr., C.D.: Moving-load stability of a circular plate on a floating central collar. J. Acoust. Soc. Am. **61**(2), 439–447 (1977)
8. A. Mohammadpanah and S. G. Hutton, "Flutter instability speeds of guided splined disks: an experimental and analytical investigation," Shock Vib. **2015** (2015)
9. Schajer, G.S.: CSAW 4.0 Computer software for optimizing circular saw design. Wood Mach, Institute, Berkeley, CA USA (2007)
10. Hagedorn, P., DasGupta, A.: Vibrations and waves in continuous mechanical systems. Wiley (2007)

Chapter 46
Investigations on the Influence of Radial Run-Out on Cutting Forces for Serrated Cutters

P. Bari, M. Law and P. Wahi

Abstract Serrated cutters with their complex local tool geometries result in continuously changing chip thickness and engagement conditions due to a change in local radius along the serration height. This change in geometry results in a reduction of cutting forces, and hence these cutters are favoured in the rough cutting of difficult-to-cut materials. The change in local radius of the cutter is also influenced by the radial run-out on the tool. Run-out, though undesirable, is inevitable. How this run-out influences the local radius, chip thickness, and forces is not entirely understood for serrated cutters. This paper fills that gap, by presenting a model that factors in the influence of radial run-out on cutting forces for serrated cutters. Investigations with different sets of cutting parameters show that for cutting with a low feed, and high axial and radial depths of cut, the influence of run-out on the cutting forces is negligible. For the case of cutting with high feeds however, we find that run-out adversely impacts cutting forces for any combination of axial and radial depths of cut. We also observe that with certain combinations of cutting parameters and run-outs, serrated cutters lose their advantage over regular end mills with similar levels of run-outs. Our findings suggest that serrated cutters may be used even with high levels of run-outs when the feed is low, and that their use should be avoided for any level of run-out when cutting with high feed rates.

Keywords Serrated cutters · Run-out · Chip thickness · Cutting forces

46.1 Introduction

Serrated cutters with their complex local tool geometries result in continuously changing chip thickness and engagement conditions due to a change in local radius along the serration height. This change in geometry results in a preferential

P. Bari · M. Law (✉) · P. Wahi
Department of Mechanical Engineering, Indian Institute of Technology Kanpur, Kanpur, Uttar Pradesh 208016, India
e-mail: mlaw@iitk.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_46

reduction of cutting forces, and hence these cutters are favoured in the rough cutting of difficult-to-cut materials. The change in local radius of the cutter is also influenced by the radial run-out on the tool. Run-out is characterized as the difference in the geometrical axis of the tool from the spindle's rotational axis. Run-out, though undesirable, is inevitable due to how the tool is clamped in the tool holder and also due to manufacturing tolerances and errors. How this run-out influences the local radius, chip thickness, and forces is not entirely understood for serrated cutters. The main focus of this paper is hence to incorporate run-out in predictive force models for serrated cutters and present systematic investigations on how different levels of run-out influence cutting forces for serrated cutters for different combinations of cutting parameters.

Since run-out changes the nature of cutting forces, contributes to surface roughness errors on machined workpieces, and also contributes to premature tool wear and potential breakage of the tool, a lot of research attention has been paid on modelling and understanding the influence of run-out for milling cutters. Early seminal work by Martellotti [1, 2] reported on the influence of run-out in milling processes. Later, Kline and DeVor [3] presented mathematical models incorporating the influence of run-out on force prediction for regular end mills. Zheng et al. [4] reported on face milling force profile changes with run-outs. Wan et al. [5] proposed new techniques to calibrate cutting force coefficients and run-out parameters in peripheral milling. The effect of run-out on cutting force variation and surface roughness generation was also studied by Krüger and Denkena [6], in which they identified run-out parameters from measured process forces using model-based run-out identification methods. Schmitz et al. [7] explored the effect of milling cutter run-out on surface topography, surface location error, and stability in end milling operations. More recently, Li et al. [8] addressed the issue of run-out in five-axis milling processes. Most literature addressing the influence of run-out in milling has focused only on regular end mills, and even though run-out is an important factor for end mills with serrations, the reported research on serrated end mills has focused more on modelling the serration geometry and its influences than on run-out.

Some classical early work on straight fluted serrated end mills was presented by Tlusty et al. [9], in which it was shown that due to the reduced contact between the tool and the workpiece on account of serrations, cutting forces reduce. Merdol and Altintas [10, 11], and Dombovari et al. [12] reported modelling of cylindrical and tapered serrated end mills with a generalized representation of serration profiles using cubic splines without considering the effect of run-out. Later, Campomanes [13] presented a detailed mechanics-based model that included the influence of helix with sinusoidal serrations, and used an approximate chip thickness model. He also made passing reference to how run-out does not have a significant effect on the cutting forces for serrated end mills at lower feed rates. His work, however, or the work of others before, and/or after him, does not support the claim based on model-based systematic investigations. Since serrated cutters preferentially reduce cutting forces, the influence of run-out on cutting forces, if any, and if run-out negates the positive influence of serrations in reducing forces, needs to be

understood comprehensively based on systematic model-based investigations and is the main motivation of the present work.

We present an expanded geometric model for serrated cutters that includes the influence of radial run-out, in Sect. 46.2. We limit our discussions in this paper to serrations only of the sinusoidal kind. We ignore axial run-out presently. We also assume that run-out is measurable on serrated tools. Measuring run-out on regular end mills is possible and has been reported on elsewhere; see, for example, [14–16]. However, measuring run-out on serrated cutters due to constantly changing radius along the serration wavelength is non-trivial and needs to be addressed separately. Assuming run-out to be measurable, we systematically investigate its influence on force profiles for different combinations of cutting parameters in Sect. 46.3. We systematically discuss the influence of run-out for different levels of feed, and axial and radial depths of cut. This is followed by the main conclusions in Sect. 46.4.

46.2 Geometric Model of Serrated Cutter with Run-Out

This section describes the geometry of serrated cutters considering the influence of radial run-out. Models described in this section build on the classical work done by Merdol and Altintas [11], and by Dombovari et al. [12], and also on our own earlier reported work [17], by including the influences of run-out.

46.2.1 Modelling of Serration Profile

A schematic and a cross-sectional view of a sinusoidal serrated cutter with run-out is shown in Fig. 46.1. Due to the run-out, cutter geometrical centre, denoted by O_g, deviates from spindle rotation centre O_r, by a constant radial deviation, ρ.. The run-out angle between direction of radial offset (deviation) and the nearest tooth at the bottom of the tool is denoted by δ. We assume that there is no axial run-out or any cutter geometric axis tilt; i.e. ρ is constant along the cutter height. Due to the run-out, since the whole system rotates about the spindle axis through O_r, it is convenient to model all things with respect to O_r, and hence the xyz coordinate frame is attached to O_r as shown.

The cutter can have N number of flutes (teeth), but as an example, only three (ith, $(i+1)$th, and $(i+l)$th) flutes are shown in Fig. 46.1. As there is a wavy surface along the flute of the serrated cutter, the local geometrical radius changes along the flute and the height. The local geometrical radius which is measured with respect to geometrical axis (dotted line through O_g shown in Fig. 46.1a) which is parallel to z-axis for the ith flute at the height z is defined as:

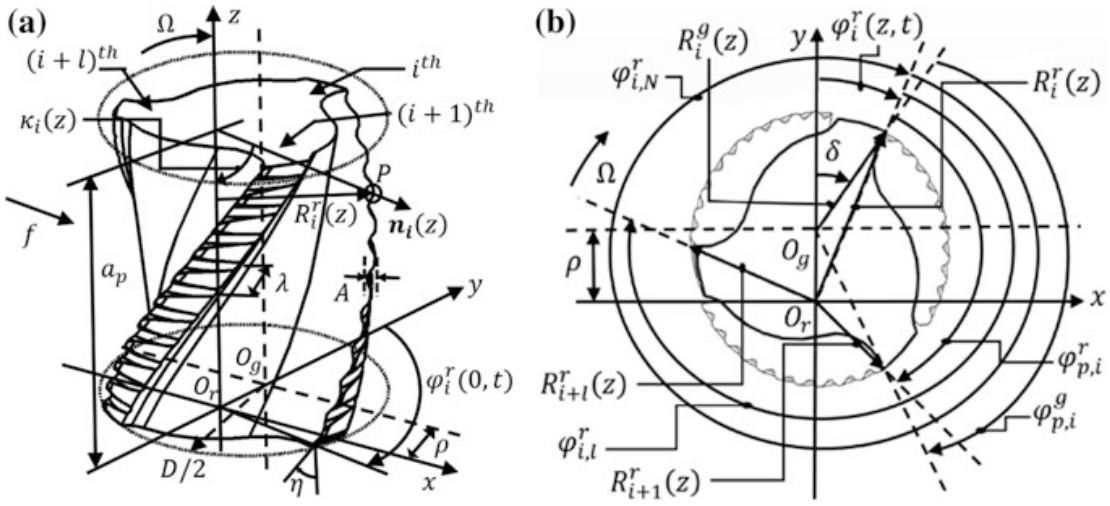


Fig. 46.1 **a** Geometry of the serrated cutter with run-out. Figure is adapted and modified from [17]. **b** Cross-sectional view at height z with run-out

$$R_i^g(z) = \frac{D}{2} - \Delta R_i^g(z) \tag{46.1}$$

where D is shank diameter (nominal diameter) of the cutter, and $\Delta R_i^g(z)$ is the variation in local geometrical radius for the sinusoidal serration profile:

$$\Delta R_i^g(z) = \frac{A}{2} - \frac{A}{2}\sin\left(\frac{2\pi z}{\lambda \cos\eta} - \psi_i + \frac{\pi}{2}\right) \tag{46.2}$$

where A is the serration amplitude, which is half of the peak-to-peak serration height; λ is the wavelength; η is the helix angle; and ψ_i is the phase shift due of the serrations on different flutes, explained in [17]. All geometries are measured with respect to spindle rotational axis which do not align with cutter geometrical axis due to eccentricity between them caused by the run-out. Due to the run-out, we define another local radius called the rotational radius with respect to the z-axis through O_r as follows:

$$\begin{aligned} R_i^r(z) = \Bigg[& \left(\rho + R_i^g(z)\cos\left(\delta - \sum_{k=1}^{i-1}\varphi_{p,k}^g - \frac{2z\tan\eta}{D}\right)\right)^2 \\ & + \left(R_i^g(z)\sin\left(\delta - \sum_{k=1}^{i-1}\varphi_{p,k}^g - \frac{2z\tan\eta}{D}\right)\right)^2 \Bigg]^{0.5} \end{aligned} \tag{46.3}$$

The local geometrical radius $R_i^g(z)$ for a representative four-fluted sinusoidal serrated cutter without and with run-out is shown in Fig. 46.2a and b, respectively. Cutter parameters are as follows: shank diameter D = 16 mm, serration amplitude

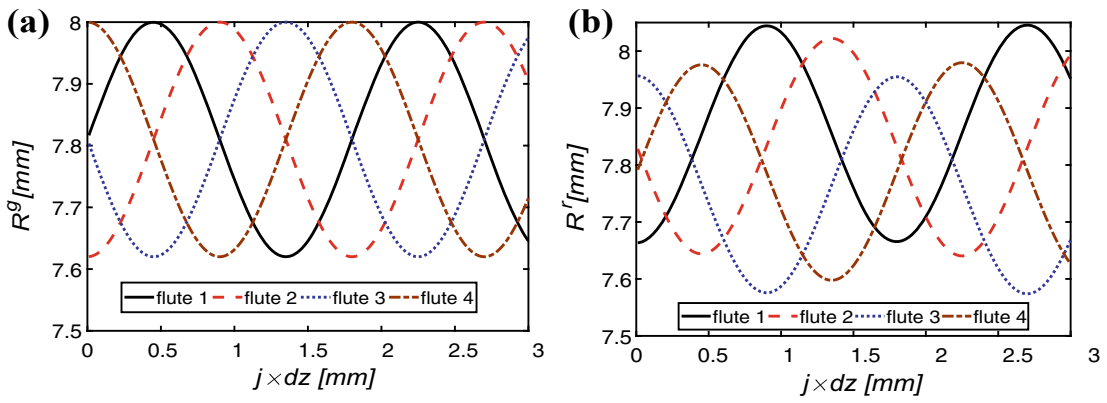


Fig. 46.2 **a** Variation of local geometrical radius, R^g, along height without run-out. **b** Variation of local rotational radius, R^r, along height with run-out: $\rho = 50\,\mu m, \delta = 30^\circ$

and wavelength $A = 0.38$ mm and $\lambda = 1.8$ mm; and, helix angle $\eta = 20^\circ$. Run-out is assumed as $\rho = 50\,\mu$m and $\delta = 30$. Comparing Fig. 46.2a and b, we see that run-out makes the radius irregular.

In addition to causing a change in the local radius, run-out also changes the instantaneous radial immersion angle. The angular position for the ith flute at height z, measured from the y-axis in a clockwise direction considering the run-out, called the instantaneous radial immersion angle, is calculated as follows

$$\varphi_i^r(z,t) = \Omega t + \sum_{k=1}^{i-1} \varphi_{p,k}^r - \frac{2z\tan\eta}{D} \tag{46.4}$$

where Ω is the clockwise spindle speed (rad/sec) and $\varphi_{p,k}^r$ is new pitch angle with respect to O_r due to the run-out. The relationship between the new pitch angle $\varphi_{p,k}^r$ with respect to O_r and the old geometrical pitch angle $\varphi_{p,k}^g$ with respect to O_g is shown in Fig. 46.3a. From Fig. 46.3a, using the cosine triangle formula, $\varphi_{p,k}^r$ is calculated as follows:

$$\varphi_{p,k}^r = \cos^{-1}\frac{\left(R_i^r(z)\right)^2 + \left(R_{i+1}^r(z)\right)^2 - \left(L_i^r(z)\right)^2}{2R_i^r(z)R_{i+1}^r(z)} \tag{46.5}$$

where

$$L_i^r(z) = \sqrt{\left(R_i^g(z)\right)^2 + \left(R_{i+1}^g(z)\right)^2 - 2R_i^g(z)R_{i+1}^g(z)\cos\varphi_{p,k}^g}$$

In addition to a change in the instantaneous radial immersion angle and pitch angles, run-out also changes the axial immersion angle, and that is calculated as detailed in [17]. Furthermore, the changed pitch angle, due to run-out, may result in a delay between formation of the current surface being generated and the previous

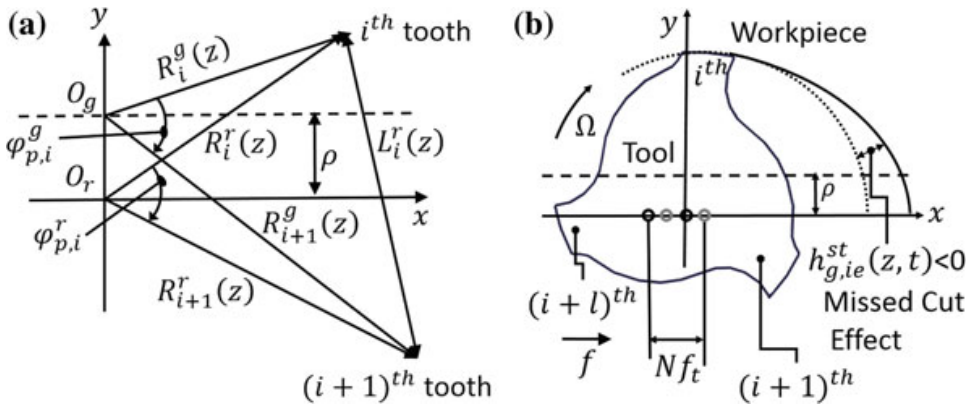


Fig. 46.3 **a** Calculation of the new rotational pitch angle, $\varphi_{p,k}^r$, due to the run-out on serrated cutters. **b** Missed cut effect with run-out for serrated cutter

surface generated, due to which redistribution of multiple delays may occur, and this can be factored in as discussed in [17] and as discussed next in Sect. 2.2.

46.2.2 Local Chip Thickness

Elemental physical static chip thickness is defined as the local distance between previous and current cut surfaces in the direction of normal vector $\boldsymbol{n}_i(z)$ of the flute considering run-out with circular tool path approximation as:

$$h_i^{st}(z,t) = g_i(z,t) \min_{l=1}^{N} \left[\left(R_i^r(z) - R_{i+l}^r(z) \right) + f_{i,l}(z,t) \sin\varphi_i^r(z,t) \right] \sin\kappa_i(z) \qquad (46.6)$$

where $f_{i,l}(z,t)$ is the corresponding feed motion during delay time $\tau_{i,l}$, $\kappa_i(z)$ is the axial immersion angle, $g_i(z,t)$ is the screening function due to radial immersion and missed cut (Ref. Fig. 46.3b) effect, explained in [17], and $R_i^r(z)$ is calculated from Eq. (46.3). The missed cut effect considering the run-out creates a non-uniform chip thickness profile as shown schematically in Fig. 46.3(b).

Variation in the chip thickness profile for a four-fluted sinusoidal serrated cutter considering zero run-out is shown in Fig. 46.4a, and considering run-out ($\rho = 50\,\mu\text{m}, \delta = 30^\circ$) is shown in Fig. 46.4b. The cutting conditions chosen are: depth of cut, $a_p = 3\,\text{mm}$; feed, $f_t = 0.25\,\text{mm/tooth}$; and speed, 5000 rpm, for slotting, i.e. 100% immersion. Comparing Fig. 46.4a and b, we see that due to the run-out, the flute chip thickness for the first and second flutes increases by 150%, and for the other two flutes, it decreases by 5%. Since chip thickness is different with run-outs, cutting forces which depend on chip thickness are also expected to be different.

The force model presented in [17] is used incorporating Eq. (46.6) to investigate the influence of run-out on force profiles, as discussed next, in Sect. 46.3. Furthermore, since the chip thickness also depends on the instantaneous entry and

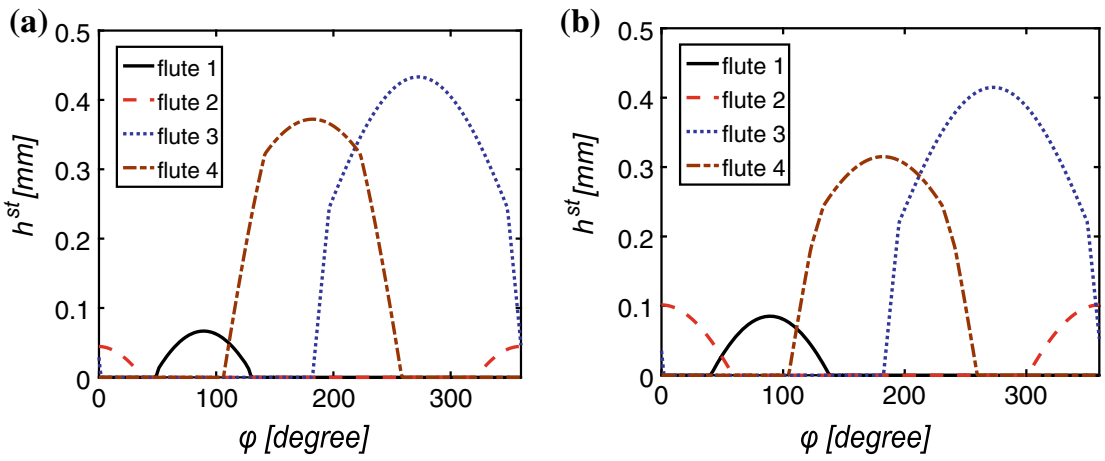


Fig. 46.4 Chip thickness (hst) distribution at tool tip ($z = 0$ mm) **a** without run-out and **b** with run-out: $\rho = 50\,\mu$m, $\delta = 30^\circ$

exit angles, these too will change with run-out, and they will also result in the force model being updated. However, for brevity, those formulations are not reported herein, since we observed that even without the updated entry and exit angle formulations, the force profiles do not change much.

46.3 Effect of Run-Out for Serrated Cutter at Different Operating Conditions

In this section, we discuss the influence of run-out on cutting forces for different combinations of cutting parameters. Our serrated cutting force model without run-out was numerically verified against other reported results and was also experimentally validated in our earlier reported work in [17]. We have also numerically verified our run-out model, albeit for the regular end mill case (i.e. by putting $A = 0$ in force model), using a commercially available advanced machining process simulation software, CutPRO [18], but have not reported those results here for brevity. Herein, we only present numerical investigations for our serrated model with run-out.

Although serrated cutters are more effective at reducing forces with low feeds than at high feeds, the effect of run-out is investigated for all possible operating conditions. Two types of feeds, one low (0.02 mm/tooth) and the other high (0.25 mm/tooth), are chosen. We present results with a low feed in Sect. 3.1 and present a subset of results with a high feed in Sect. 3.2. At each feed, two levels of engagement conditions, one low (10% up milling) and the other high immersion (100% slotting), are taken. Again, for each engagement condition, two levels of depths of cut, one low ($a_p = 3$ mm) and the other high ($a_p = 8$ mm), are chosen. Other cutting conditions for all cases are as follows: run-out angle $\delta = 30^\circ$, 5000 rpm, helix $\eta = 20^\circ$. For each case, cutting force is compared for zero run-out,

small run-out (5 μm), and large run-out (50 μm), as discussed in subsequent sub-sections. To also understand, how/if run-out on serrated cutters can potentially negate the positive influence of serrations, which is to otherwise reduce force, we also discuss in Sect. 46.3.3 results for regular end mills with run-outs.

46.3.1 *Influence of Run-Out at Low Feed (0.02 mm/tooth) Rates*

Effect of run-out on serrated cutting forces (F_x) at low feed (0.02 mm/tooth) rates is shown in Fig. 46.5. From Fig. 46.5, we observe that the ratio of maximum to average force increases with an increase in run-out. Maximum difference occurs between zero run-out and large run-out force profiles. For 10% up milling, at a low depth of cut ($a_p = 3$ mm) (Fig. 46.5a), the increase of the peak amplitude of up to 25% from the case of zero run-out is observed, and at a high depth of cut ($a_p = 8$ mm) (Fig. 46.5b), the increase of the peak amplitude of up to 12% is observed.

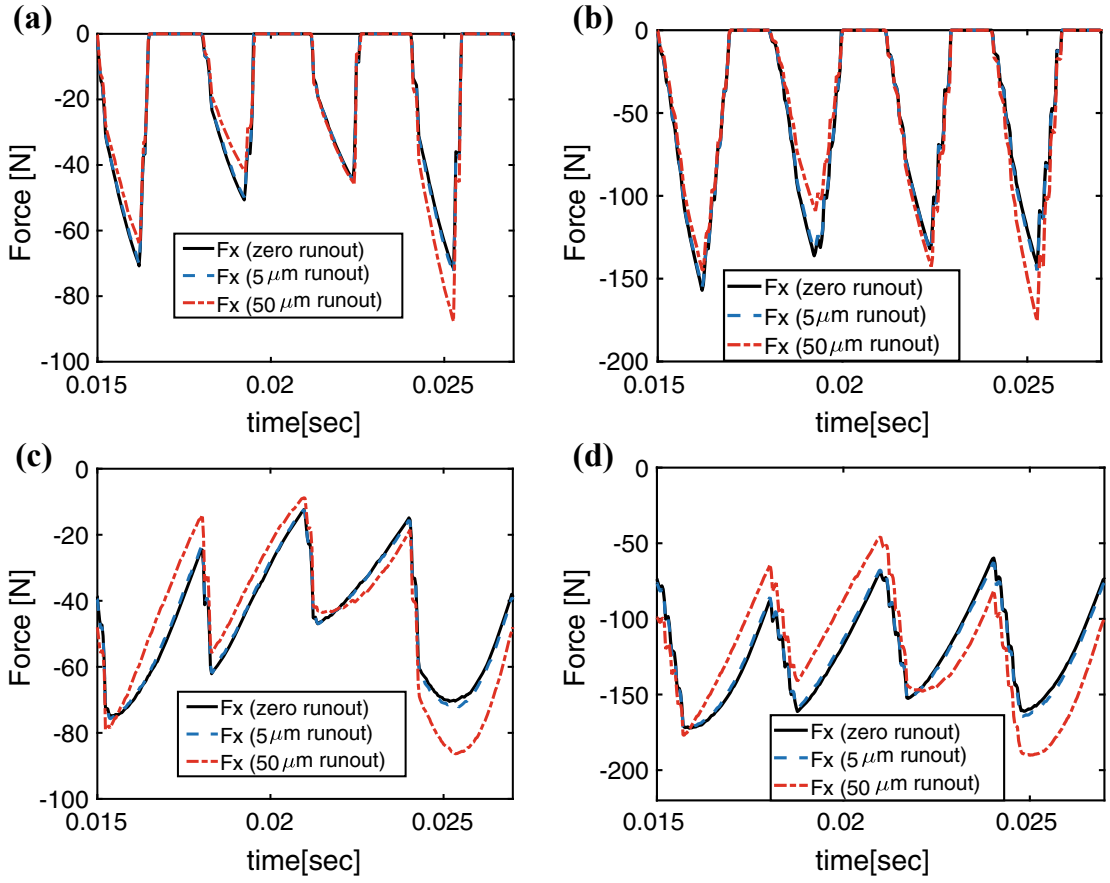


Fig. 46.5 Influence of run-out on cutting force (F_x) for a sinusoidal serrated end mill for low feed rates: **a** $a_p = 3$ mm, at 10% up milling; **b** $a_p = 8$ mm, at 10% up milling; **c** $a_p = 3$ mm, at 100% milling (slotting); and **d** $a_p = 8$ mm, at 100% milling (slotting)

For 100% milling (slotting) at a low depth of cut ($a_p = 3\,\text{mm}$) (Fig. 46.5c), the increase of the peak amplitude of up to 13% is observed, and at a high depth of cut ($a_p = 8\,\text{mm}$) (Fig. 46.5d), the increase of the peak amplitude of up to 10% is observed.

From these investigations, we can conclude that at low feeds, there is no substantial increase in the maximum force, or a significant change in the force profile, even with large values of run-out. This is interesting and is consistent with the insightful observations made in passing by Campomanes [13]. We also see that for the cases investigated, there is no significant change in the mean value of the forces, even with large run-outs on serrated cutters. These observations suggest that serrated end mills may be used even with large run-outs on them at low feeds, and potentially any combination of axial and radial depths of cuts, without resulting in any potentially significant damage to the tool, or the surface profile.

46.3.2 *Influence of Run-Out at High Feed (0.25 mm/tooth) Rates*

Effect of run-out on serrated cutting forces (F_x) at high feed (0.25 mm/tooth) rates is shown in Fig. 46.6. Here, we report on only a subset of the cutting combinations reported on in Sect. 46.3.1. As before, in this case too, we observe that the ratio of maximum to average force increases with an increase in run-out. For 10% up milling at low depth of cut ($a_p = 3\,\text{mm}$) (Fig. 46.6a), the increase of the peak amplitude of up to 35% is observed. For the 100% milling (slotting) at high depth of cut ($a_p = 8\,\text{mm}$) (Fig. 46.6b), the increase of the peak amplitude of up to 25% is observed.

Contrasting the results for high immersion with those of the low immersion case, we see that with high feed rates and for low immersion cutting, run-out appears to

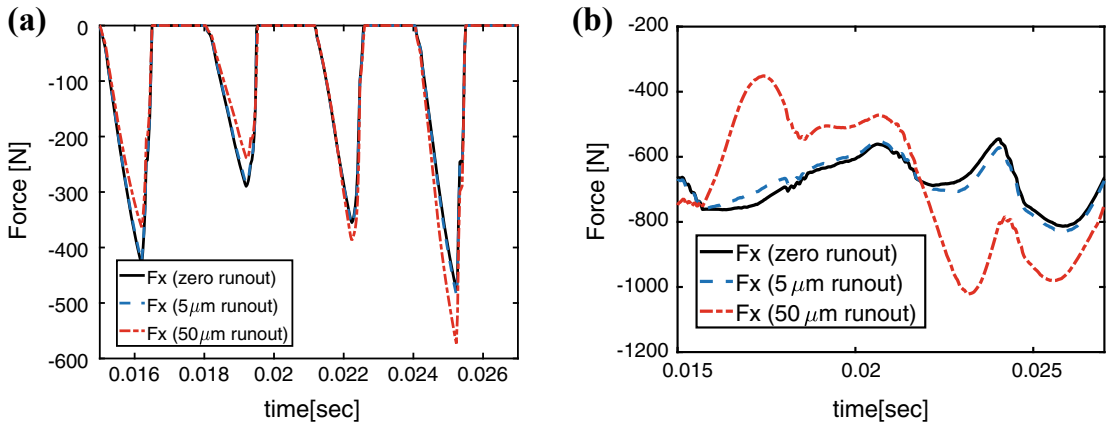


Fig. 46.6 Influence of run-out on cutting force (F_x) for sinusoidal serrated end mill for high feed rates: **a** $a_p = 3\,\text{mm}$, at 10% up milling, and **b** $a_p = 8\,\text{mm}$, at 100% milling (slotting)

have a more severe effect than at the high feed and high immersion case. Moreover, for the high feed and high axial and radial depths of cut, there is also a significant change in the force profile for high levels of run-out as compared to the case of zero run-out. These observations suggest that if serrated cutters have high levels of run-out, they best be avoided to be used with high feed rates. A sudden increase in forces, may cause impact loading on the cutter, which may cause it to prematurely break.

Seeing that serrated cutters with high run-outs are not advisable to use with high feed rates, to understand if/how run-out on serrated cutters used at high feed rates can potentially negate the positive influence of serrations, which is to otherwise reduce forces, we discuss results for regular end mills with run-outs next in Sect. 3.3 and contrast those findings with those observed here.

46.3.3 *Influence of Run-Out at High Feed (0.25 mm/tooth) Rates for Regular End Mills*

Effect of run-out on regular end mill cutting forces (F_x) at high feed (0.25 mm/tooth) rates is shown in Fig. 46.7. Here too, we report on only a subset of the cutting combinations reported on in Sect. 3.1. Results herein are obtained by setting $A = 0$ in the force model, i.e. essentially making it a regular end mill. All other geometries are the same as for the serrated end mill. As observed for serrated end mills, we observe that the ratio of maximum to average force increases with an increase in run-out. As evident in Fig. 46.7, for 10% up milling at a low depth of cut $(a_p = 3\,\text{mm})$ (Fig. 46.7a), the increase of peak amplitude of up to 37% is observed, whereas for 100% milling (slotting) at a high depth of cut $(a_p = 8\,\text{mm})$ (Fig. 46.7b), difference of the peak amplitude up to 46.5% is observed. These observations, as before, suggest that if regular end mills have high levels of run-out, they best be avoided to be used with high feed rates.

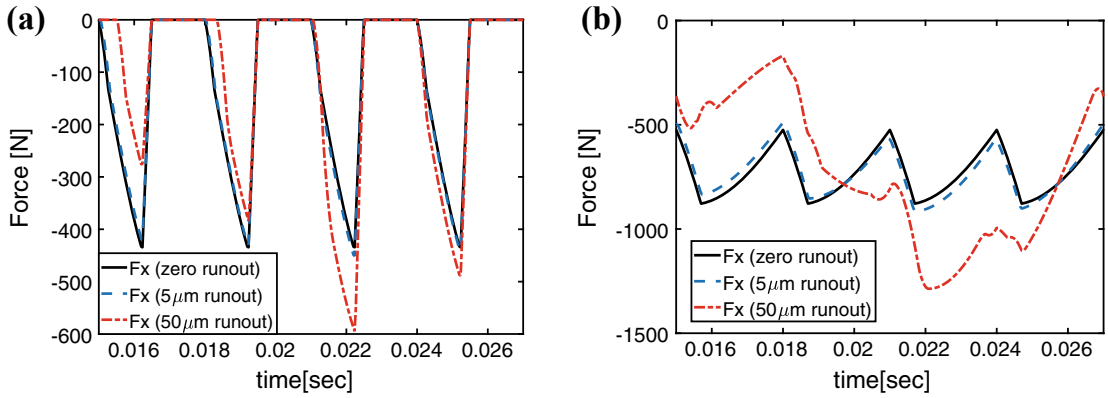


Fig. 46.7 Influence of run-out on cutting forces (F_x) for regular end mill at high feed rates: **a** $a_p = 3\,\text{mm}$, at 10% up milling, and **b** $a_p = 8\,\text{mm}$, at 100% milling (slotting)

Comparing the results for serrated cutters with run-out and with cutting at high feed and low immersion in Fig. 46.6a with those in Fig. 46.7a for regular end mills with run-out and high feed, with also low immersion, it is clear that even in the case of no run-out, the maximum force for the serrated end mill considered is 11% higher than regular end mills, which suggests that serrated cutters are ineffective at high feed rates, since there is no preferential reduction in cutting forces. The trend, however, reverses for the case with large run-out, which is strange, and needs to be investigated further. For the case of serrated end mill cutting with high feed and high immersion (axial, and radial), the peak force from Fig. 46.6b is lower than the case of regular end mills from Fig. 46.7b, without and with large run-outs. These observations are interesting and suggest that at certain combination of cutting and run-out parameters, serrated end mills retain their advantage over regular end mills, while at certain other combinations of cutting and run-out parameters, serrated cutters lose their advantage over regular end mills.

46.4 Conclusions

This paper discussed the influence of run-out on serrated cutters. We expanded the geometric force model for sinusoidal serrated cutters by including the effect of run-out. We systematically investigated how the run-out influences local chip thickness distribution, and how that in turn can potentially change the force profile for serrated cutters with run-out. Comprehensive investigations on the influence of run-out on forces were presented for a combination of feed rates, and axial and radial depths of cut. We observed that high levels of run-out increase the ratio of maximum to average forces and also change the nature of the cutting force profile. Interestingly, we observed that at low feeds, and large axial and radial depths of cut, run-out had no significant effect on the serrated force profiles, suggesting that even for large levels of run-outs, serrated cutters may still be used with low feed and high axial and radial depths of cuts, without worrying too much about if run-out may damage the workpiece quality or cause tool breakage. For cutting with high feed rates however, we observed that run-out adversely impacts cutting forces for any combination of axial and radial depths of cut. We also observed that with certain combinations of cutting parameters and run-outs, serrated cutters lose their advantage over regular end mills with similar levels of run-outs, and hence their use in such cases should be avoided.

Our findings were only for the case of sinusoidal serrations and having neglecting axial run-out. How run-out potentially influences force profiles for other serration types needs to be also investigated. Furthermore, numerical investigations presented herein need experimental validation, which will also include actually measuring run-out on serrated cutters, and these form part of our planned future investigations.

References

1. Martellotti, M.E.: An analysis of the milling process. Trans. Am. Soc. Mech. Eng. **63**, 677 (1941)
2. Martellotti, M.E.: An analysis of the milling process: Part II down milling. Trans. Am. Soc. Mech. Eng. **74**, 233 (1945)
3. Kline, W.A., DeVor, R.E.: The effect of runout on cutting geometry and forces in end milling. Int. J. Mach. Tool Design Res. **23**, 123–140 (1983)
4. Zheng, H.Q., et al.: Theoretical modeling and simulation of cutting forces in face milling with cutter runout. Int. J. Mach. Tools Manuf. **39**, 2003–2018 (1999)
5. Wan, M., et al.: New procedures for calibration of instantaneous cutting force coefficients and cutter runout parameters in peripheral milling. Int. J. Mach. Tools Manuf. **49**, 1144–1151 (2009)
6. Krüger, M., Denkena, B.: Model-based identification of tool run-out in end milling and estimation of surface roughness from measured cutting forces. Int. J. Adv. Manuf. Technol. **65**(5–8) (2013)
7. Schmitz, T.L., et al.: Runout effects in milling: Surface finish, surface location error, and stability. Int. J. Mach. Tools Manuf. **47**, 841–851 (2007)
8. Li, Z., et al.: Mechanistic modeling of five-axis machining with a general end mill considering cutter runout. Int. J. Mach. Tools Manuf. **96**, 67–79 (2015)
9. Tlusty, j., et al.: Use of Special Milling Cutters against Chatter. Proceedings of NAMRC 11, SME (1983)
10. Altintas, Y., Ber, A.: Manufacturing Automation: Metal Cutting Mechanics, Machine Tool Vibrations, and CNC Design (2001)
11. Merdol, S.D., Altintas, Y.: Mechanics and dynamics of serrated cylindrical and tapered end mills. J. Manuf. Sci. Eng (2004)
12. Dombovari, Z., et al.: The effect of serration on mechanics and stability of milling cutters. Int. J. Mach. Tools Manuf (2010)
13. Campomanes, M.L.: Kinematics and dynamics of milling with roughing tools. Metal Cutt. High Speed Mach (2002)
14. Wang, J.-J.J., Zheng, C.M.: Identification of cutter offset in end milling without a prior knowledge of cutting coefficients. Int. J. Mach. Tools Manuf. **43**, 687–697 (2003)
15. Hekman, K.A., Liang, S.Y.: In-process monitoring of end milling cutter run-out. Mechatronics **7**(1), 1–10 (1997)
16. Lee, K.Y., et al.: A run-out measuring method using modeling and simulation in four-fluted end milling. J. Mater. Process. Technol. **187–188**, 207–211 (2007)
17. Bari, P., Wahi, P., Law, M.: Investigations on the influence of serration parameters on cutting forces. Proceedings of the AIMTDR, 2018, India (to appear)
18. CUTPRO © Advanced Machining Process Measurement and Simulation Software (2000)

Chapter 47
Modeling and Optimization of WEDM Parameters Using Design of Experiments

P. Bharathi, G. Srinivasarao and P. G. Krishnaiah

Abstract In the current work, modeling and optimization of input parameters in wirecut electric discharge machine (WCEDM) of titanium [Ti-6A-4V] (Grade 5) were attempted. WCEDM working variables such as pulse-on-time (T_{ON}), pulse-off-time (T_{OFF}), servo-voltage (SV), peak-current (IP), and wire feed were considered to study the responses of cutting speed or speed of cut (CS) and wire wear ratio (WWR). Each input parameter was set at three levels. Experiments were conducted as per central composite face (CCF)-centered design. Based upon the runs' data, model equations for CS and WWR were developed using multiple linear regression. Optimization of practice parameters had been performed with the help of level mean analysis, response graphs, and model equations. From the analysis, it was observed that pulse-on-time was the substantial considerable input variable followed by servo-voltage, peak-current, and pulse-off-time.

Keywords WCEDM · Modeling · Optimization · Cutting speed · Wire wear ratio

47.1 Introduction

The WCEDM is a non-classical thermoelectric machining practice used for cutting newer, harder materials like titanium, composites, and ceramics. Titanium and its alloy are winning materials owing to their unique amalgamation of properties like high hardness and chemical wear resistance to substantial environments and tendency to form a stable, tightly adherent, protective surface film [1]. Due to the excellent properties of titanium alloys, the usage is being increased in applications

P. Bharathi (✉)
NRI Institute Technology Agiripalli, Krishna, AP, India
e-mail: nitbharathi307@gmail.com

G. Srinivasarao
R.V.R & J.C. College of Engineering Chowdavaram, Guntur, AP, India

P. G. Krishnaiah
PVPS Institute of Technology, Kanuru, Krishna, AP, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_47

like aerospace, chemical processing, surgical implants, and missile component [2]. Titanium and its alloys are the popular choice for fixation devices and are used to make artificial hips and knee replacement surgeries and other biological implants [3]. [Ti-6A-4V] is an alpha–beta alloy group, in which α-stabilizer possesses excellent creep resistance and β-stabilizers exhibit high hardenability and a higher density [4]. It is a substantial suitable process for machining very complex profiles of conductive hard materials where the high degree of precision and surface finish is the crucial factor. In WCEDM, material removal mechanism is very complex with a series of electric sparks generating the tremendous amount of heat between workpiece material and conductive tool which melts and evaporates minute amount of the workpiece and electrode material. The dielectric medium apart from providing the necessary dielectric strength continuously flushes away the solidified debris and also dissipates the heat from the machining zone. The tool wire (usually less than 0.3 mm diameter) is continuously fed from a supply spool, and the tool electrode is controlled and monitored by the CNC system. Researchers have focused on the input variables, pulse-on-time, pulse-off-time, servo-voltage, and peak-current, to avoid severe problems like the short-circuit of wire, wire fracture, and work exterior surface spoil leading to decrease in the overall efficiency of the process. Ghodsiyeh et al. [5] studied the effect of cutting parameters counting pulse-on-time, pulse-off-time, and peak-current on exterior surface roughness, sparking gap, and material removal rate of titanium (Ti-6Al-4V) alloy. Statistically the optimization model is the best manufacturing tool. Kaojy et al. [6] conducted the input variables' optimization of electric liberation cutting process of Ti-6Al-4V alloy considering multiple recital attribute wire wear ratio, material removal rate [MRR], and exterior surface coarseness using the Taguchi technique and gray relational coefficient grades. The process parameters chosen in this revision are discharge current, open voltage, pulse duration, and duty factor. The optimized process parameters prove an improved electrode wear ratio of 0.15, MRR of 0.12, and surface roughness of 0.19. Krishna et al. [7] built a hybrid replica by artificial neural networks [ANN] and applied genetic algorithm technique to optimize the exterior surface coarseness in electric discharge cutting of [Ti-6Al-4V] alloy by allowing for the concurrent effect of peak-current and voltage.

Xess et al. [8] performed concurrent optimization of the practice accomplishment material removal rate, surface roughness with aid of gray Taguchi's technique for cutting of titanium [Ti-6A-4V] alloy in EDM. Jabbaripour et al. [9] investigated the response distinctiveness of material removal rate [MRR], ratio of tool rapture [TWR], and diverse aspects of integrity of surface for Ti-6Al-4V on electrical discharge machining (EDM) specimens like topography of cutting surface, crack initiation, white layer (recast layer) height, and microhardness. Here, three process machining variables including pulse-current, pulse-on-time, and open-circuit voltage were altered through EDM runs. The effects of MRR and TWR versus process cutting variables were carried out based on foremost and combine outcome plots and too confirmed by analysis of variance results. The practical values exposed that broad aspects of exterior surface reliability for machined specimens are substantially exaggerated by pulse-current and pulse-on-time. Spedding and Wang et al.

[10] developed models through RSM and ANN to characterize the responses in WCEDM ultimate performance obtained by two techniques. Susanta et al. [11] applied weighted principal component [WPC] procedure to optimize the numerous outputs of the WCEDM process. The outcome offered good improvement in overall quality level by the WPC rather than other approaches. Rupesh et al. [12] had conducted studies on pure titanium with input variables such as pulse-on-time, pulse-off-time, peak-current, wire feed, wire tension, and servo-voltage. Dual-response optimization was performed on performance characteristics like cutting speed and surface coarseness. Anish-Kumar et al. [13] investigated on WCEDM using RSM with pure titanium as source material and copper alloy brass electrode. The responses of material removal rate, surface coarseness, dimensional deviation, and wire wear ratio were optimized by using multi-output optimization through desirability approach. Nayak et al. [14] calculated the optimal combination of variables for maximum MRR, minimum SR, and kerf width. Amitesh et al. [15] carried out studies on surface integrity, material removal rate, and wire wear ratio in WCEDM of Nimonic 80 alloys. Pulse-on-time and pulse-off-time were identified as considerable parameters. Further, single- and multi-response optimization were compared and validated. Advanced microscope, i.e., SEM analysis, concluded that the effect of the higher pulse-on-time causes height of recast layer. Sanjeev et al. [16] used the desirability function approach to optimize surface coarseness and velocity of cut while cutting alloys of composite (MMC). It was identified that the pulse width (PW), the time among pulses, small pulse time, and servo manage to mean reference voltage have the individual and interactive influence on the response variables. Srinivasarao et al. [17] conducted experiments on AISI 52100 steel using a CCF design and studied the effect of the input variables like pulse-on-time, pulse-off-time, de-ionized water pressure, and wire feed on surface roughness. Srinivasarao et al. [18] run the titanium Grade 5 material with CCF design with input variables like pulse-on-time, pulse-off-time, peak-current (IP), servo-voltage (SV), and wire feed (WF). Mathematical models were formulated for material subtraction rate and surface coarseness, and multi-objective optimization has been performed using desirability approach. Bharathi et al. [19] focused on some cutting limits T_{ON}, T_{OFF}, WF, and voltage while machining SS304 material for high MRR, lower SR, and kerf width using multi-objective optimization. Nihat et al. [20] investigated on wire wear ratio in WCEDM with the practice parameters like pulse duration, open-circuit voltage, wire speed, dielectric flushing pressure on 'AISI-4140' steel (work material), and brass wire (tool). From this study, it was observed that WWR increased by amplifying pulse duration and open-circuit voltage, while it is decreased by wire speed and non-conductive flushing pressure.

47.1.1 Materials and Processes

The work material used for experimentation is Ti-6A-4V. The annealed brass wire was used as a cutting tool with a diameter 0.25 mm and 1000 MPa tensile strength

Table 47.1 Element composition of the specimen

Element	N	C	Fe	O	Al	V	Ti
Composition	0.1	0.1	0.3	0.2	6.8	4.5	88.08

to manufacture very excellent, exact, and spotless cuts. The chemical composition of the workpiece is shown in Table 47.1 [21]. Workpieces in the form of the rectangle plate are considered for conducting the runs.

The speed of the cutting determined by dividing the cutting length by time is taken for each experiment. Wire wear ratio was determined as follows:

$$\text{wire wear ratio} = \left[\frac{\text{IWW} - \text{FWW}}{\text{IWW}}\right]$$

IWW = wire weight before machining
FWW = wire weight after machining

The tool electrode was weighted before and after the machining, and the weight reduction was determined to get the output response, i.e., WWR. An electronic weighing scale with 0.001 g accuracy was taken for weight dimension. The mean value was used to minimize the error in weighing.

47.1.2 Design of Experiments

Five input variables, specifically pulse-on-time, pulse-off-time, servo-voltage, peak-current, and wire feed, were accounted to revise the responses cutting speed (CS) and wire wear ratio [WWR]. Each input variables was located as three levels which are shown in Table 47.2. Experiments run using CCF design for five factors, and the CCF layout includes 28 experiments in which there are 16 fractional points, 10 axial points, and 2 center-points. All the experiments are carried out in random order in order to avoid bias. The runs along with responses, i.e., cutting speed and wire wear ratio, are shown in Table 47.3.

Table 47.2 Factors and corresponding levels

Factor	Factors	Level 1 (−1)	Level 2 (0)	Level 3 (+1)
T_{ON} (μs)	Pulse-on-time	107	113	119
T_{OFF} (μs)	Pulse-off-time	51	57	63
IP (amp)	Peak-current	10	11	12
SV (V)	Servo-voltage	30	40	50
WF (meter/minute)	Wire feed	3	6	9

Table 47.3 Experimental design with response

S. No	T_{ON}	T_{OFF}	IP	SV	WF	CS (mm/minute)	WWR
1	−1	−1	−1	−1	1	1.2245	0.011
2	−1	−1	−1	1	−1	0.59	0.008
3	−1	−1	1	−1	−1	1.943	0.003
4	−1	−1	1	1	1	0.9891	0.006
5	−1	1	−1	−1	−1	0.6645	0.006
6	−1	1	−1	1	1	0.05	0.004
7	−1	1	1	−1	1	0.805	0.006
8	−1	1	1	1	−1	0.2	0.001
9	1	−1	−1	−1	−1	1.653	0.009
10	1	−1	−1	1	1	1.152	0.003
11	1	−1	1	−1	1	3.22	0.009
12	1	−1	1	1	−1	2.1	0.022
13	1	1	−1	−1	1	1.059	0.077
14	1	1	−1	1	−1	0.741	0.013
15	1	1	1	−1	−1	2.5	0.009
16	1	1	1	1	1	1.179	0.009
17	−1	0	0	0	0	0.531	0.001
18	1	0	0	0	0	1.5	0.024
19	0	−1	0	0	0	2.865	0.011
20	0	1	0	0	0	1.151	0.008
21	0	0	−1	0	0	0.784	0.008
22	0	0	1	0	0	1.228	0.011
23	0	0	0	−1	0	1.9	0.016
24	0	0	0	1	0	1.152	0.006
25	0	0	0	0	−1	1.861	0.009
26	0	0	0	0	1	1.737	0.011
27	0	0	0	0	0	1.1	0.008
28	0	0	0	0	0	1.11	0.004

47.2 Examination of Data and Analysis of the Outcome

The foremost objective of this work is to examine the influence of variables on the response by means of response tables, response plots, and mathematical modeling. The main effects are studied from the factorial segment of the CCF layout which consists of the experiments from 1 to 16. One-factor-at-a-time analysis has been made using the experiments from 17 to 28. The performance of each control variable on the output assumed, i.e., cutting speed has been executed with level mean analysis. A level mean of a factor is the average of the cutting speed value of experiments in which the factor is at a particular level. For the factorial portion of CCF design, the mean value of the cutting speed for the T_{ON} at level '−1' and '+1'

can be determined by a mean response for runs 1–8 and 9–16, respectively. The average of the response for every level of the cutting variables can be computed in a similar manner. The control variable with the strongest influence is calculated by the difference between the average values of the factor at large and small levels. For one-factor-at-a-time analysis, the level means' values of T_{ON} are obtained by considering the experimental values for experiments 17, an average of 27 and 28, and 18 (i.e., 0.531, the average value of 1.1 and 1.11, 1.50). For T_{OFF}, the level means' values are obtained by considering the experimental values for experiments 19, an average of 27 and 28, and 20 (i.e., 2.865, the average value of 1.1 and 1.11, 1.151). From one-factor-at-a-time analysis, the change in each parameter is observed when the factor is changing from level '−1' to level '0' and to level '+1'. The level averages obtained from the fractional factorial part of CCF layout and one-factor-at-a-time analysis for cutting speed and wire wear ratio are represented in Tables 47.4, 47.5, 47.6, and 47.7 respectively.

Table 47.4 Level means: The factorial segment of [CCF] layout

Level	T_{ON}	T_{OFF}	IP	SV	WF
−1	0.8082	1.608	0.891	1.633	1.298
1	1.7	0.899	1.617	0.875	1.209
Difference	0.8918	0.709	0.716	0.758	0.089
Rank	1	4	3	2	5

Table 47.5 Level means: individual factor at a time scrutiny

	T_{ON}	T_{OFF}	IP	SV	WF
Level [−1]	0.531	2.865	0.784	1.9	1.861
Level [0]	1.105	1.105	1.105	1.105	1.105
Level [+1]	1.5	1.151	1.228	1.152	1.737
Difference	0.969	1.714	0.444	0.748	0.124
Rank	2	1	4	3	5

Table 47.6 Level means: the factorial part of CCF layout for WWR

Level	T_{ON}	T_{OFF}	IP	SV	WF
–1	0.0056	0.0089	0.0164	0.0163	0.0089
1	0.0189	0.0156	0.0081	0.0083	0.0156
Difference	0.0133	0.0067	0.0083	0.008	0.0067
Rank	1	4	2	3	4

Table 47.7 Level means: single-factor simultaneous analysis for WWR

	T_{ON}	T_{OFF}	IP	SV	WF
[Level. −1]	0.001	0.011	0.008	0.016	0.009
[Level. 0]	0.006	0.006	0.006	0.006	0.006
[Level +1]	0.024	0.008	0.011	0.006	0.011
Difference	0.023	0.003	0.003	0.01	0.002
Rank	1	3	3	2	4

47.2.1 Cutting Speed (CS)

The level means from Table 47.4 show that the T_{ON} is the substantial considerable parameters followed by servo-voltage, peak-current, and pulse-off-time. The wire feed rate is observed as not considerable. The same behavior was observed in the case of single-factor simultaneous analysis, but the pulse-off-time is identified as substantial considerable instead of the pulse-on-time. This might be due to the existence of combined effects (Fig. 47.1).

47.2.2 Wire Wear Ratio (WWR)

The level means from Table 47.6 shows that the T_{ON} is the considerable parameters followed by peak-current, servo-voltage, pulse-off-time, and wire feed rate. From Table 47.7, it can be observed from the single-factor simultaneous analysis that the pulse-on-time is identified as substantial considerable and remaining all factors are having more or less the similar effect. The reverse trend in case of pulse-off-time and peak-current has been identified, i.e., in the factorial portion if the pulse-off-time is escalating from level '−1' to level '+1', the wire wear ratio is increasing, whereas in one-factor-at-a-time analysis, if the pulse-off-time is escalating from level '−1' to level '+1' the wire wear ratio is decreasing. The similar trend can also be observed in the case of peak-current. From factorial portion when the peak-current is growing from starting level '−1' to end level '+1', the wire wear ratio is decreasing, whereas in one-factor-at-a-time analysis, if the peak-current is increasing from level '−1' to level '+1', the wire wear ratio is increasing. This is due to the existence of combined effects. The combined effects are studied by means of model equations and surface graphs.

47.2.3 Development of Mathematical Models

Based upon the experimental data, second-order models are developed for cutting speed and wire wear ratio. The data is given in the coded appearance of factors,

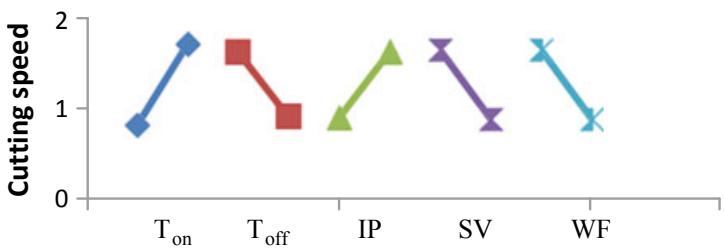


Fig. 47.1 Response plot for fractional factorial of CCF design for CS

i.e., minus one to plus one to SPSS software to estimate the model equation. The model equation for cutting speed is shown below:

$$\begin{aligned} \mathrm{CS} &= 1.425 + 0.45T_{\mathrm{ON}} - 0.41T_{\mathrm{OFF}} + 0.347\mathrm{IP} - 0.3795\mathrm{SV} \\ &\quad - 0.482(T_{\mathrm{ON}})^2 + 0.511(T_{\mathrm{OFF}})^2 - 0.491(\mathrm{IP})^2 \\ &\quad + 0.302(\mathrm{WF})^2 + 0.187T_{\mathrm{ON}}\mathrm{IP} - 0.121\mathrm{IPSV} \\ R^2 &= 0.928 \end{aligned}$$

The R^2, the value of 0.928, exhibits that variability in cutting speed is 92.8%, described by the model.

Based on the model, the response of the CS decreases as an increase in the T_{OFF} and SV and increases with an increase in T_{ON} and IP. The combination of IP with T_{ON} and SV is also identified as considerable. Further, the presence of square terms of the pulse-on-time, pulse-off-time, peak-current, and wire feed indicates the quadratic nature in output.

Combined effects mean the dual effect of two input variables on the response. The combined effect of pulse-on-time and peak-current on speed of cut gives that CS goes to a maximum value 1.436 mm/minute at larger value of pulse-on-time as well as peak-current at the same time as it reaches minimum value 0.158 mm/minute, where pulse-on-time is smaller level and peak-current is also smaller level as shown in Fig. 47.2. The combined effect of peak-current with servo-voltage on cutting speed gives that cutting speed of maximum value 1.781 mm/minute at the larger level of peak-current and smaller level of servo-voltage, while it reaches the least amount value of 0.845 mm/minute at the smaller level of peak-current and smaller level of servo-voltage. Decreasing SV decreases the gap between the electrode and work material that results in high ionization of dielectric medium and hence generates high discharge energy.

47.2.3.1 Optimum Values of Speed of Cut

The optimum values corresponding to the utmost cutting speed of 3.3719 mm/minute are the pulse-on-time 119 ms µs, pulse-off-time 51 ms µs, peak-current 12 A, servo-voltage 30 V, and wire feed 3 m/minute (Figs. 47.3 and 47.4).

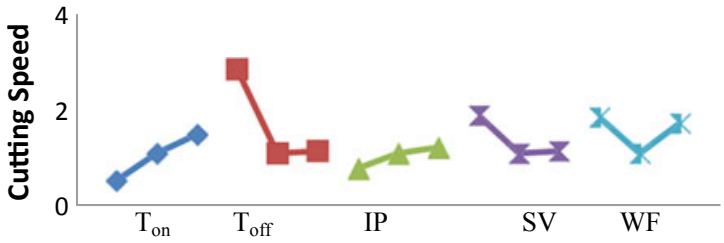


Fig. 47.2 Response plot for one-factor-at-a-time analysis for CS

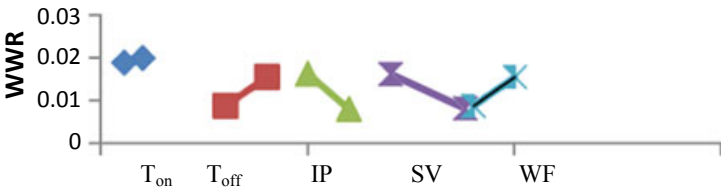


Fig. 47.3 Response plot for factorial of CCF design (WWR)

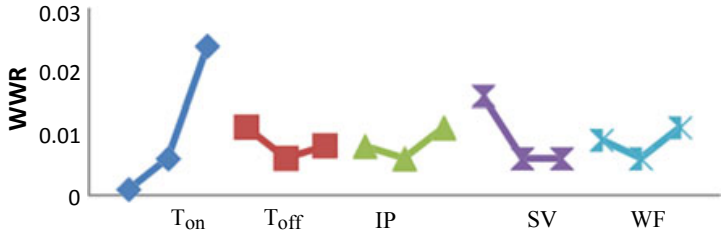


Fig. 47.4 Response plot for one-factor-at-a-time analysis (WWR)

The model equation for wire wear ratio is shown below:

$$\begin{aligned}\text{WWR} = {} & (0.01) + 0.004\,T_{\text{ON}} + 0.003\,T_{\text{OFF}} - 0.003\,\text{IP} - 0.004\,\text{SV} \\ & + 0.003\,\text{WF} + 0.003(T_{\text{ON}})^2 + 0.005\,T_{\text{ON}}\,T_{\text{OFF}} - 0.0\,T_{\text{ON}}\text{IP} - 0.003T_{\text{ON}}\text{SV} + 0.002T_{\text{ON}}\text{WF} \\ & - 0.005T_{\text{OFF}}\text{IP} - 0.005\,T_{\text{OFF}}\text{SV} + 0.005\,T_{\text{OFF}}\text{WF} + 0.005\,\text{IP}\,\text{SV} - 0.004\,\text{IP}\,\text{WF} \\ & - 0.006\,\text{SV}\,\text{WF} - (R^2 - 0.963)\end{aligned}$$

The R^2 value 0.963 implies that 96.3% of the variability in wire wear ratio is described by the model as given in the above equation (Figs. 47.5 and 47.6).

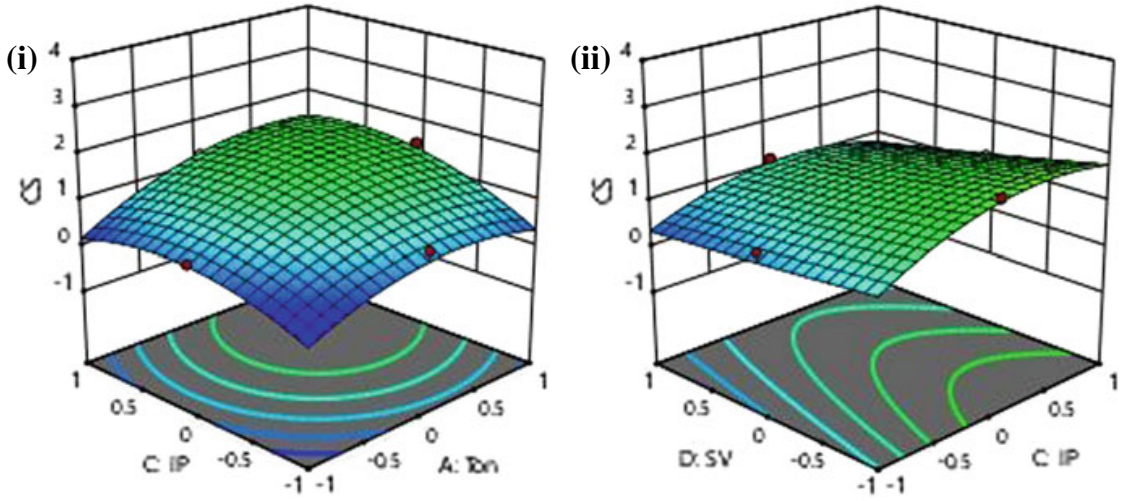


Fig. 47.5 (i) Effect of IP and T_{ON} on cutting speed (T_{OFF} = 0, SV = 0, WF = 0) and (ii) effect of IP and SV on cutting speed (T_{ON} = 0, T_{OFF} = 0, WF = 0)

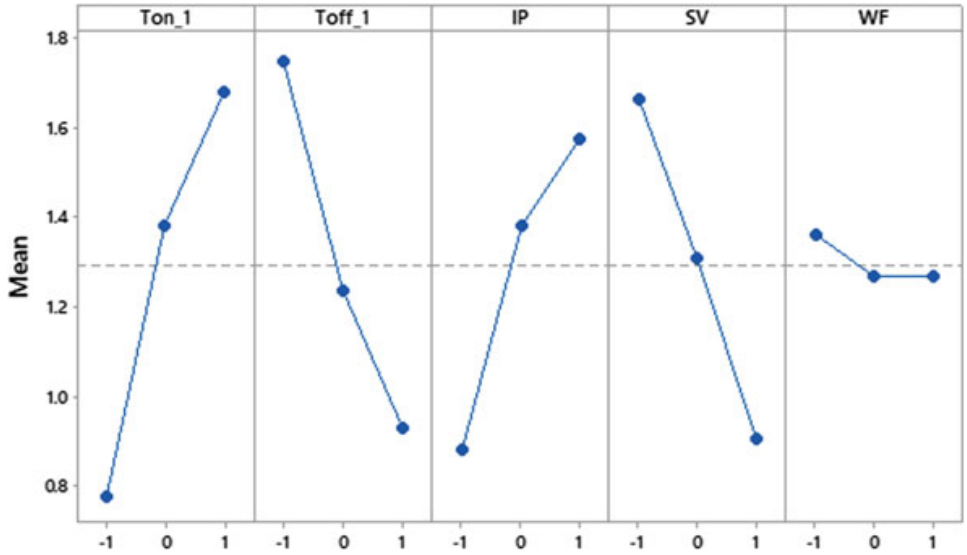


Fig. 47.6 Main effect plots for cutting speed

The combined effect of the pulse-on-time and pulse-off-time on WWR during their interaction gives that WWR goes to a maximum value of 0.028 at the higher value of pulse-on-time same as pulse-off-time, while it reaches the minimum value of 0.008, where the pulse-on-time and pulse-off-time were at lower level. The combined effect of pulse-off-time with peak-current on WWR shows that WWR's maximum value is 0.012 at the higher level of pulse-off-time and lower level of peak-current while it reaches minimum value 0.005 at the smaller level of pulse-off-time as well as lower level of peak-current. The optimum input variables equivalent to the minimum WWR 0.0099 were the pulse-on-time = 107 microseconds; pulse-off-time = 63 microseconds, peak-current = 12 A, servo-voltage = 30 V, and wire feed rate = 3 m/minute (Figs. 47.7 and 47.8).

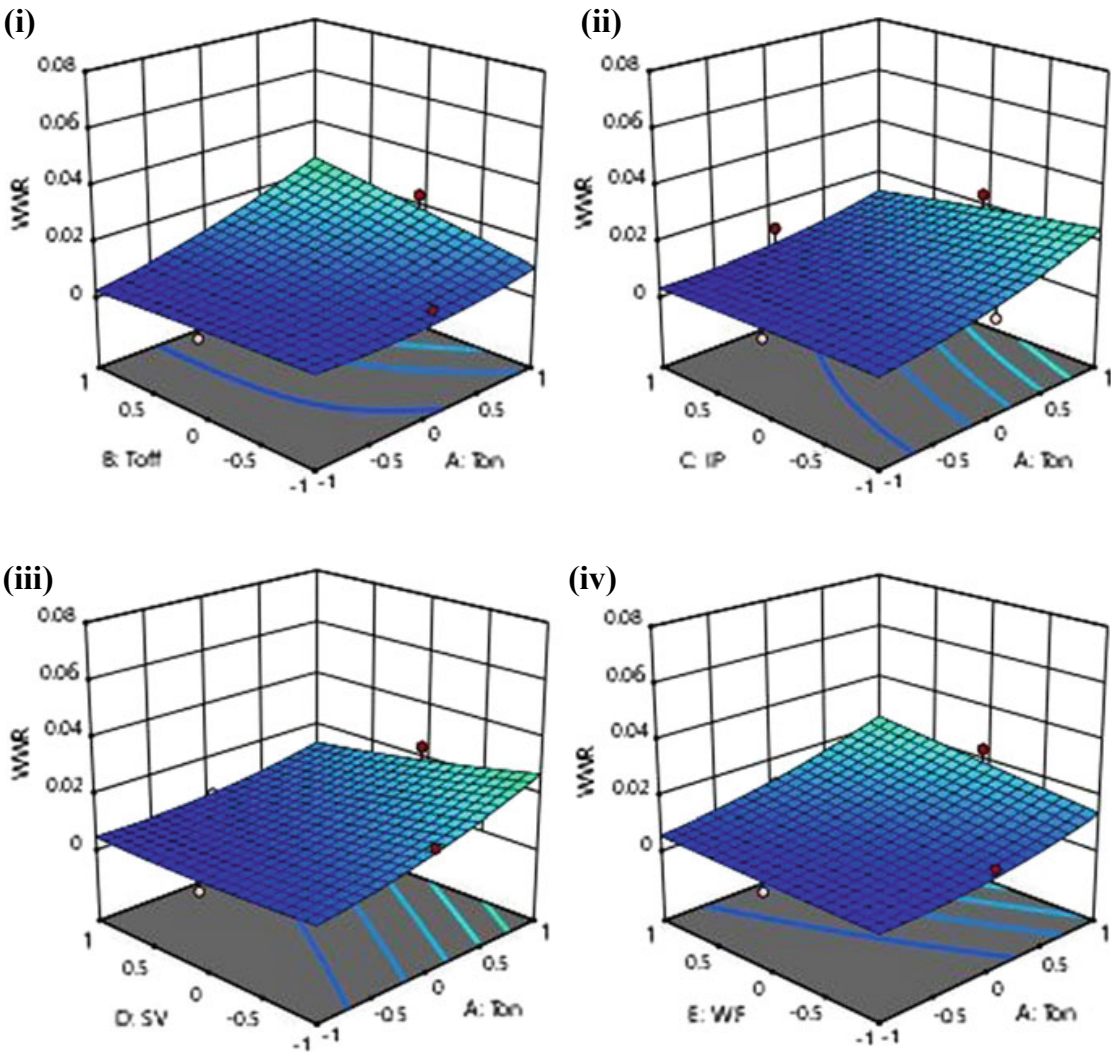


Fig. 47.7 (i) Effect of T_{OFF} and T_{ON} on WWR, (ii) effect of IP and T_{ON} on WWR, (iii) effect of SV and T_{ON} on WWR, and (iv) effect of WF and T_{ON} on WWR

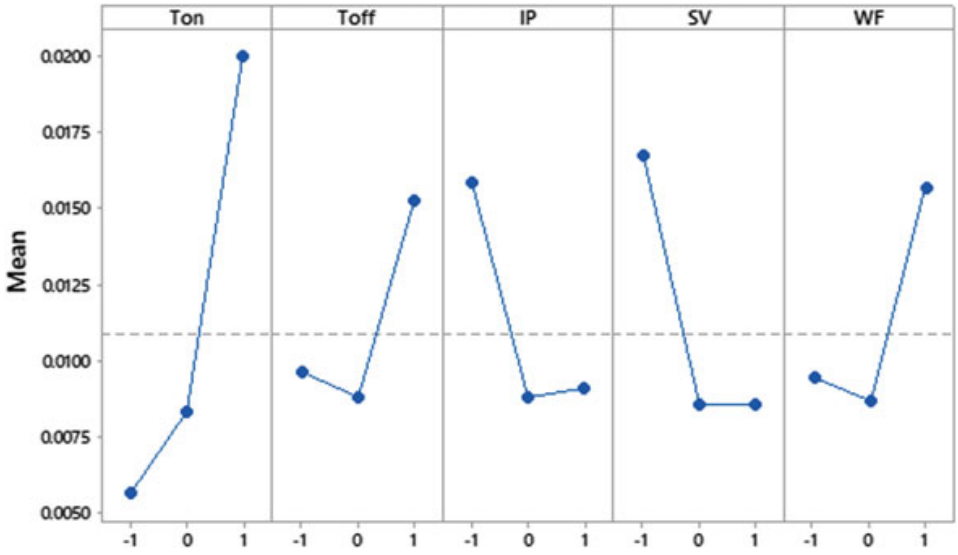


Fig. 47.8 Major outcome plots for wire wear ratio

47.3 Conclusions

Modeling and optimization of input variables have been performed in WCEDM of titanium [Ti-6A-4V] material with annealed brass wire as the electrode. The following observations are made:

1. All input variables having a considerable effect on cutting speed and wire wear ratio.
2. The factors pulse-on-time and peak-current have a positive effect, whereas pulse-off-time and servo-voltage have a negative effect on cutting speed.
3. The factors pulse-on-time, pulse-off-time, and wire feed have a positive outcome, whereas peak-current and servo-voltage have a negative effect on WWR.
4. The importance of interaction and second-order terms which shows the quadratic nature of the response is also observed on both responses.
5. The optimum values corresponding to the high cutting speed of 3.3719 mm/minute are the pulse-on-time 119 µs, pulse-off-time 51 µs, peak-current 12A, servo-voltage 30 V, and wire feed 3 m/minute.
6. The optimum values related to the minimum WWR 0.0099 were pulse-on-time = 107 µs; pulse-off-time = 63 µs, peak-current = 12 A, servo-voltage = 30 V, and wire feed rate = 3 m/minute

References

1. Yang, X., Richard Liu, C.: Machining titanium and its alloys. Mach. Sci. Technol. **3**(1), 107–139 (1999)
2. Myers, J.R., Bomberger, H.B., Froes F.H.: Corrosion behavior and use of titanium and its alloys. J. Metals (1984)
3. Sharma, A., Sharma, M. D., Sehgal, R.: Experimental study of machining characteristics of titanium alloy (Ti–6Al–4V). Arab. J. Sci. Eng. **38**(11), 3201–3209 (2013)
4. Catia S., Lopes, D., Mariana, T., Donato, Ramgi. P.: Comparative corrosion behavior of titanium alloys (Ti15MO and [Ti-6A-4V] for dental implants applications: a Review. Cor-Ros. Prot. Mater. **35**, 5–14 (2016)
5. Ghodsiyeh, D., Lahiji, M. A., Ghanbari, M., Shirdar, M. R., Golshan, A.: Optimizing Material Removal Rate (MRR) in WEDMing Titanium alloy (Ti6Al4V) using the Taguchi method. Res. J. Appl. Sci. Eng. Technol. **4**(17), 3154–3161 (2012)
6. Kao, J.Y., Tsao, C.C., Wang, S.S., Hsu, C.Y.: Optimization of the EDM variables on machining Ti–6Al–4V with multiple quality characteristics. Int. J. Adv. Manuf. Technol. **47**, 395–402 (2010)
7. Rao, K.M., Rangajanardhaa, G., Rao, H.D., Rao, S.M.: Development of hybrid model and optimization of surface roughness in electric discharge machining using artificial neural networks and genetic algorithm. J. Mater. Process. Technol. **209**(3), 1512–1520 (2009)
8. Xess, P.A., Sandhyarani, B., Manoj, M.: Optimization of the EDM variables on machining [Ti-6A-4V] alloy with multiple quality characteristics. Appl. Mech. Mater. **619**, 89–93 (2014)
9. Jabbaripour, B., Sadeghi, M.H., Faridv, S., Shabgard, M.R.: Investigating the effects of EDM variables on surface integrity, MRR and TWR in machining of Ti–6AL–4V. Mach. Sci. Technol. **16**, 419–444 (2012)
10. Spedding, T.A., Wang, Z.Q.: Study on modeling of wire EDM process. J. Mater. Process. Technol. **69**, 18–28 (1997)
11. Susanta, K.G., Shankar, C.: Optimization of multiple response for WCEDM processes using weighted principal components. Int. J. Adv. Manuf. Technol. **40,** 1102–1110 (2009)
12. Chalisgaonkar, R., Kumar, J.: Optimization of WCEDM process of pure titanium with multiple performance characteristics using Taguchi's DOE approach and utility concept. Front. Mech. Eng. **8**(2), 201–214 (2013)
13. Kumar, A., Kumar, V., Kumar, J.: Multi-response optimization of process parameters based on response surface methodology for pure titanium using WEDM process. Int. J. Adv. Manuf. Technol. **68**(9–12), 2645–2668 (2013)
14. Nayak, B.B., Mahapatra, S.S.: Multi-response optimization of WCEDM process variables us ing the AHP an TOPSIS method. **2**, 2319–3182 (2013)
15. Goswami, A., Kumar, J.: Investigation of surface integrity, material removal rate and wire wear ratio for WEDM of Nimonic 80A alloy using GRA and Taguchi method. Eng. Sci. Technol. Int. J. **17**(4), 173–184 (2014)
16. Garg, S. K., Manna, A., Jain, A.: An investigation on machinability of Al/10% ZrO_2 (P)-metal matrix composite by WEDM and parametric optimization using desirability function approach. Arab. J. Sci. Eng. **39**(4), 3251–3270 (2014)
17. Srinivasarao, G., Santhipriya, P., Suneel, D.: Modelling of surface roughness for AISI 52100 steel in WCEDM by Design of Experiments. Int. J. Eng. Res. Technol. **5** (2016)
18. Srinivasarao, G., Suneel, D.: Parametric optimization of WCEDM on α-β titanium alloy using desirability approach. Mater. Today Proc. **5**, 7937–7946 (2018)
19. Bharathi, P., Priyanaka, T.G.L., Srinivasa Rao, G., Nageswara Rao, B.: Optimum WCEDM Process variables of SS304 using Taguchi Method. Int. J. Ind. Manuf. Syst. Eng. **1**,69–72 (2016)
20. Tosun, N., Cogun, C.: An investigation on wire wear in WEDM. J. Mater. Process. Technol. **134**(3), 273–278 (2003)
21. ATI, Technical Data Sheet.: ATI [Ti-6A-4V], Grade 5, version 1, Pittsburg, USA, pp. 1–4 (2012)

Chapter 48
Analysis on Thermal Characteristics of Micro-Drilled Glass Using Microwave Energy at 2.45 GHz

Gaurav Kumar and Apurbba Kumar Sharma

Abstract Use of microwave energy at 2.45 GHz for micromachining offers promising nontraditional machining technique for brittle materials like glass. However, process is yet to be analyzed for better understanding of the process mechanism to minimize the defects associated with the thermal damages in the machined materials. In the present paper, the process was studied using the finite element method, and a 3D model of the microwave drilling setup was developed using COMSOL Multiphysics v5.2 software. Simulation was carried out to study the thermal characteristics of the materials with time while irradiated during the process. The main affecting parameters were analyzed to find out the possible reasons for the defect associated with this process. Simulation of the microwave drilling process was carried out for a combination of graphite concentrator (diameter = 500 μm) and borosilicate glass (thickness = 1.2 mm) at 700 W. Simulation results revealed that the electric field intensity near the concentrator tip was higher (approximately 8.99×10^6 V/m) enough to ionize the air dielectric media which causes the generation of plasma around the tip of concentrator. The maximum temperature in the machining zone was observed approximately 1100 °C in 6 s. The rapid rise in temperature in the specimen induces a very high thermal stress (approximately 85 MPa) which reaches beyond the fracture strength of the borosilicate glass. The experimental investigation of the same concentrator-specimen combination demonstrated the similar pattern of fracture while drilling in the air using the same processing parameters.

Keywords Microwave drilling · Glass · Stress

G. Kumar
NIT Uttarakhand, Srinagar (Garhwal) 246174, Uttarakhand, India

G. Kumar · A. K. Sharma (✉)
Microwave Material Processing Laboratory, IIT Roorkee, Roorkee 247667, Uttarakhand, India
e-mail: akshafme@gmail.com; akshafme@iitr.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_48

48.1 Introduction

Microwave drilling is a thermal energy-based advanced machining method in which electromagnetic energy concentrated at the tip of the concentrator is used to ablate/melt the material from the workpiece just beneath the concentrator. Microwave drilling was reported successfully for the first time by Jerby et al. in the year 2002. An extendable monopole antenna was used to concentrate the microwave into a small hot spot with the help of a coaxial waveguide. Due to the creation of hot spot, material became soft and antenna was made to penetrate inside the softened material to create a hole [1–3]. Later on, they reported that microwave drilling takes place due to thermal run-away effect at the tip of metallic drill. Theoretical analysis of thermal run-away effect was done by coupling the Maxwell's and heat equations to understand the temperature variation of hot spot with respect to time. It was reported that temperature of hot spot rises very rapidly [4]. In the year 2015, Lautre et al. performed drilling on various materials using microwave energy in a domestic microwave applicator using gravity feed. Drilling on various materials like polymer, glass, etc. was reported [5, 6]. But, a significant amount of thermal damage was present around the hole. To understand the mechanism, the process was simulated using COMSOL Multiphysics. It was reported that the jamming of the concentrator was taking place due to solidification of glass melt as the upper plasma sphere was relatively cooler. The high temperature of the plasma sphere around the concentrator tip caused it to melt and get blunt. So, as the cylindricity or conicity of the concentrator tip is lost, plasma formation at concentrator tip stops [7]. Later on, precursors were used to reduce thermal shock on the glass specimen but problems of concentrator jamming and blunting were still present [8]. In the year 2018, Kumar and Sharma used dielectrics to reduce thermal damage and avoid problem of concentrator melting and jamming while drilling glass [9].

Finite element-based modeling has been used for predicting the liquid and solid-state material transformation, the transient temperature distribution, and residual stresses induced in the workpiece during thermal energy-based machining processes [10–12]. Dhanik and Joshi studied the effect of inter-electrode gap on plasma temperature. It was found that temperature lies within 8100 ± 1750 K with a change in the inter-electrode gap [11]. Jiao and Wang simulated the laser cutting of glass plate to study the temperature and thermal stress distribution using ANSYS. It was reported that a dual-laser beam method where a laser beam was focused to machine the undesired material from the glass and an off-focused CO_2-laser beam was used to preheat the glass sample helped in reducing the thermal gradients. Due to reduced thermal gradient, the thermal stress became less than the critical fracture strength of the glass [12]. Krötz et al. simulated the single discharge of electrochemical arc machining for analyzing the heat-affected zone. It was found that duration of discharge and heat transfer from discharge to workpiece governs the diameter of heat-affected zone rather than the arc spot diameter [13].

Therefore, the objective of the present study is to simulate the microwave drilling process to understand the mechanism of the process and to find out the reason behind thermal damage induced in the workpiece during machining.

48.2 Modeling and Simulation

Figure 48.1 shows the experimental setup of microwave drilling. The pin vice is made up of structural steel and is used to hold concentrator. The material of the concentrator is also of structural steel. Microwave drilling is achieved by concentrating electric field at the tip of a metallic concentrator which rests just above the top surface of the workpiece (borosilicate glass) with the help of pin vice. As the electric field strength around concentrator tip reaches above the ionization potential of the medium around it, plasma is formed at the concentrator tip as shown in Fig. 48.2. The plasma formed ablates/melts the material from the workpiece just beneath it and the concentrator penetrates inside the workpiece to create a hole. But, it is very difficult to have a clear understanding of the process mechanism of microwave drilling due to the limitation in experimental assessment of distribution of electric field inside the cavity, temperature, and thermal stress along the surface of the workpiece, etc. So, the software tool like COMSOL Multiphysics can help in having a better understanding of the process mechanism.

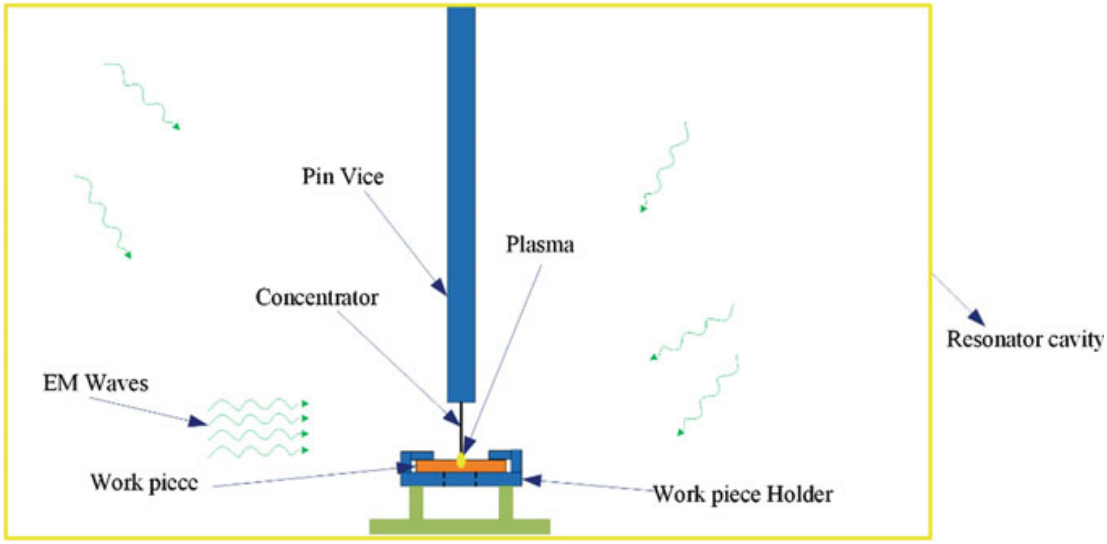


Fig. 48.1 Schematic diagram of the microwave drilling experimental setup

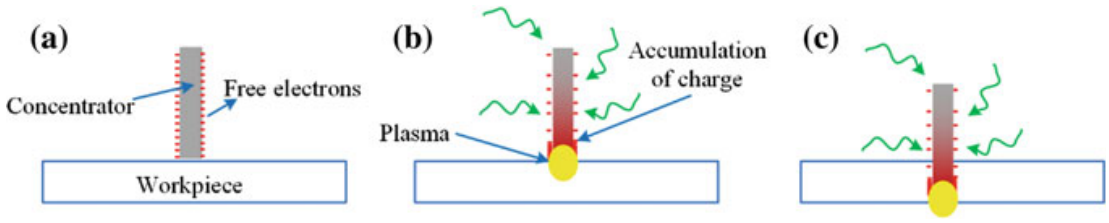


Fig. 48.2 Schematic of: **a** concentrator resting just above workpiece, **b** plasma formation after the microwave oven is switched on, and **c** concentrator penetrating inside the workpiece

48.2.1 Model Development

A 3D model of the experimental setup of microwave drilling as shown in Fig. 48.3 was developed using COMSOL Multiphysics 5.2 software to analyze the process and the behavior of workpiece while machining. The dimensions of the above model are shown in Table 48.1.

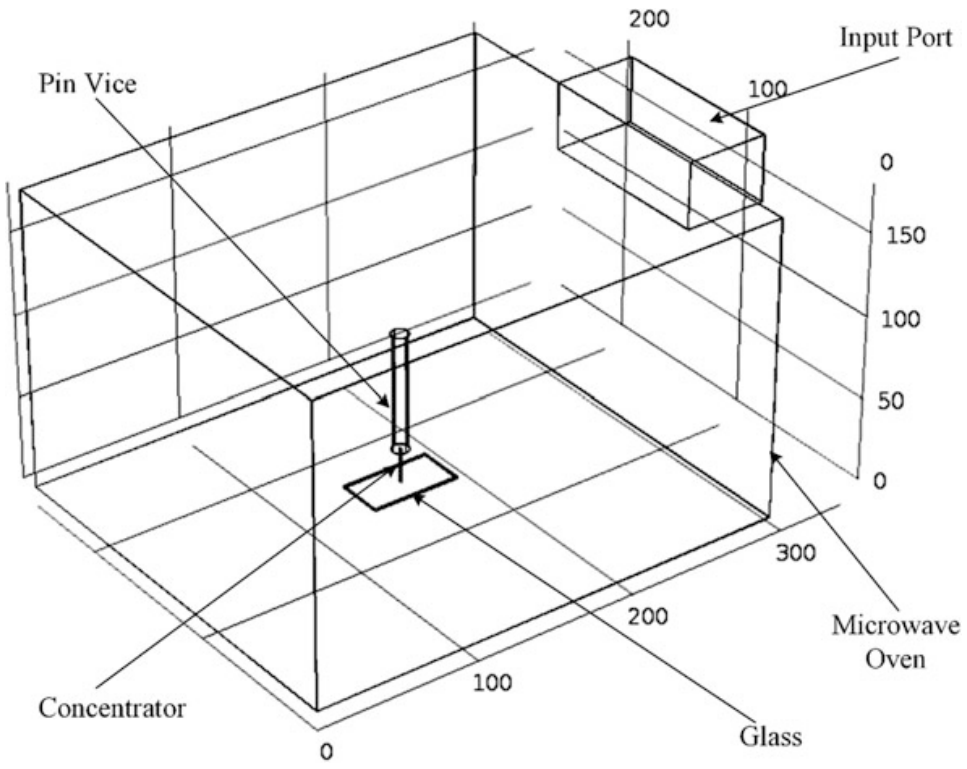


Fig. 48.3 Schematic of the 3D model of the microwave drilling

Table 48.1 Dimensions of the component used to generate the 3D model

Component	Oven	Waveguide	Pin vice	Concentrator	Workpiece
Dimensions (mm)	292 × 256 × 180	110 × 40 × 50	70 × 4.2(ϕ)	20 * 0.5(ϕ)	53 × 28 × 1.2

48.2.2 Meshing

Physics-controlled meshing is used to mesh the 3D model of the microwave drilling. Figure 48.4 shows the meshed 3D model. Fine meshing has been used to mesh the cavity (Fig. 48.4a), and extremely fine meshing has been used to mesh the concentrator and workpiece (Fig. 48.4b). The number of vertex elements, edge elements, triangular elements, and tetrahedral elements is 40, 557, 4504, and 45,401, respectively. The degrees of freedom solved for are 292,505. Figure 48.4c shows the plot of element quality of meshed geometry. The domain element statistics are as follows: average element quality = 0.653; element volume ratio = 1.924×10^{-7}; mesh volume = 1.368 × 107; average growth rate: 2.077.

48.2.3 Assumptions

The phenomenon of microwave drilling is governed by the formation of plasma at the tip of the concentrator and its subsequent interaction with the workpiece material. The formation of plasma depends upon the coupling of electromagnetic waves with the material of workpiece and concentrator. Besides it, shape of concentrator and its absorption behavior also influence the intensity of plasma formed. Therefore, following assumption has been made in order to simplify the model: (a) The workpiece, concentrator, and concentrator holder material used in the study are homogeneous and isotropic. (b) The atmosphere temperature inside the system is considered to be 27 °C, initially. (c) The air domain and jigs used are not

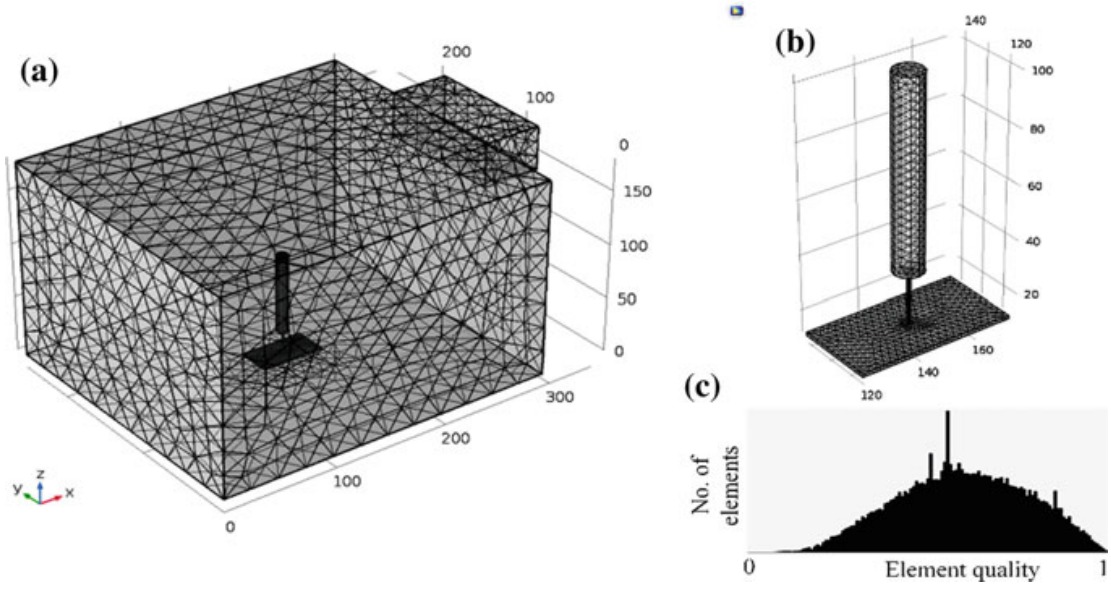


Fig. 48.4 Image showing: **a** the fine-meshed resonating cavity, **b** extremely fine-meshed workpiece and concentrator, and **c** plot of element quality of meshed geometry

Table 48.2 Property of the material of different component of experimental setup

Property	Materials		
	glass	Steel	Copper
Density (kg/m^3)	2210	7850	8960
Thermal Conductivity (W/(m × K))	1.4	44.5	400
Heat Capacity at constant pressure (J/(kg × K))	730	475	385
Relative permeability	1	1	1
Relative permittivity	4.2	1	1
Electrical conductivity (S/m)	1×10^{-14}	4.032×10^{6}	5.998×10^{7}
Young's modulus (N/m^2)	70×10^{9}	200×10^{9}	110×10^{9}
Poisson's ratio	0.22	0.30	0.35
Coefficient of thermal expansion(1/K)	3.3	12.3×10^{-6}	17×10^{-6}

considered for thermal analysis as they have zero dielectric losses. (d) The walls of the cavity and waveguide are considered to be made up of copper. (e) The port is excited by a transverse electromagnetic standing wave fields in the Z direction. The property of the material of the different component of the experimental setup is shown in Table 48.2 which has been taken from the material library of the software.

48.2.4 Governing Equations

COMSOL Multiphysics (version 5.2), a finite element-based software has been used to solve the governing equations. The phenomenon of interaction of microwave with the concentrator and glass can be explained with the help of Maxwell's equations. The following equation based on standard Maxwell's EM equations with the stipulated excitation at the rectangular port is solved for electric field ($\boldsymbol{E}$) inside the waveguide and oven:

$$\nabla \times \left(\mu_r^{-1} \nabla \times \boldsymbol{E}\right) - k_0^2 \left(\varepsilon_r - \frac{j\sigma}{\omega \varepsilon_0}\right) \boldsymbol{E} = \boldsymbol{0} \tag{48.1}$$

where μ_r = relative permeability, k_0 = wave number in the free space, ε_r = relative permittivity, j = imaginary unit, ω = angular frequency, σ = conductivity, and ε_0 = permittivity of free space.

Due to the externally applied electromagnetic field, a very intense electric field ($\sim 10^6$ V/m) around the tip of the concentrator gets developed. This high electric field propagates through dielectric materials by ionizing it and creating a hot spot just beneath the tool tip. The heat generation in the concentrator, at the interface of concentrator and workpiece and in the dielectric media, can be expressed by the heat transfer equation as:

$$\rho C_p \frac{\partial T}{\partial t} = \nabla.(k\nabla T) + P(r,T) - P_{\text{Loss}}(T_S) \tag{48.2}$$

where t = time (s), ρ = material density (kg m^{-3}), C_p = specific heat capacity (J kg^{-1} K^{-1}), and k = thermal conductivity (Wm^{-1} K^{-1}), $P(r,T)$ = power absorbed by the dielectric, $P_{\text{Loss}}(T_S)$ = power lost to the surrounding due to convection and radiation respectively.

Due to the creation of hot spot on the surface of the tool tip, a thermal gradient is developed on the surface of workpiece. Due to this thermal gradient, a thermal stress is induced inside the surface of workpiece which is expressed using the following equation:

$$\sigma_{\text{th}} = \frac{E\alpha}{1-\vartheta}(\Delta T) \tag{48.3}$$

where σ_{th} = residual thermal stress (MPa), E = Young's modulus, ϑ = Poisson's ratio and α = coefficient of temperature expansion, ΔT = temperature differential on the surface of the workpiece.

48.2.5 Validation

The result obtained after simulating the above experimental condition was validated experimentally. Figure 48.5a shows the optical image of a hole drilled in air. It is evident from the figure that circular and radial cracks are present around the hole. These cracks originate from the critically stressed zone and propagate through the materials around it. It is evident from the figure that an area of $\sim$4.09 mm^2 is critically stressed. Figure 48.5b shows the simulated result of the distribution of around the hole created. It is evident that an area of 3.878 mm^2 is having Von

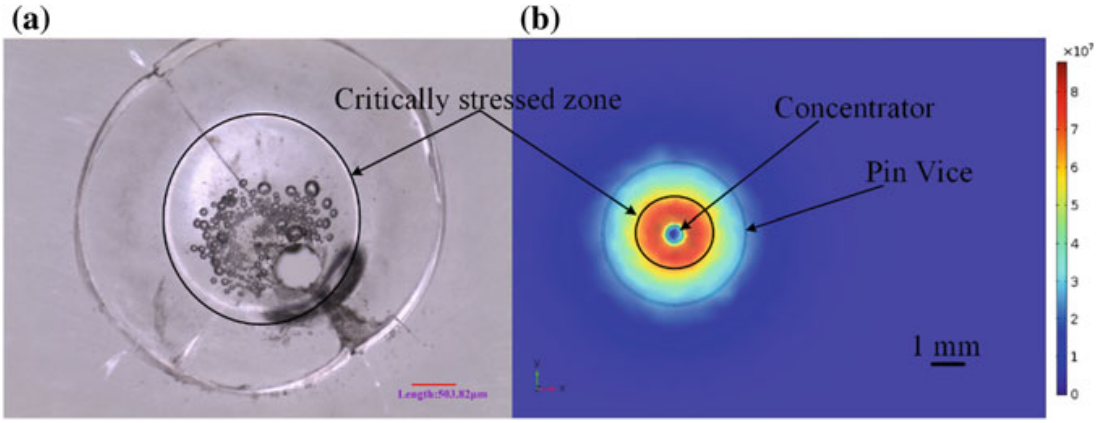


Fig. 48.5 **a** Optical image of the workpiece drilled in air and **b** Top view of the simulated stressed zone around the hole drilled on the borosilicate glass

Mises stress value greater than 60 MPa (critical stress limit for fracture of glass) which is called critically stressed zone. Both the experimental and simulated results are almost close to each other. Thus, the above-proposed model can be used for further study.

48.3 Results and Discussion

Figure 48.6 shows the distribution of electric field around concentrator tip. It can be observed from the zoomed view of electric field distribution around corners of concentrator tip that electric field strength is more than 10^6 V/m. Due to the high strength of electric field around concentrator tip, medium around the concentrator tip is ionized and plasma formation takes place. Consequently, maximum temperature in the surface of the workpiece just beneath the concentrator tip and concentrator rises rapidly up to 1100 °C and 950 °C in 6 s, respectively as shown in Fig. 48.7. It is also evident that maximum temperature in workpiece is higher as compared to the concentrator. This shows that core of the plasma formed, i.e., the zone of maximum temperature is around the surface of the workpiece. This is also evident from Fig. 48.8. It takes around 2 s to reach the softening temperature of borosilicate glass (~820 °C). As the glass melts, the concentrator starts to penetrate inside the workpiece. But, after few seconds, when the temperature is around 1100 °C, phenomena of ablation also takes place. Figure 48.8 shows the distribution of temperature along the different direction of the surface. It can be observed from Fig. 48.8 that temperature drops rapidly along the surface away from concentrator. The temperature just beneath the concentrator is around 950 °C and increases to around 1100 °C up to a distance of 0.4 mm. But after 0.4 mm, temperature decreases rapidly, and after 0.8 mm, the temperature comes below 800 °C. Due to this, the material removal takes place only up to a radial distance of 0.4 mm.

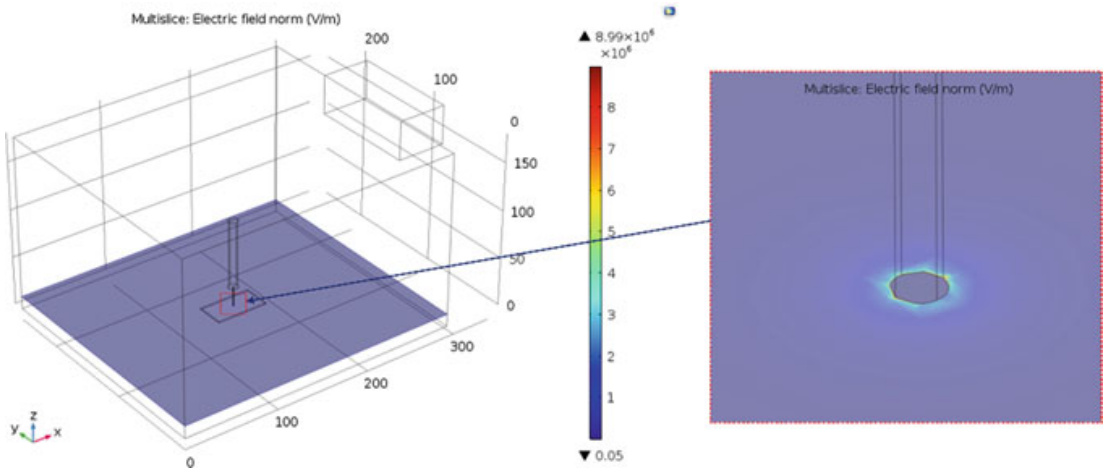


Fig. 48.6 Distribution of electric field around concentrator (inset: zoomed view of electric field distribution around concentrator tip)

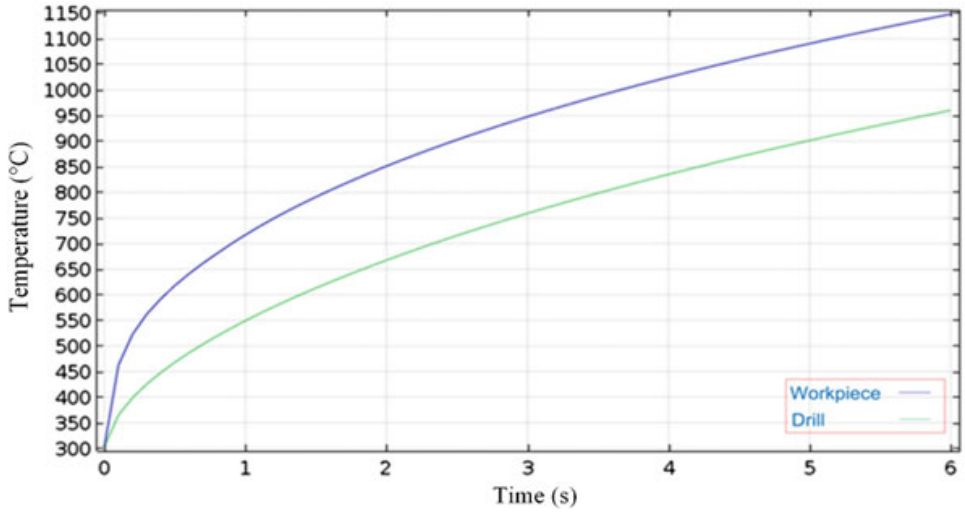


Fig. 48.7 Distribution of maximum temperature in concentrator and workpiece with respect to time

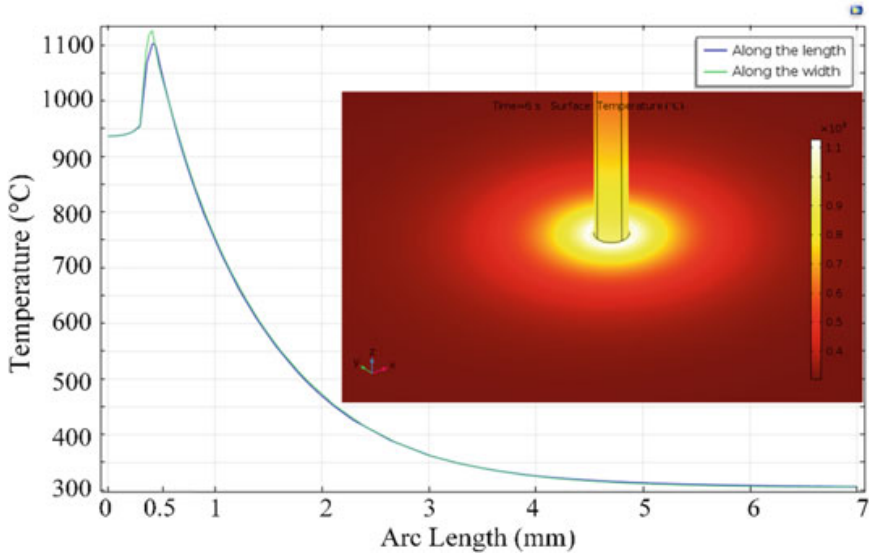


Fig. 48.8 Distribution of temperature along the surface with respect to arc length in different direction

The area after radial distance of 0.4 mm is heat-affected zone and highly stressed. Due to the high thermal gradient after 0.4 mm, the value of Von Mises stress rises above the critical stress limit for fracture of glass (60 MPa) as shown in Fig. 48.9. The region up to a radial distance of $\sim$1.3 mm is highly stressed zone. Crack initiates in this zone and propagates radially up to certain distance on the surface of the workpiece and after that circumferentially around the hole as shown in Fig. 48.5a. Due to the high thermal stress, glass often cracks while drilling in air. The hole drilled in glass in air at a controlled feed rate of 1.2 mm/s is shown in

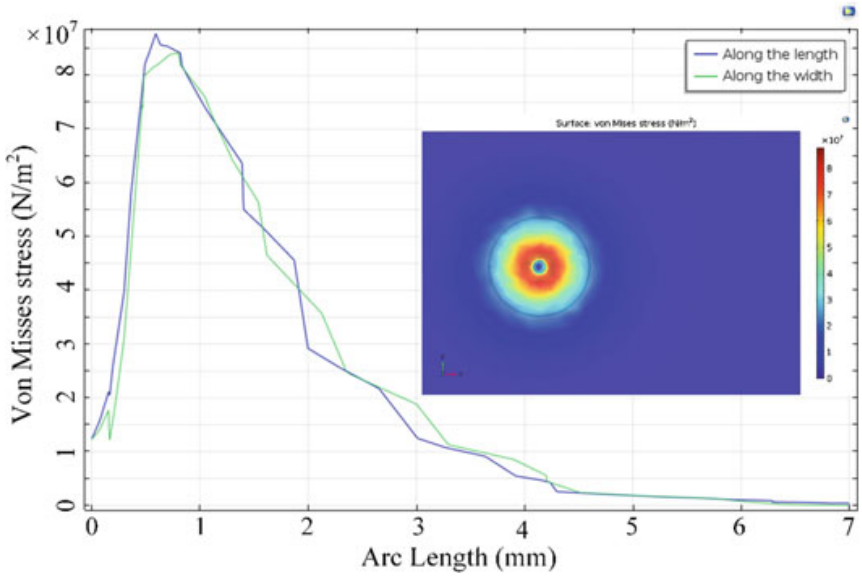


Fig. 48.9 Stress distribution along the surface just below the tip of the concentrator

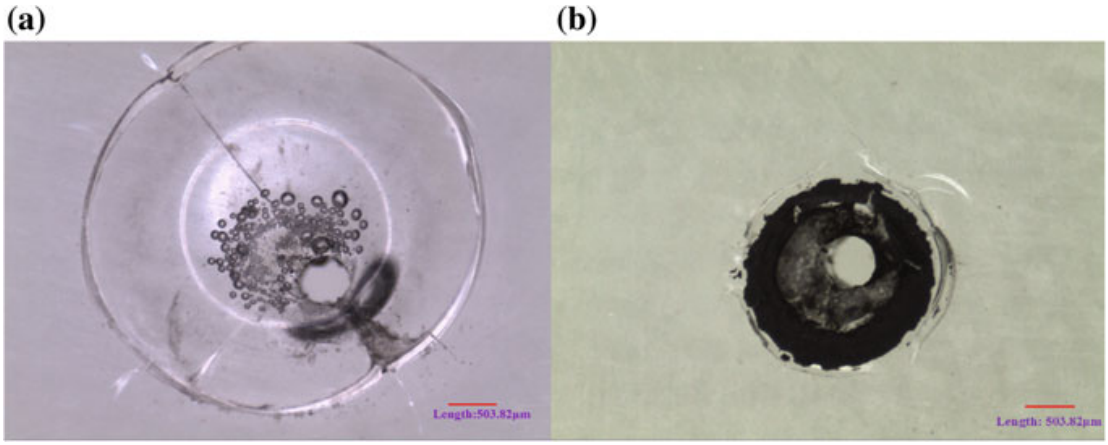


Fig. 48.10 Optical image of the hole drilled in (**a**) air (**b**) dielectric at a feed rate of 1.2 mm/s

Fig. 48.10a. It can be observed that significant amount of cracks is present around hole drilled. But as the medium around the concentrator tip is changed by a dielectric like transformer oil, the cracks and heat-affected zone are reduced significantly as shown in Fig. 48.10b. Dielectrics help in reducing the thermal gradient due to which cracks around the hole are reduced. Dielectric also helps in confining the shape of the plasma which ultimately reduces the thermal damage area as shown in Fig. 48.10b. But, a significant amount of carbon deposit is seen around the hole.

48.4 Conclusion

COMSOL Multiphysics 5.2 was used to simulate the microwave drilling process. Followings are the conclusions based on the simulation study:

- High electric field strength around the concentrator tip ($\sim 8.99 \times 10^6$ V/m) which is more than dielectric breakdown strength of air ionizes the medium around it to produce plasma.
- The maximum temperature on the surface of workpiece just beneath the concentrator tip is ~ 1100 °C which is higher than the ablation temperature of the borosilicate glass.
- The mechanism of material removal from the target area consists of combined action of ablation and melting.
- A very high thermal stress (~ 85 MPa) around the concentrator tip is induced in the surface of the workpiece which is the prime reason behind the crack generation while drilling.
- Dielectric helps in reducing the thermal gradient which ultimately lessens the crack and thermal damage around the hole.

References

1. Jerby, E., Dikhtyar, V., Aktushev, O., Grosglick, U.: The microwave drill. Science **298**, 587–589 (2002)
2. Jerby, E., Dikhtyar V.: Drilling into hard non-conductive materials by localized microwave radiation. In Advances in Microwave and Radio Frequency Processing, pp. 687–694. Springer, Berlin, Heidelberg (2006)
3. Jerby, E., Thompson, A.M.: Microwave drilling of ceramic thermal-barrier coatings. J. Am. Ceram. Soc. **87**(2), 308–310 (2004)
4. Jerby, E., Aktushev, O., Dikhtyar, V.: Theoretical analysis of the microwave-drill near-field localized heating effect. J. Appl. Phys. **97**(3), 034909 (2005)
5. Lautre, N.K., Sharma, A.K., Kumar, P., Das, S.: A photoelasticity approach for characterization of defects in microwave drilling of soda lime glass. J. Mater. Process. Technol. **225**, 151–161 (2015)
6. Lautre, N.K., Sharma, A.K., Kumar, P.: Distortions in hole and tool during microwave drilling of perspex in a customized applicator. In Microwave Symposium (IMS), 2014 IEEE MTT-S International, pp. 1–3 (2014)
7. Lautre, N.K., Sharma, A.K., Pradeep, K., Das, S.: A simulation approach to material removal in microwave drilling of soda lime glass at 2.45 GHz. Appl. Phys. A **120**(4), 1261–1274 (2015)
8. Lautre, N.K., Sharma, A.K., Das, S., Kumar, P.: On crack control strategy in near-field microwave drilling of soda lime glass using precursors. J. Thermal Sci. Eng. Appl. **7**(4), 041001 (2015)
9. Kumar, G., Sharma, A.K.: Role of dielectric fluid and concentrator material in microwave drilling of borosilicate glass. J. Manuf. Process. **33**, 184–193 (2018)

10. Das, S., Klotz, M., Klocke, F.: EDM simulation: finite element-based calculation of deformation, microstructure and residual stresses. J. Mater. Process. Technol. **142**(2), 434–451 (2003)
11. Dhanik, S., Joshi, S.S.: Modeling of a single resistance capacitance pulse discharge in micro-electro discharge machining. J. Manuf. Sci. Eng. **127**(4), 759–767 (2005)
12. Jiao, J., Wang, X.: A numerical simulation of machining glass by dual CO_2-laser beams. Opt. Laser Technol. **40**(2), 297–301 (2008)
13. Krötz, H., Roth, R., Wegener, K.: Experimental investigation and simulation of heat flux into metallic surfaces due to single discharges in micro-electrochemical arc machining (micro-ECAM). Int. J. Adv. Manuf. Technol. **68**(5–8), 1267–1275 (2013)

Chapter 49
A Study on the Effect of Polarity Change on Various Parameters on Ti6Al4V in Powder-Mixed Micro-EDM Using Multi-objective Grey Fuzzy Optimization

Deepak Sharma, Anusha Roohi Siddique, Vishnu Kumar, Shalini Mohanty and Alok Kumar Das

Abstract The study emphasizes on the effect of polarity change while conducting experiments using tungsten disulphide in powder-suspended micro-EDM process on Ti6Al4V alloy workpiece with brass tool electrode. For design of experiment, Taguchi method with L9 orthogonal array has been adopted. For each set of experiments, three levels of factors such as voltage, duty factor and machining time are considered. The responses such as MRR (higher the better) and SR (lower the better) are calculated, and a grey fuzzy logic model is built for optimization of the parameters. ANOVA table shows that the voltage plays the most significant role when machining in straight polarity followed by machining time and duty factor, whereas for reverse polarity machining time is the most significant one followed by voltage and duty factor. In case of straight polarity, confirmatory tests were performed at $V = 25$, DF = 55 and $T = 8$ and in reverse polarity at $V = 25$, DF = 45 and $T = 8$ and GFRG in case of straight polarity was found to increase by 0.216 and increased by 0.182 for the reverse polarity.

Keywords Micro-EDM · Surface roughness · Grey fuzzy · ANOVA · MRR

D. Sharma · A. R. Siddique · V. Kumar · S. Mohanty (✉) · A. K. Das
Department of Mechanical Engineering, Indian Institute of Technology (Indian School of Mines), Dhanbad 826004, India
e-mail: shalinimohanty5@yahoo.com

D. Sharma
e-mail: mr.deepaksharma7s7@gmail.com

A. R. Siddique
e-mail: anusha.roohi@gmail.com

V. Kumar
e-mail: vishnu.bit15@gmail.com

A. K. Das
e-mail: eralok@yahoo.co.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_49

49.1 Introduction

With the everyday advancement in the field of engineering and technology, one gets to witness novel methods of machining at every step. Now is the time when the researchers and industries are moving towards the non-traditional machining and hybrid machining at a very fast rate. With growing requirements of precision, tolerances and other high aspect ratio micro-features, micro-EDM has served as a very important technique to achieve the former. Micro-EDM is nothing but just an evolution of the process of EDM where the size of the tool and movement of the axis of resolution are small to the level of microns and the discharge energy which is released is very less as compared to the conventional EDM [1].

Various parameters are considered for carrying out experiments in micro-EDM, but polarity change has been a challenging parameter by researchers. The difference in polarities of the workpiece and electrode and comparing them help in comprehending the results better. Generally, the aim of a machining process is to provide better surface finish with optimized parameters. Polarity of the process holds great significance in such a case. To this aspect, Schluze [2] studied that EDM with reverse polarities (tool positive) proves to be important when small material is required to be removed during finishing operations. It can also be used to remove oxide layers on the surface when used in hybrid processes with ECM. Parshuramalu et al. [3] have provided with a methodology for testing and finding out how the different tools and workpiece can be used with alteration of polarity. It is clear that if the polarity is not chosen properly, serious effects on wear and stability can occur. The experiments were carried out on a steel workpiece in a sinking EDM with copper electrode. From the results, it was found that the removal of steel was greatly affected by the polarity and the current settings. Researchers have been trying to improve the EDM process for a long time now. From changing the electrode to improving the dielectric, everything is tried to increase the efficiency of this process. Kumar et al. [4] mixed nanopowder of Al_2O_3 in dielectric. In this process, reverse polarity was used and it was found out that there was a reduction in micro-cracks with enhanced surface finish. Tyagi et al. [5] produced green sintered electrodes along with copper powder in die-sinking EDM. They performed EDC in straight polarity. As a consequence, coating with enhanced thickness and better surface properties were formed. It is evident that research progress with reference to the polarity change has not been much prevalent and hence, in the present work, an effort is done to study the effect of different parameters of the machining process on surface roughness of Ti6Al4V workpiece when machined with mixing powder in dielectric in micro-EDM.

Whenever machining is done, it is very necessary that the parameters at which it is done are optimum. It is an arduous work to ensure that the process parameters are best. Hence, certain optimization methods like principal component analysis, Taguchi methodology and GRA are used and have proved to be of great importance for obtaining desirable results. Various researchers have used GRA for optimization of the performance parameters while experimenting. Gopal and Chakradhar [6]

optimized the parametric settings of feed rate, voltage applied and the concentration of the electrolyte for an ECM process. Durairaj et al. [7] made use of the GRA method in WEDM for finding out the vital process parameters which affected the optimal values of kerf width, and SR. Kumar and Kumar [8] determined optimal values of different parameters like pulse on duration (Ton), discharge current and gap voltage for an EDM process using GRA method considering MRR, electrode wear, SR as response values. Natarajan and Arunachalam [9] used Taguchi system and GRA to improve the EDM machining inputs by many regular attributes. Gopalakannan et al. [10] conducted experiments by selecting process parameters like gap voltage, pulse current, pulse on and off duration and used Taguchi-based GRA for EDM process with several outputs like EWR, MRR and SR. They decided that current and pulse on duration were the parameters of highest significance. Dewangan et al. [11] conducted experiments by means of RSM on composites with wire EDM and applied grey fuzzy method. Optimum values of process parameters were acquired, and consequent values of the MRR and SR were established. Ramanan and Shanavas [12] applied Taguchi and grey fuzzy logic to improve parameters of performance in turning of the AISI D2, while Soepangkat et al. [13] used the same for optimization in submerged arc welding.

Though a lot of researchers have used the above-mentioned methodology for optimizing difficult parameters of various non-traditional machining processes and electrical discharge machining in general [14], a optimization of parameters specifically for a micro-EDM has not been done. The purpose of the present study is to highlight the multi-objective optimization of powder-mixed micro-EDM under different polarities using grey fuzzy technique.

49.2 Methodology

49.2.1 Grey Relational Generation

In grey method, some interrelated information is well known and other is unknown. These multiple interrelated performance characteristics can be solved by grey relational analysis. In this analysis by using certain formulas, experimental data (MRR and SR) was normalized in the range of 0 and 1 and these values were used in generating grey relational coefficient which will direct the connection among the desired and actual experimental numbers. According to the type of normalized data required, GRA is of three types, namely lower the better, higher the better and nominal the better. To get normalized higher-the-better MRR values, the following equation was used:

$$X_a(b) = \frac{y_a(b) - \min y_a(b)}{\max y_a(b) - \min y_a(b)} \tag{49.1}$$

Similarly, normalized smaller-the-better SR values can be obtained by:

$$Y_a(b) = \frac{\max y_a(b) - y_a(b)}{\max y_a(b) - \min y_a(b)} \tag{49.2}$$

where $a = 1, 2, 3, \ldots k$ and $b = 1, 2, 3, \ldots q$; k is the number of experimental data; q is the number of factors; $y_a(b)$ is the actual sequence; $y_a(b)$ is the value post-grey relational generation; min $y_a(b)$ and max $y_a(b)$ are the smallest and largest values of $y_a(b)$, respectively.

Grey relational coefficient (GRC)

$$\epsilon_a(b) = \frac{\Delta\min + \omega\Delta\max}{\Delta oi(k) + \omega\Delta\max} \tag{49.3}$$

where $\epsilon_a(b)$ is the grey relational coefficient; Δ_{oa} is difference among $y_o(b)$ and $y_a(b)$; $y_o(b)$ is the best sequence; Δmax is uppermost value of $\Delta_{oa}(b)$; Δmin is smallest value of $\Delta_{oa}(b)$; ω is the distinctive coefficient which is in the range of $0 \leq \omega \leq 1$. To get the results closer to optimal normalized value, highest value of grey relational coefficient has been used.

Grey Relation Grade (GRG)

The average of GRC value corresponding to each criterion is the GRG value.

$$\Gamma_a = \frac{1}{n}\sum_{b=1}^{n} \epsilon_a(b) \tag{49.4}$$

where n is no. of characteristics; greater the value of GRC, more will be its priority. The greater value of GRC suggests that corresponding process parameter is close to the optimal parameter.

49.2.2 Fuzzy Logic

Fuzzy logic is a scheme where the perceptive analysis of certain data by a human brain is represented. Fuzzy logic helps in making certain deductions which are similar to the results from a certain scheme of proposals which show considerable similarity with the initial data. A fuzzy logic division comprises of membership functions, fuzzifier, a fuzzy rule base, the inference engine (Mamdani) and a defuzzifier. The fuzzy interface engine (Mamdani) is shown in Fig. 49.1. Membership functions are fed to the fuzzifier which generates the grey relational coefficients. The membership function of MRR and SR is shown in Fig. 49.2, and membership function of GFRG is shown in Fig. 49.3. The fuzzifier interface engine converts these values to fuzzy values, and at end, defuzzifier alters fuzzy values in the form of grey fuzzy reasoning grade.

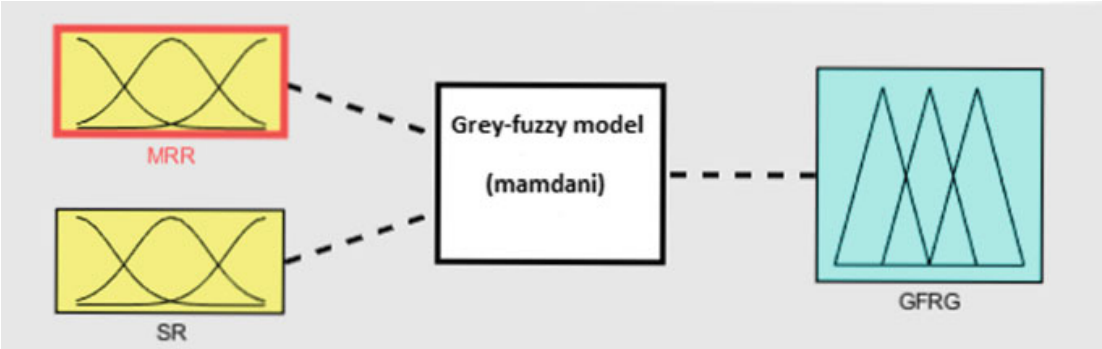


Fig. 49.1 Fuzzy logic model for process responses

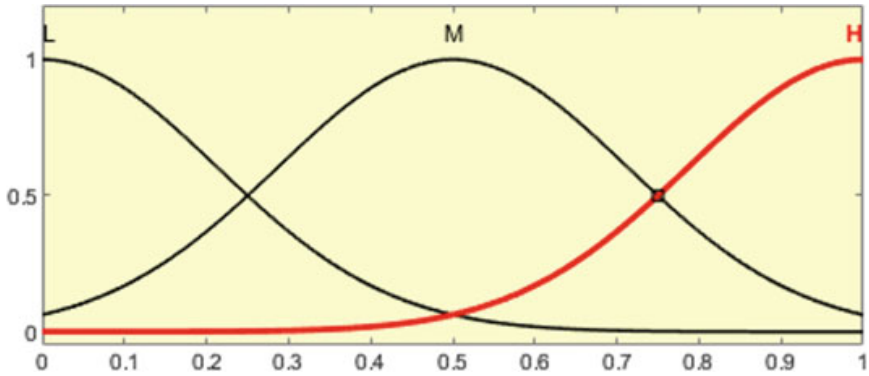


Fig. 49.2 Membership function for MRR and SR

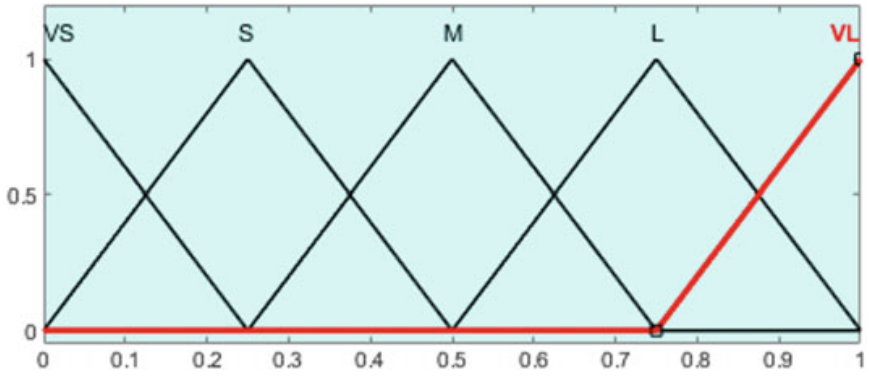


Fig. 49.3 Membership function for grey fuzzy reasoning grade

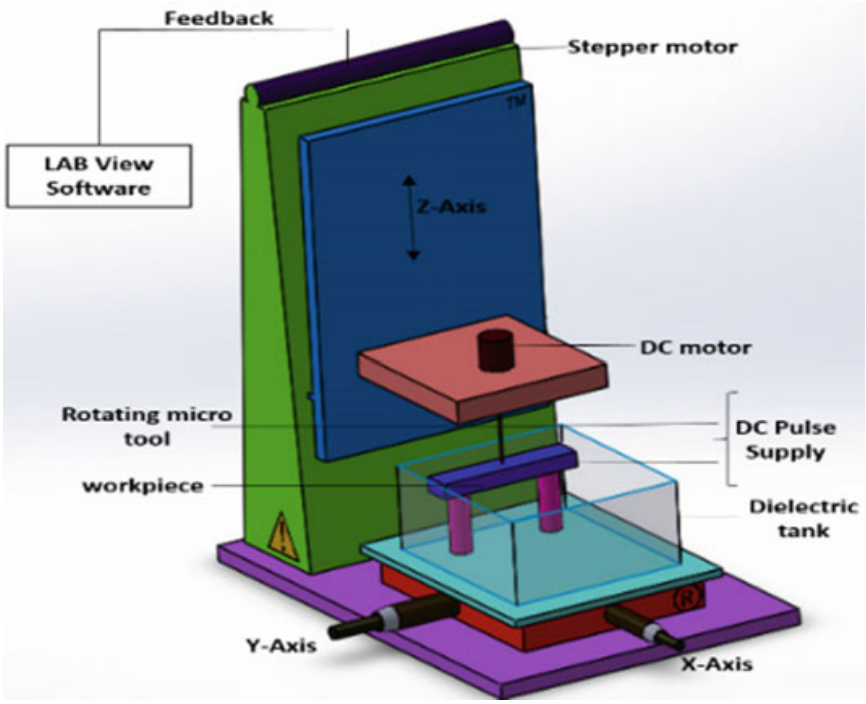


Fig. 49.4 Three-dimensional figure of micro-EDM system

49.3 The Experimental Set-Up and Procedure

Experiment was conducted on Ti6Al4V workpiece of dimensions 10 × 10 × 1.5-mm thickness with brass tool of 1-mm diameter and 30-mm length. De-ionized water was used as dielectric medium and tungsten disulphide as suspended powder particles. To conduct the experiments, a three-axis system fitted out with the micro-EDM set-up is used as shown in Fig. 49.5. XY-stage is used to provide longitudinal and transverse motion to the workpiece. A DC motor is used to impart rotational motion to the tool, and the workpiece is submerged in the dielectric with proper clamping. Another motor is devoted to the set-up to provide vertical motion to the tool. The 3D model of micro-EDM set-up is presented in Fig. 49.4. Transistor-based power supply was used, and LabVIEW software allows the z-axis motion.

A total of 9 × 2 experiments were directed (for reverse and straight polarities) by choosing L9 Taguchi DOE as shown in Table 49.2. The uncontrolled input parameters are voltage (V), duty factor (%) and machining time (min). Tool is fixed to be positive and the workpiece is connected to negative terminal for reverse polarity, and tool is negative and workpiece is positive for straight polarity. Trial experiments were conducted to decide different levels of parameter values, and it was observed that as the difference between the values of parameters for consecutive levels was small it was difficult to differentiate between the performance parameter values so middle-range parameters with sufficient difference were nominated for the experiments as shown in Table 49.1. The performance of

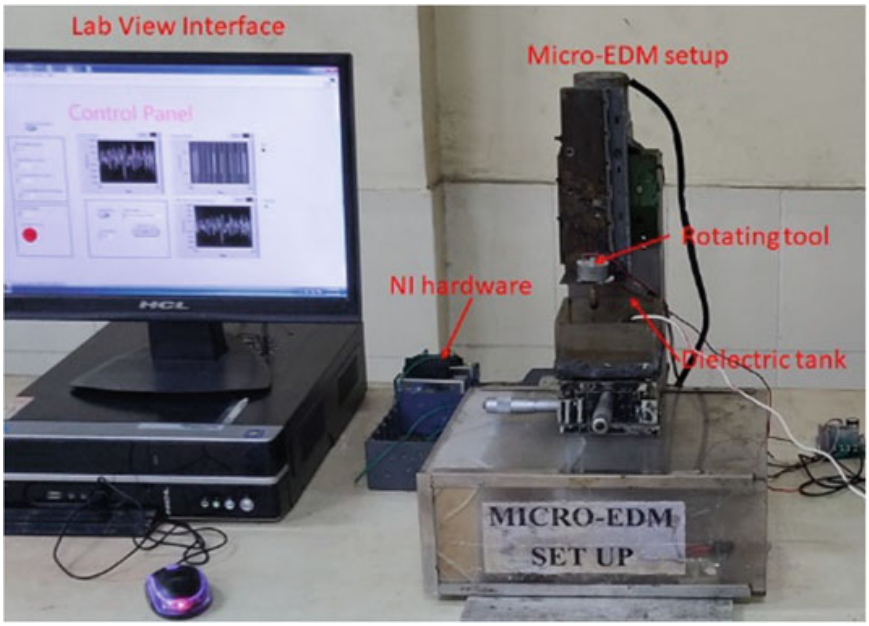


Fig. 49.5 Micro-EDM set-up

Table 49.1 Parameters of the process and their levels

Factors	Parameters	Level 1	Level 2	Level 3
A	Voltage, *V*	25	40	55
B	Duty factor, DF	35	45	55
C	Machining time, *T* (min)	6	8	10

experiments was measured by MRR and SR. Each of these values is considered at once to get GFRG which yield the optimum MRR and SR. Design of experiment is generated by L9 Taguchi array system.

The flow chart shown in Fig. 49.6 gives a concise idea about how the grey fuzzy logic works. It starts with the construction of design of experiment with the help of Taguchi L9 design. Experimental data was then converted to normalized data with the help of formulas. This normalized data was used to calculate grey relational coefficients for both MRR and SR, and the GRCs for both straight and reverse polarities are then fed to the fuzzifier as input membership functions. With the help of certain "if and then" rules, fuzzy interface engine (Mamdani) converts these values to fuzzy values and then the defuzzifier converts them into grey fuzzy reasoning grades. GFRG values decide the optimal solution, and then confirmatory tests were done to verify the optimal solution.

Table 49.2 Design of experiment with responses

Exp. no.	V	DF	T	MRR_{st} (g/s)	MRR_{rev} (g/s)	SR_{st} (μm)	SR_{rev} (μm)
1	25	35	6	0.00101	0.00023	0.998	0.266
2	25	45	8	0.00172	0.00015	1.556	0.177
3	25	55	10	0.00347	0.00032	1.856	0.379
4	40	35	8	0.00077	0.00031	2.085	0.481
5	40	45	10	0.00523	0.00038	2.83	0.75
6	40	55	6	0.00054	0.00006	1.85	0.489
7	55	35	10	0.00282	0.00020	3.248	0.588
8	55	45	6	0.00395	0.00023	2.041	0.575
9	55	55	8	0.00887	0.00038	2.755	0.696

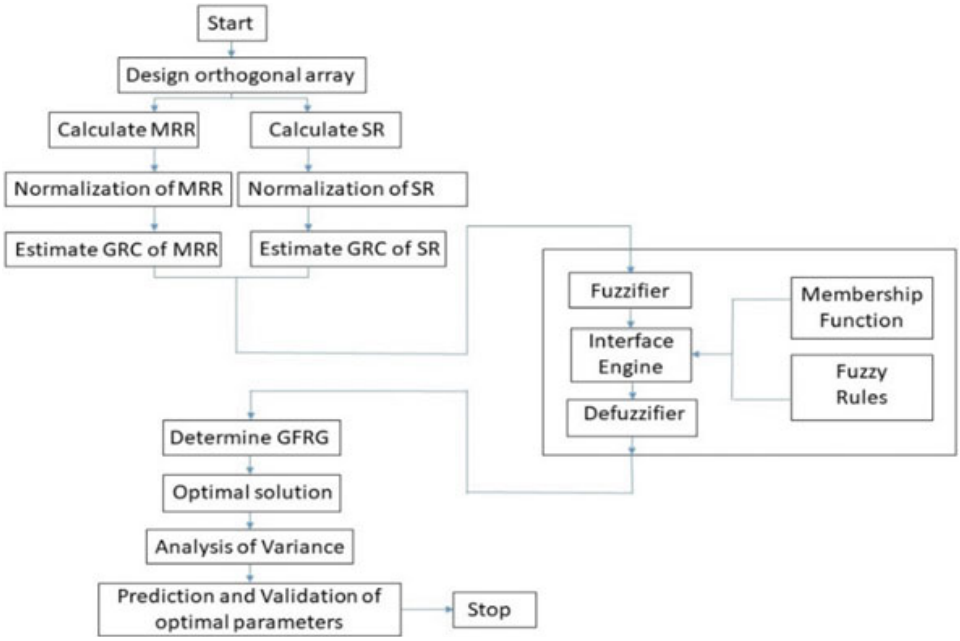


Fig. 49.6 Flow chart of grey fuzzy-based algorithm

49.4 Results and Discussion

49.4.1 The Grey Reasoning Coefficient

For each performance parameter, normalized values were calculated. Table 49.3 displays the calculated normalized value of each quality characteristic, grey relational coefficients and grey relational grade.

Table 49.3 Normalized and GRC table

Exp no.	Normalized value of		Normalized value of		Grey relational coefficient of		Grey relational coefficient of		GRG_{st}	GRG_{rev}
	MRR_{st}	MRR_{rev}	SR_{st}	SR_{rev}	MRR_{st}	MRR_{rev}	SR_{st}	SR_{rev}		
1	0.05664	0.51586	1	0.84467	0.34641	0.50805	1	0.76298	0.673	0.636
2	0.14220	0.25953	0.752	1	0.36824	0.40307	0.66844	1	0.518	0.702
3	0.35161	0.80422	0.61866	0.64746	0.43539	0.71862	0.56732	0.58648	0.501	0.653
4	0.02760	0.77218	0.51688	0.46945	0.33958	0.68699	0.50859	0.48518	0.424	0.586
5	0.56342	0.99647	0.18577	0	0.53385	0.993	0.38045	0.33333	0.457	0.663
6	0	0	0.62133	0.45549	0.33333	0.33333	0.56904	0.47869	0.451	0.406
7	0.27457	0.43255	0	0.28272	0.40802	0.46840	0.33333	0.41075	0.371	0.440
8	0.40921	0.52162	0.53644	0.30541	0.45838	0.51105	0.51891	0.41855	0.489	0.465
9	1	1	0.21911	0.09424	1	1	0.39035	0.35568	0.695	0.678

49.4.2 The Grey Fuzzy Reasoning Analysis

MRR and surface roughness are two inputs, and GFRG is one output of fuzzy logic method. The GRC for MRR and surface roughness are considered as the inputs to the fuzzy logic method. To establish the fuzzy if and then rule as shown in Table 49.4, the language-based membership operators L, M and H were used to indicate the grey relational coefficient of the entered variables. Similarly, the output grey relational grade has been signified through the member functions VS, S, M, L, VL on triangular shape membership function.

The common methods for defuzzification are centre of maximum (COM) and centre of area (COA). The relationship between two inputs and output is developed by certain "if-then" rules that are:

Rule 1: If $m1$ is $V1$ and $n2$ is $W1$ and $p3$ is $X1$ and $q4$ is $Y1$, then output (G) is $Z1$.
Rule 2: If $m1$ is $V2$ and $n2$ is $W2$ and $p3$ is $X2$ and $q4$ is $Y2$, then output (G) is $Z2$.
Rule n: If $m1$ is Vn and $n2$ is Wn and $p3$ is Xn and $q4$ is Yn, then output (G) is Zn.

where Vi, Wi, Xi and Yi are the fuzzy subsections distinct by the equivalent member functions, i.e. μVi, μWi, μXi and μYi, respectively. The inference engine (Mamdani) does fuzzy reasoning on rules defined by fuzzy system while considering max and min inference for generating a specific fuzzy value, $\mu c_o(G)$.

$$\begin{aligned}\mu c_o(G) = &(\mu V_1(m_1) \wedge \mu W_1(n_2) \wedge \mu X_1(p_3) \wedge \mu Y_1(q_4) \wedge \mu Z_1(G)) \\ &\vee (\mu V_1(m_1) \wedge \mu W_2(n_1) \wedge \mu X_2(p_2) \wedge \mu Y_2(q_3) \wedge \mu Z_2(G)) \\ &\vee (\mu V_n(m_1) \wedge \mu W_n(n_2) \wedge \mu X_n(p_3) \wedge \mu Y_n(q_3) \wedge \mu Z_n(G)).\end{aligned}$$

where $\vee$ is the maximum operator and $\wedge$ is the minimum operator. The process of defuzzification has been used to convert the fuzzy interface output into exact value of GFRG(G_o).

Table 49.4 Fuzzy if and then rule

I/C criteria	Input (if)		MPCI (then)
	Individual desirability values of MRR	Individual desirability values of SR	
1	L	L	VS
2	L	M	S
3	L	H	M
4	M	L	S
5	M	M	M
6	M	H	L
7	H	L	M
8	H	M	L
9	H	H	VL

$$G_o = \frac{\sum G\mu c_o(G)}{\sum \mu c_o(G)} \tag{49.5}$$

Graphical illustration of fuzzy logic reasoning is shown in Fig. 49.7 (straight polarity) and Fig. 49.8 (reverse polarity). The nine rows represent rules, and three columns represent MRR, SR and GFRG, respectively. Both the graphs are taken at optimal parameters, and in Fig. 49.7 the input values of grey relational coefficient of SR and MRR are 0.39 and 1, respectively, and defuzzified output is 0.666 which shows the combined effect of both the input values. Similarly, in Fig. 49.8 the input values are 0.403 and 1 and corresponding output value is 0.671. The GFRG values with their respective ranks are shown in Table 49.5.

The response analysis in Tables 49.6 and 49.7 determined that for straight polarity the parameter combination $V = 25$, DF = 55 and $T = 8$ gives the optimal value and for reverse polarity the parameter combination $V = 25$, DF = 45 and $T = 8$ gives the optimal value of MRR and SR. Figures 49.9 and 49.10 show the mean of mean, response graph of GFRG for straight and reverse polarities, respectively.

49.4.3 Analysis of Variance

ANOVA is formulated for both straight and reverse polarities using GFRG values to calculate the percentage by which each parameter is affecting the process. The results of ANOVA for straight polarity are given in Table 49.8 and for reverse polarity in Table 49.9. From ANOVA tables, it is justified that in case of straight polarity voltage (27.19%) influences more on alloying process followed by

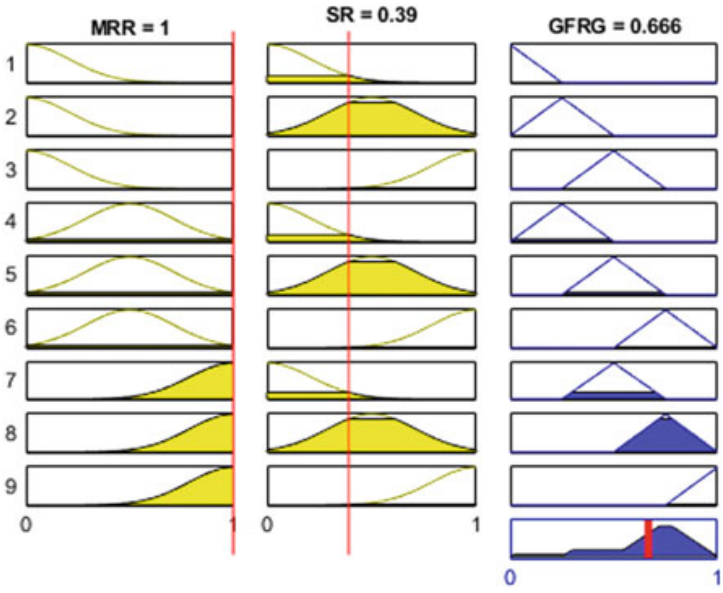


Fig. 49.7 Fuzzy logic rules for straight polarity

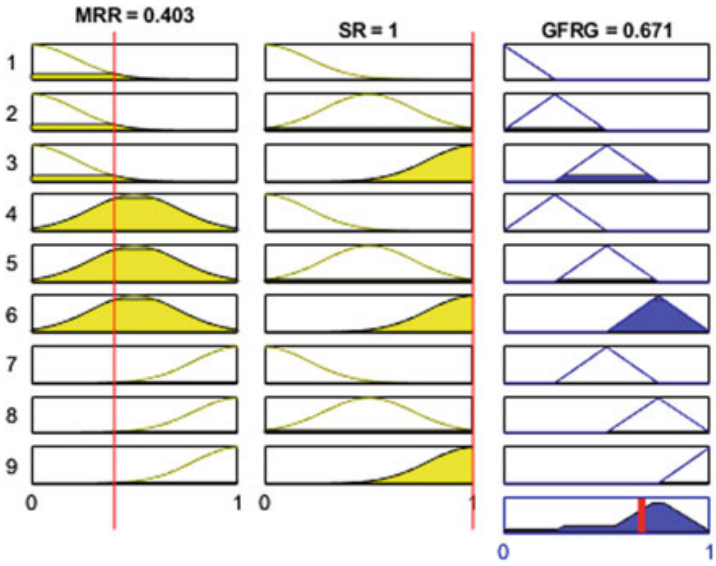


Fig. 49.8 Fuzzy logic rules for reverse polarity

Table 49.5 GFRG with the respective rank

Exp. no.	$GFRG_{st}$	$Rank_{st}$	$GFRG_{rev}$	$Rank_{rev}$
1	0.649	2	0.611	4
2	0.516	3	0.671	1
3	0.501	4	0.604	5
4	0.445	8	0.566	6
5	0.47	6	0.641	3
6	0.46	7	0.435	9
7	0.422	9	0.466	8
8	0.493	5	0.478	7
9	0.666	1	0.653	2

Table 49.6 Response values for the GFRG straight polarity

Factors	Level 1	Level 2	Level 3
V	0.5553	0.4583	0.5270
DF	0.5053	0.4930	0.5423
T	0.5340	0.5423	0.46432

Table 49.7 Response values for the GFRG reverse polarity

Factors	Level 1	Level 2	Level 3
V	0.6287	0.5473	0.5323
DF	0.5477	0.5967	0.5640
T	0.5080	0.6300	0.5703

machining time (20.04%) and duty factor (7.201). Similarly, from table in case of reverse polarity machining time (33.63%) affects more followed by voltage (24.27%) and duty factor (5.62%).

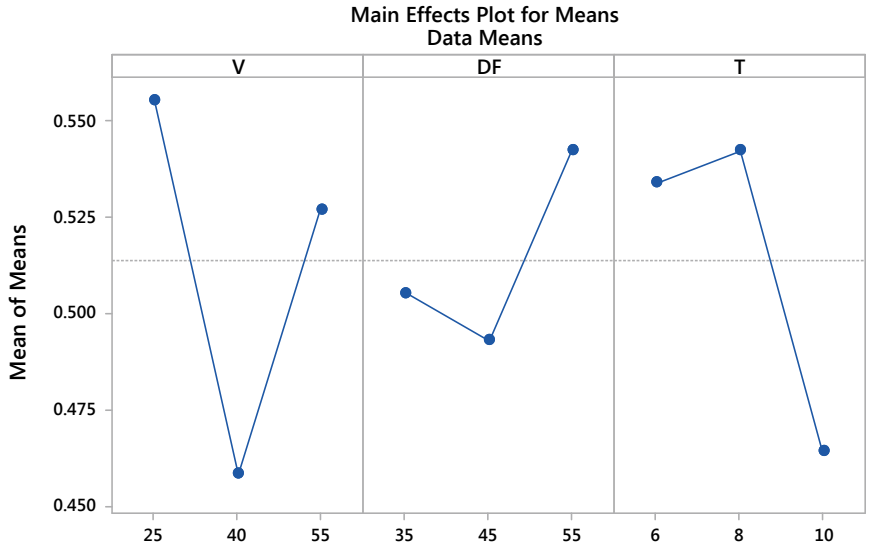


Fig. 49.9 Response graph for straight polarity

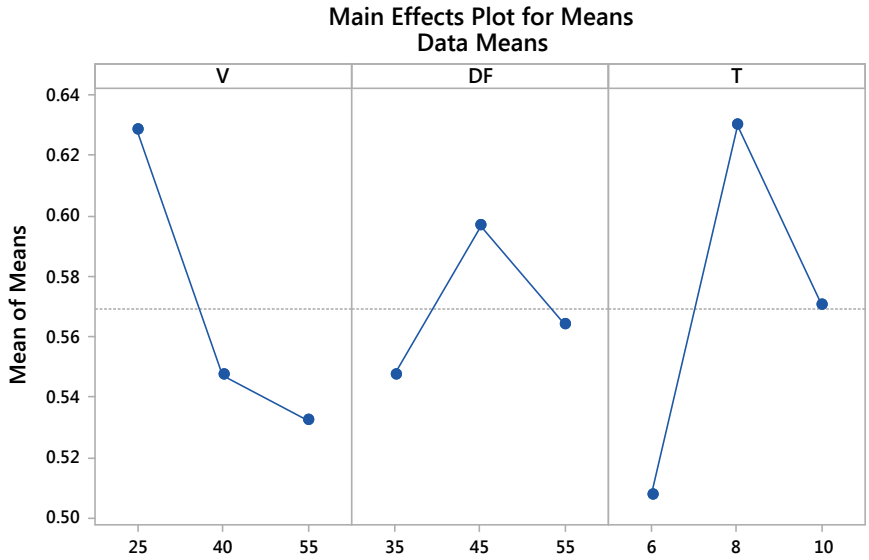


Fig. 49.10 Response graph for reverse polarity

Table 49.8 Analysis of variance of grey fuzzy relation grade for straight polarity

Factor	DOF	Sum of squares	Mean squares	F value	P value	% contribution
V	2	0.01493	0.007463	0.99	0.424	27.19
DF	2	0.003955	0.001977	0.21	0.815	7.201
T	2	0.01101	0.005503	0.67	0.545	20.046
Error	2	0.025027	0.025028	–	–	45.519
Total	8	0.054922	–	–	–	100

Table 49.9 Analysis of variance of grey fuzzy relation grade for reverse polarity

Factor	DOF	Sum of squares	Mean squares	*F* value	*P* value	% contribution
V	2	0.016120	0.008060	1.04	0.409	24.2778
DF	2	0.003735	0.001867	0.19	0.831	5.62516
T	2	0.0223	0.011165	1.67	0.265	33.6305
Error	2	0.024213	0.024213	–	–	36.4664
Total	8	0.066398	–	–	–	100

49.4.4 Theoretical Prediction and Confirmation of Experiments

After obtaining optimal level of design parameters in case of straight and reverse polarities, the last step is to forecast and settle the upgradation of performance features. Table 49.10 shows the contrast of original process parameters and optimal process parameters in case of straight polarity, and as noted the MRR improved from 0.00347 to 0.009225 g/min, surface roughness (SR) decreased from 1.856 to 1.83 μm, and GFRG is improved by 0.216. Similarly, Table 49.11 shows the comparison for reverse polarity, and as noted MRR increased from the value of 0.00034 to 0.00035 g/min and GFRG is improved by 0.182. It is clearly revealed that multiple performance parameters are improved using grey fuzzy optimization technique. MRR is calculated by the formula as given in Eq. 49.6.

$$\text{MRR} = \frac{W_\text{i} - W_\text{f}}{t} \tag{49.6}$$

where W_i and W_f are the weights before and after the experiments and t is the time of machining in seconds.

FESEM images of both straight and reverse polarities at 40 V, 35% duty factor and 8 min machining time are shown in Figs. 49.11 and 49.12, respectively. It is found that the recast layer is more uneven in case of straight polarity whereas a

Table 49.10 Results of process performance by means of the original and optimal factors for straight polarity

	The initial process parameters	The optimal process parameters	
		Prediction	Experiment
Level	A1B3C3	A1B3C2	A1B3C2
MRR	0.00347	0.009005	0.009225
Surface roughness	1.856	1.85	1.83
GFRG	0.501	0.704	0.717
Improvement of GFRG		0.203	0.216

Table 49.11 Results of process performance by means of the original and optimal factors for reverse polarity

	The initial process parameters	The optimal process parameters	
		Prediction	Experiment
Level	A1B3C3	A1B2C2	A1B2C2
MRR	0.00032	0.00034	0.00035
Surface roughness	0.379	0.15	0.177
GFRG	0.604	0.767	0.786
Improvement of GFRG		0.163	0.182

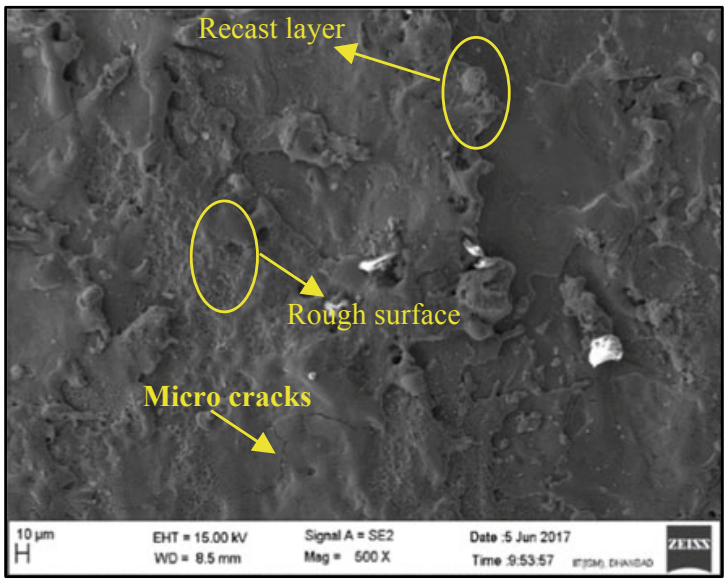


Fig. 49.11 FESEM for straight polarity

comparatively lower surface roughness is seen in reverse polarity. This may be due to less discharge energy available at workpiece in case of reverse polarity as compared to straight polarity. In addition to that, WS_2 is a solid lubricating powder; thus, it plays an important role to improve surface properties of workpiece.

49.5 Conclusions

The combination of optimization method and grey fuzzy logic is used in current research work for concurrently optimizing the numerous performance characteristics like MRR and SR both in straight and reverse polarities. After summarizing the key attributes of this experimental investigation, the following conclusions were drawn:

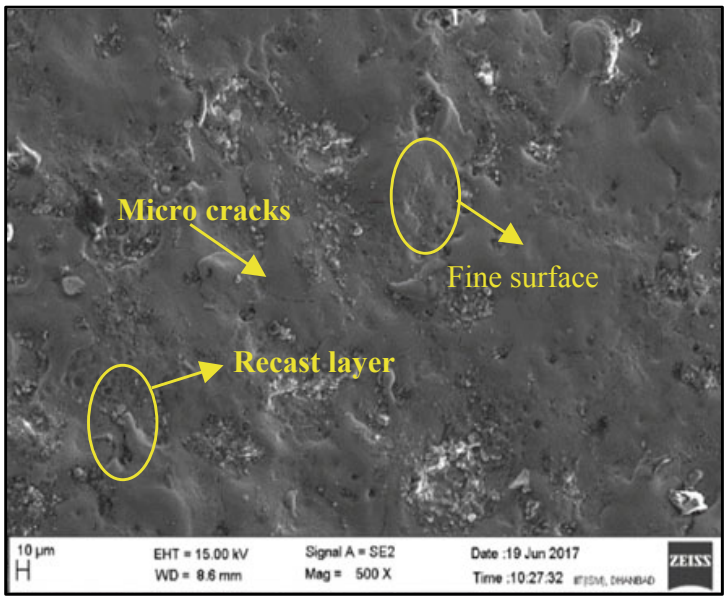


Fig. 49.12 FESEM for reverse polarity

1. The optimal parametric combination in straight polarity is $V = 25$, DF = 55 and $T = 8$ and in reverse polarity is $V = 25$, DF = 45 and $T = 8$ determined by grey fuzzy logic.
2. The ANOVA results determine that in case of straight polarity voltage (27.19%) influences more and in case of reverse polarity machining time (33.63%) affects more.
3. It has been concluded from the results that the lower surface roughness was obtained in case of reverse polarity as compared to what obtained in straight polarity as the craters formed in case of straight polarity were of comparatively greater size due to the direction of current which in turn resulted in high energy flow to the workpiece.
4. It is indeed complicated to convert multiple parametric optimizations into a single GFRG. So, an amalgamated technique like grey fuzzy logic method along with orthogonal array system provides an efficient way for optimizing the process parameters in micro-EDM.

References

1. Masuzawa, T.: State of the art of micromachining. CIRP Ann. Manuf. Technol. **49**(2), 473–488 (2000)
2. Schulze, H.P.: Importance of polarity change in the electrical discharge machining. AIP Conf. Proc. **1896**(1), 050001. AIP Publishing (2017)
3. Parshuramulu, A., Buschaiah, K., Laxminarayana, P.: A study on influence of polarity on the machining characteristics of sinker EDM. Int. J. Adv. Res. Eng. Technol. (IJARET) **4**(3), 158–162 (2013)

4. Kumar, A., Mandal, A., Dixit, A.R., Das, A.K.: Performance evaluation of Al2O3 nano powder mixed dielectric for electric discharge machining of Inconel 825. Mater. Manuf. Processes **33**(9), 986–995 (2018)
5. Tyagi, R., Das, A.K., Mandal, A.: Electrical discharge coating using WS2 and Cu powder mixture for solid lubrication and enhanced tribological performance. Tribol. Int. **120**, 80–92 (2018)
6. Gopal, A.V., Chakradhar, D.: Parametric optimization in electrochemical machining of EN-31 steel based on grey relation approach. Appl. Mech. Mater. **110**. Trans Tech Publications (2012)
7. Durairaj, M., Sudharsun, D., Swamynathan, N.: Analysis of process parameters in wire EDM with stainless steel using single objective Taguchi method and multi objective grey relational grade. Procedia Eng. **64**, 868–887 (2013)
8. Kumar, V.S., Kumar, P.M.: Optimization of cryogenic cooled EDM process parameters using grey relational analysis. J. Mech. Sci. Technol. **28**, 3777–3784 (2014)
9. Natarajan, N., Arunachalam, R.M.: Optimization of micro-EDM with multiple performance characteristics using Taguchi method and Grey relational analysis (2011)
10. Gopalakannan, S., Senthilvelan, T., Ranganathan, S.: Statistical optimization of EDM parameters on machining of aluminium hybrid metal matrix composite by applying Taguchi based grey analysis (2013)
11. Dewangan, S., Gangopadhyay, S., Biswas, C.K.: Multi-response optimization of surface integrity characteristics of EDM process using grey-fuzzy logic-based hybrid approach. Eng. Sci. Technol., Int. J. **18**(3), 361–368 (2015)
12. Ramanan, G., Shanavas, S.: Multi objective optimization of machining parameters for AA7075 metal matrix composite using grey-fuzzy technique. Int. J. Appl. Eng. Res. **12**(8), 1729–1735 (2017)
13. Soepangkat, B.O.P., Soesanti, A., Pramujati, B.: The use of Taguchi-grey-fuzzy to optimize performance characteristics in turning of AISI D2. Appl. Mech. Mater. **315**. Trans Tech Publications (2013)
14. Winarni, S., Indratno, S.W.: Application of multi response optimization with grey relational analysis and fuzzy logic method. J. Phys.: Conf. Ser. **948**(1). IOP Publishing (2018)

Chapter 50
Multi-output Response Optimization for Overall Enhancement of Mechanical Characteristics Using Utility Approach for AISI 316L Austenitic Stainless Steel Using Microplasma Arc Welding

Srikant Prasad, Dibakor Boruah and Pintu Thakur

Abstract This paper presents multi-output advancement for microplasma arc welding (MPAW) process using utility methodology. It is an advanced unconventional welding process in which very thin (usually less than 1 mm) metal surfaces are joined using plasma arc. A successful trial has been made to weld 0.5-mm-thick 316L stainless steel sheet. An attempt has been made to optimize the MPAW parameters for a desired equal-weighted multi-output such as ultimate tensile strength, yield strength, percentage elongation, Young's modulus, and weld hardness at fusion zone and heat-affected zone. The design of experiments L_{27} orthogonal matrix was conducted using levels of several visually successful trials. An analytical tool analysis of variance and interaction has been clubbed with utility approach to get desired aim of experimentation. As per different work application, the different preferable weight can be incorporated into the output. The confirmation test correlated with the best-desired aim of the output parameter. It was observed that overall utility for all mechanical properties and corresponding S/N ratio is enhanced with 0.28 and 0.42 values, respectively. It was found that welding speed with 49.01% and pulse time with 13.21% are the most significant parameters.

S. Prasad (✉)
Mechanical Engineering, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India
e-mail: shrikant@iitg.ac.in

D. Boruah
Faculty of Engineering, Environment and Computing, Coventry University, Coventry CV1 5FB, UK
e-mail: boruahd@uni.coventry.ac.uk

P. Thakur
Girijananda Chowdhury Institute of Management and Technology, Tezpur 784501, Assam, India
e-mail: pintuthakur07@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_50

The enhancement in the responses confirmed the approach of the feasibility of utility analysis which can help to enhance the product reliability for an industrial application.

Keywords MPAW · Optimization · Overall utility · ANOVA

50.1 Introduction

Microplasma arc welding (MPAW) was invented first by Robert Gage in 1957, which was introduced in 1964 to the welding industry for better control of the arc-welding process in lower current ranges. The plasma gas (the fourth state of matter) is used to generate arc with the help of argon gas. An ionization potential of plasma gas is 15.6 eV [1]. It takes very low currents to generate plasma arc. The argon gas is preferred as plasma gas because its low ionization potential ensures reliable starting and dependable pilot arc [2]. This arc is utilized to weld very thin plates even less than 0.5 mm. Low-alloy stainless steel sheet 316L is the useful engineering materials for industrial applications since they enriched with an excellent combination of properties like high strength-to-weight ratio, corrosion resistance, good toughness, fatigue life, which make it attractive for different applications through advanced joining process like MPAW. During welding, several input parameters which influence the welding properties need to be optimized for the better performance in the field application. The welding input parameters influence weld joint quality during the welding process. Therefore, it can be considered as a multi-input and multi-output process. Recently, a few researchers have reported optimization of 304L and Inconel considering effect of grain size, bead geometry, hardness, and ultimate tensile strength (UTS) using pulse microplasma [3]. The pulsed current microplasma arc-welding parameters were optimized to maximize UTS for AISI 304L sheets material using response surface method [4] and Hooke and Jeeves Algorithm by Shiv Prasad et al. [5]. The latest research work on the optimization of hydroxyapatite synthesis and microplasma spraying of porous coatings onto titanium implants has been reported by Alontsevadalontseva et al. [6]. According to the authors best of knowledge, no literature reported optimization of AISI 316L austenitic stainless steel sheet of using pulse MPAW. The various tests have been conducted to evaluate mechanical properties like tensile test, hardness, and weld bead geometry. As a multi-output response ultimate tensile strength (UTS), yield strength (YS), percentage elongation (%E), young modulus (YM), hardness of fusion zone (H_{FZ}) and hardness of heat-affected zone (H_{HAZ}) were used to study the optimization process using utility approach which helped to quantify the integrated performance of the MPAW. The confirmation experiment was conducted to validate the optimized parameters of the process. The predicted parameters have been correlated with the experimental result of the optimized parameters. The optimal welding conditions help to increase the productivity, minimize the total operating cost and strengthening of weld joint.

50.2 Experimental Work

The welding experiments were conducted using Aircraft plasma MP-50. The stainless steel 316L sheet of size (100 mm × 150 mm), thickness 0.5 mm was welded without filler material. The experiment had been designed as per design of experiment (DOE), using orthogonal array L27 Taguchi fractional factorial design method. The Argon gas was used as an orifice gas and shielding gas. The shielding gas prevents the weld from oxidation during welding and welds remain shiny. The coupling of orifice and shielding gas made with help of "T" joint attached at the end of cylinder exit pipe. The transferred arc mode in which the torch is brought in close to the workpiece, welding current is initiated, and the arc is transferred from the electrode to the workpiece which produces excess temperature up to 1700 °C to meet the welding requirement, then it cooled down to permissible temperature by means of circulating water through the nozzle. The complete experimental setup and welding fixture have been developed at IIT Guwahati central workshop. The setup has been shown in Fig. 50.1.

50.2.1 Weld Direction Determination for Rolled Sheet

To enhance percentage elongation, the weld direction has been determined with special light scattering technique which always shines in the direction perpendicular to the rolling direction. Due to rolling, grains are elongated in the direction of rolling resulting in less percentage elongation and hardness in the rolling direction. Therefore, care has been taken during the experiment and welding direction has been proposed normal to the light scattering direction [2]. Figure 50.2a illustrates the weld direction and rolling direction with the help of light scattering direction.

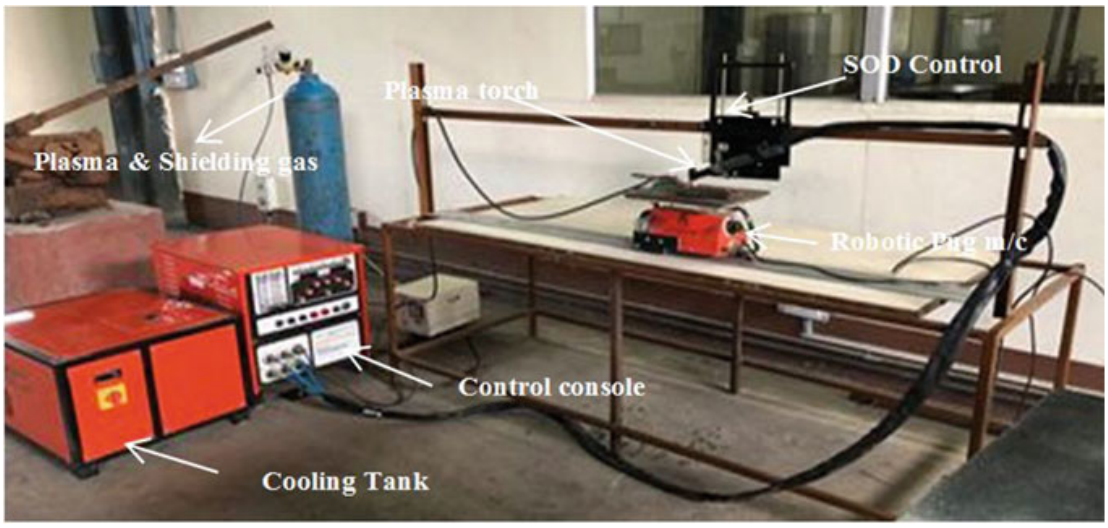


Fig. 50.1 Microplasma arc-welding (MPAW) setup

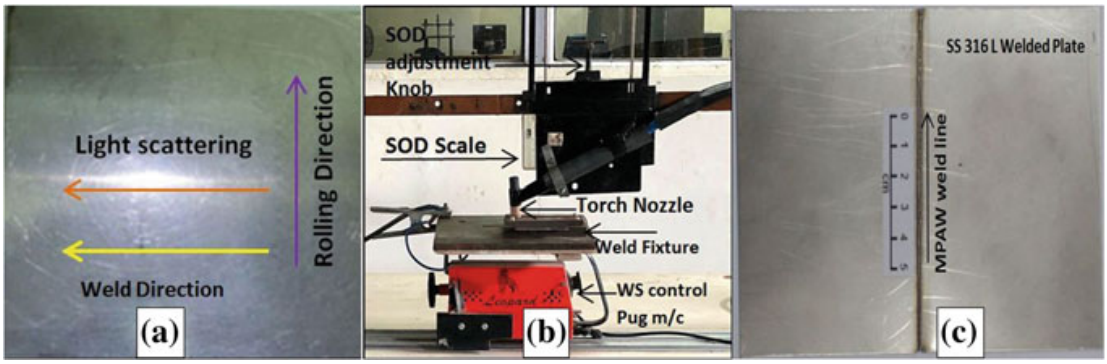


Fig. 50.2 **a** Weld direction and rolling direction, **b** torch and fixture equipped with SOD adjustment, **c** welded plate

Table 50.1 Stainless steel 316L chemical composition

Chemical composition (%) of SS 316L	C	Mn	Si	P	S	Cr	Mo	Ni	N
	0.03	2.0	0.75	0.045	0.03	18.0	3.00–2	10–14	1.0

The yellow arrow shows the rolling direction, and the blue arrow shows the welding direction.

The welding setup, fixture, and standoff distance adjustment features for MPAW setup were fabricated at IIT Guwahati central workshop for accuracy of weld quality, which has been shown in Fig. 50.2b, a welded sample was shown in Fig. 50.2c. The weld fixture was also designed as per dimension of weld sheet and less arc deflection. The interface gap between two weld plates, which is not permissible was ensured properly before welding. The chemical composition of the weld material used in these experiments has shown in Table 50.1. The significant six parameters in three levels which have an impact on weld quality, taken into consideration and shown in Table 50.2.

During post-weld process, tensile samples were cut through CO_2 laser-cutting machine, make: model no: LVD Orion3015, power rated 2.5 kW. The hardness samples (5 mm × 5 mm) and tensile samples as per the dimension of standard

Table 50.2 Details regarding the welding parameters and their levels

Sl. no.	Parameters (unit)	Abbreviations	Level 1	Level 2	Level 3
1	Current (A)	IP	11.9	13.8	15.3
2	Welding speed (mm/s)	WS	2.7	3.7	4.7
3	Plasma gas flow rate (L/s)	PGFR	0.6	0.7	0.8
4	Pulse time (ms)	PT	30	40	45
5	Pulse frequency (Hz)	PF	35	45	50
6	Standoff distance (mm)	SOD	1	1.5	2

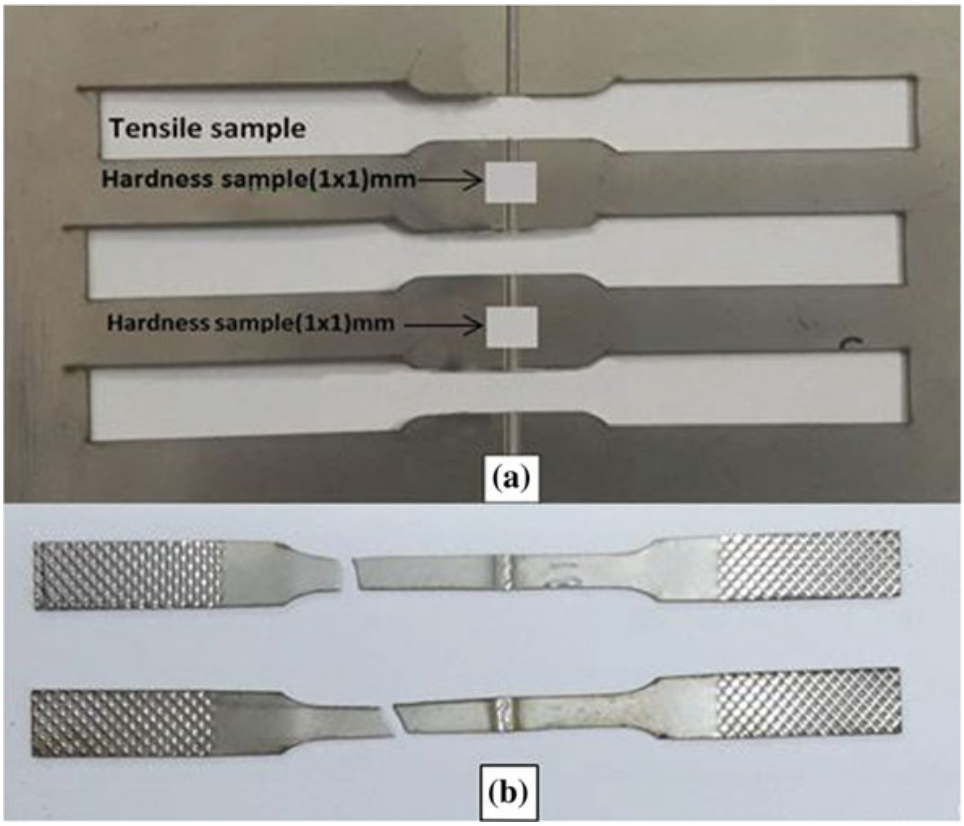


Fig. 50.3 **a** Extraction of tensile, hardness, and microstructure sample from a welded plate, **b** tested tensile samples

ASTM–B557M-10 were extracted. Figure 50.3a shows the extraction of tensile, hardness, and microstructure sample, and Fig. 50.3b shows the tested tensile sample.

The material properties considered in the present investigation are: UTS, YS, YM, %E, $H_{FZ,}$ and H_{HAZ}. The bead geometry is the mirror of the mechanical properties of a weld which depends upon reinforcement height and bead width [5]. The proper reinforcement height is required to get good strength of the welding [6]. All the mechanical output properties have been shown in Table 50.3.

Table 50.3 Experimental data of response characteristics

Exp. no.	UTS (MPa)	Avg. YS (MPa)	YM (GPa)	%E	H_{FZ} (HV)	H_{HAZ} (HV)
1	794.72	385.55	97,061.72	0.62	195.00	207.00
2	768.46	395.00	90,436.35	0.61	212.00	206.00
3	769.07	419.39	10,2576.10	0.68	188.00	208.00
4	743.43	377.71	106,978.70	0.59	240.00	199.00
5	792.52	392.19	100,899.00	0.61	211.00	209.00
6	803.36	375.11	98,579.45	0.59	199.00	212.00

(continued)

Table 50.3 (continued)

Exp. no.	UTS (MPa)	Avg. YS (MPa)	YM (GPa)	%E	H_{FZ} (HV)	H_{HAZ} (HV)
7	577.31	353.26	97,777.82	0.54	190.00	189.00
8	592.03	323.37	95,458.42	0.58	199.00	196.00
9	668.22	136.43	106,095.70	0.35	187.00	199.00
10	722.15	365.03	98,463.28	0.60	194.00	193.00
11	710.71	343.73	88,487.05	0.59	218.00	217.00
12	791.57	365.46	99,683.59	0.59	209.00	196.00
13	756.22	360.67	108,467.80	0.58	222.00	210.00
14	752.82	349.91	100,781.40	0.57	227.00	202.00
15	758.56	355.83	100,782.40	0.69	183.00	198.00
16	714.19	323.44	97,160.93	0.55	200.00	195.00
17	745.25	121.60	97,984.79	0.49	206.00	200.00
18	775.23	336.01	105,061.80	0.52	201.00	203.00
19	677.32	326.22	106,704.80	0.55	222.00	245.00
20	681.98	339.42	109,134.40	0.55	228.00	221.00
21	773.71	362.02	105,543.20	0.55	206.00	201.00
22	751.79	356.41	89,591.68	0.55	235.00	235.00
23	776.90	358.99	97,847.78	0.59	193.00	192.00
24	764.09	353.50	103,777.10	0.55	209.00	199.00
25	485.53	322.49	100,041.60	0.54	240.00	199.00
26	501.92	358.29	101,397.30	0.56	214.00	212.00
27	505.21	328.08	99,956.11	0.54	195.00	198.00
AVG	709.417	340.1891	100,249.3	0.568874	208.2593	205.2222

50.3 Result and Discussion

50.3.1 Optimization Using Utility Approach

Multi-response optimization is the most prominent area in the research for the welding industries to meet satisfaction, reliability, and cost reduction. Nowadays, unconventional welding process like MPAW is a very successful welding technology to weld materials with a very low thickness. Very precise control on parameters is required for its utility. Various process parameters need to be optimized to meet several demandable responses from the customer to manufacturer. The optimization of weld parameters is necessary to get optimal response for all quality characteristics. The multi-output characteristic is needed to scale for a single-output characteristic on a single composite scale index to get a single optimal setting of MPAW parameters [7–9]. The multi-response optimization is a need of today for industrial application where more than one response is coupled simultaneously using utility analysis. Several analytical techniques were explored to

handle the need for multi-optimization problem. Previously, some work has been reported for optimization is gray relation analysis, response surface methodology (RSM), and Taguchi methods. A research paper has been published on gray-based Taguchi method for parameters optimization of submerged arc bead on plate-welding process. Till now, no publication has been reported on multi-optimization using utility tool for MPAW process. For single output response Taguchi and RSM methodology are reliable optimization, but for multi-output welding process Taguchi analysis is not feasible [10–13]. Very few works are presented in literature for multi-response optimization problems [14, 15]. In comparison with utility approach, the desirability approach is not efficient up to a level. In the desirability approach, individual desirability is scaled into single composite index, for overall multi-response. Due to the complexity of computational aspects, several desirability values appear with zero value during the calculation which is difficult to analyze with ANOVA. Few publications on multi-output optimization have been reported on bead on plate welding. However, during the experiment, it was observed that the responses of real-welding process are different than the bead on plate welding [16]. Datta et al., have reported on submerged arc bead on plate welding using gray-based Taguchi technique. This research paper has been focused on overall enhancement of weighted mechanical properties for MPAW [17]. The mechanical property of a MPAW has been shown in Table 50.3.

50.3.2 Utility Concept Approach

The utility alludes to the fulfillment of customer; it contrasts from individual to individual, time to time and place, as defined by Prof. Hobson, "Utility is the capacity of a decent to fulfill a need". The business analyst estimated the utility in cardinal (numerical terms), yet there was no standard unit for estimating utility. In this way, the financial analysts utilized a fanciful measure, known as ('Util'). The costumers assess the item based on handiness. Taguchi configuration cannot deal with multi-output responses since every execution trademark have distinctive unit [18]. The general utility is the composite record of all quality as indicated by their value, which is appointed through the weight of the particular responses. A few research papers have been reported on Taguchi-constructed utility approach in light of various welding processes like Laser, EBM, and SAW process. However, no work has been accounted for on 0.5-mm-thin plate Austenitic steel 316L sheet welding through MPAW process. Subsequently, an effort has been made to represent the enhancement using utility-based Taguchi approach for multi-objective MPAW process for a thin austenitic steel sheet of 316L. This method was utilized to decide the ideal and best mix of process parameters for this process specifically welding pulse current, welding speed, pulse width, pulse frequency, plasma gas flow rate and stand of distance to maximize UTS, YS, %E, and hardness of the weld with the weight priority index to meet customer requirement.

50.3.3 Determination of Utility Value

All the output data has been converted into overall utility index to determine the utility for quality characteristics. In this approach, a preference scale has to construct using weight and composite numbers which is called overall utility. There are several preference scales, namely (i) exponential, (ii) logarithmic, and (iii) linear, but in the present research, investigation logarithmic approach has been considered as a preference scale index.

It is the evaluation of diverse quality characteristics in a rational choice by combining to make a composite index which represents the utility of a product. The overall utility index of a weld output is a sum of utilities of individual's quality characteristics.

$$U(X_1, X_2, \ldots\ldots\ldots\ldots, X_n) = \sum_{i=1}^{n} U_n(X_i) \tag{50.1}$$

$$P_i = A \times \log\left(\frac{X_{i,}}{X_{i,}}\right) \tag{50.2}$$

$$A = \frac{9}{\log\left(\frac{X_i^*}{X_{i,}}\right)} \tag{50.3}$$

50.3.3.1 Steps in Calculation of Utility

Step 1: In the utility analysis, the minimum acceptable quality is set numeric index 0 and the best acceptable quality level is set 9. Among the several scales mentioned above, the log scale has been chosen to find the preference number (P_i). The following equation is followed to find the constant value (A).

Step 2: For finding the value (A) using (P_i) value 9 in Eq. (50.3), we have to use (X_i*) as optimal pridicted output, and (X_i,), as minimum characteristic output among all.

Step 3: We have to find preference number (P_i) using Eq. (50.4). At this stage, we have to incorporate index (A) value and (X_i) individual present output carectersitic value, and (X_i), as minimum charectersitic output among all. The chart repersented below has been prepared using above Eq. 50.4 for all mechanical responcees like UTS-P, YS-P, YM-P, %E-P, $H_{FZ\text{-}P}$, and $H_{HAZ\text{-}P}$, respectively. The equal weightage has been considered for all outputs for predicted mechanical properties which are shown in Table 50.4. Depending on the customer requirements, the attributes to preference weightage may be given priority among all property. The property can be adjusted by incorporating weight (W) into individual utility index. Sum of all weights should not more than 1.

Table 50.4 Overall utility of data processing for predicted properties after assigning weights

Exp. no.	UTS-P W	YS-P W	YM-P W	%E-P W	$H_{FZ\text{-}P}$ W	$H_{HAZ\text{-}P}$ W
1	0.69	1.98	0.71	1.23	0.1	0.2
2	0.65	2.02	0.17	1.21	0.24	0.19
3	0.65	2.13	1.13	1.44	0.04	0.21
4	0.6	1.95	1.46	1.11	0.43	0.11
5	0.69	2.01	1.01	1.19	0.23	0.22
6	0.71	1.93	0.83	1.12	0.13	0.25
7	0.24	1.83	0.77	0.94	0.06	0
8	0.28	1.68	0.58	1.1	0.13	0.08
9	0.45	0.2	1.39	0	0.03	0.11
10	0.56	1.89	0.82	1.15	0.09	0.05
11	0.54	1.78	0	1.13	0.28	0.3
12	0.69	1.89	0.91	1.14	0.21	0.08
13	0.62	1.87	1.56	1.08	0.31	0.23
14	0.62	1.82	1	1.03	0.34	0.14
15	0.63	1.84	1	1.46	0	0.1
16	0.54	1.68	0.72	0.96	0.14	0.07
17	0.6	0	0.78	0.72	0.19	0.12
18	0.66	1.75	1.32	0.86	0.15	0.15
19	0.47	1.69	1.44	0.98	0.31	0.56
20	0.48	1.76	1.61	0.98	0.35	0.34
21	0.66	1.87	1.35	0.97	0.19	0.13
22	0.62	1.85	0.1	0.97	0.4	0.47
23	0.66	1.86	0.77	1.13	0.09	0.03
24	0.64	1.83	1.22	0.98	0.21	0.11
25	0	1.68	0.94	0.91	0.43	0.11
26	0.05	1.86	1.04	1.02	0.25	0.25
27	0.06	1.7	0.93	0.94	0.1	0.1

$$U(X_1, X_2, \ldots\ldots\ldots\ldots, X_n) = \sum_{i=1}^{n} W_i U_n(X_i) \tag{50.4}$$

A methodology of combining Taguchi and utility concept was introduced for predicting the optimal parameter settings of the combined responses. The six quality characteristics, namely UTS, YS, YM, %E, H_{FZ}, H_{HAZ}, during the experimental observation ratio of top width to top reinforcement were considered in this study for best output. The response parameters were individually optimized using Taguchi's technique to obtain the optimal input settings for response characteristics of MPAW, and using corresponding value, the mean for responses was predicted. These results are tabulated in Table 50.5.

Table 50.5 Optimal parameter setting for individual response characteristics

Response characteristics	Optimal level of process parameters	Predicted optimal value quality characteristics
UTS	$IP_2 + WS_2 + PGFR_1 + PT_2 + PF_3 + SOD_1$	919.99
YS	$IP_1 + WS_1 + PGFR_2 + PT_1 + PF_1 + SOD_1$	450.65
YM	$IP_3 + WS_2 + PGFR_3 + PT_2 + PF_3 + SOD_1$	111,895.1
%E	$IP_1 + WS_1 + PGFR_4 + PT_1 + PF_2 + SOD_1$	0.66
H_{FZ}	$IP_3 + WS_2 + PGFR_5 + PT_2 + PF_1 + SOD_3$	242.48
H_{HAZ}	$IP_3 + WS_1 + PGFR_6 + PT_2 + PF_1 + SOD_3$	232.88

Analysis of utility data for optimal parameter settings: the response graphs of overall utility and corresponding S/N ratio graphs plotted are shown below in Fig. 50.4.

The iteration between parameters (WS × SOD) and (PT × SOD) has been considered in the present study. The result reveals that third level of peak current (I_{P3}), first level of welding speed (WS_1), plasma gas flow rate (PGF_1), pulse time (PT_2), and third level of pulse frequency (PF_3) are expected to produce maximum values of utility and S/N ratio within the boundaries experimental zone. The ANOVA of utility data and S/N ratio are given in following Tables 50.6.

Analysis of Variance for mean of overall utility.

Where, SSQ = Sum of squares, MSS = Mean sum of square, DOF = Degrees of freedom, F = Fisher's ratio (the test statistic).

The utility approach reveals that welding speed (49.01%), pulse time (13.21%), and interaction of WS × SOD (12.11%), PT × SOD (11%) and standoff distance (9.57%). Interaction effect of welding speed and stand of distance are high

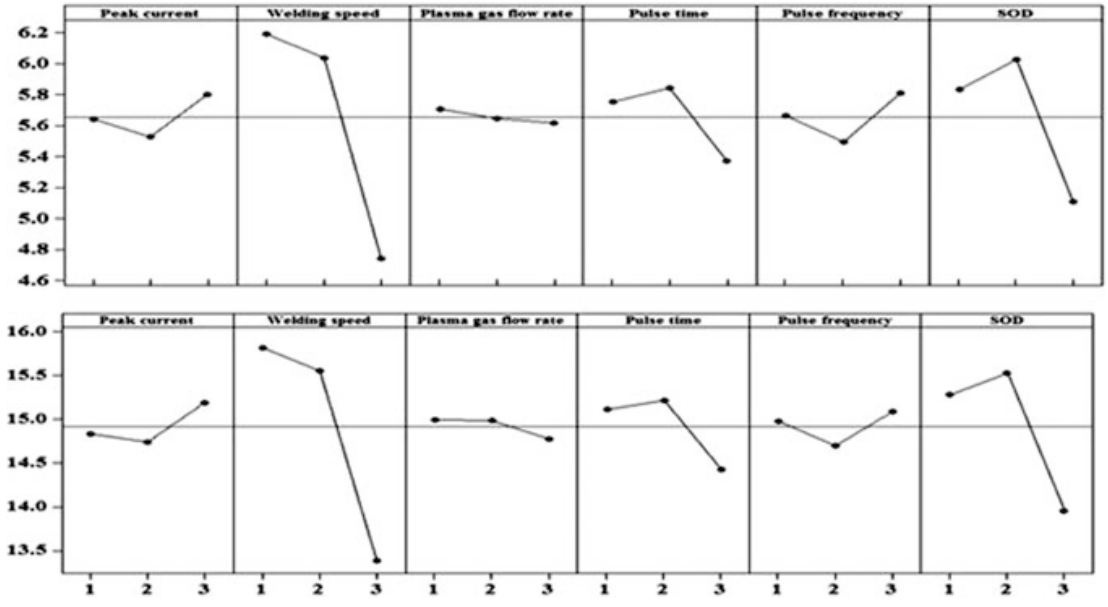


Fig. 50.4 Response graph of overall utility and S/N ratio

Table 50.6 Analysis of variance for mean of overall utility

Source (SS)	DOF	SSQ	MSS	Fisher's ratio	Percent contribution (%)
IP	2	0.23113	0.115565	0.575351	1
WS	2	9.73059	4.865295	24.22234	49
PGF	2	0.15301	0.076505	0.380888	1
PT	2	2.62276	1.31138	6.528831	13
PF	2	0.54734	0.27367	1.362492	3
SOD	2	1.89957	0.949785	4.728596	10
(WS × SOD)	4	2.134469	0.533617	2.656665	11
(PT × SOD)	4	1.326292	0.331573	1.650768	7
Residual error	6	1.205159	0.20086		6
SST	26	19.85032			100

compared with that of other input parameters even though other parameters such as peak current, plasma gas flow rate, and pulse frequency also contribute to the overall enhancement of the response characteristics. The percentage contribution of the input variables on the response characteristics is depicted by means of the pie chart Fig. 50.5.

The variation of overall utility with experiments is given in Table 50.7. It can be observed from the below table that experiment 9 and 17 have the lowest utility value, and experiment 13 and 24 have the highest utility value. The utility values are nearly uniform for experiments 1, 3, 5. From this variation, it can be inferred that the welding process is highly unpredictable. The predicted mean for overall utility has been tabulated based on Eq. (50.5) for the optimal parameter setting, where T_u is the total average of overall utility and μ_m is the predicted mean value.

$$\mu_m = \mathrm{IP}_3 + \mathrm{WS}_1 + \mathrm{PGFR}_2 + \mathrm{PT}_2 + \mathrm{FP}_3 + \mathrm{SOD}_2 - 5T_u \tag{50.5}$$

The predicted value of mean for overall utility obtained from the above formula is 5.6682 and corresponding S/N ratio is 15.0689.

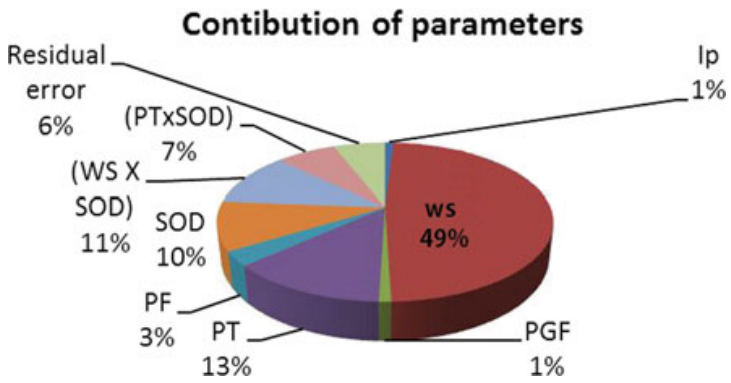


Fig. 50.5 Pie chart showing the relative contribution of process parameters

Table 50.7 Overall utility and their corresponding S/N ratio

Exp. no.	Overall utility	S/N	Exp. no.	Overall utility	S/N
1	4.914012	13.82872	15	5.029823	14.03105
2	4.469039	13.00428	16	4.108907	12.27453
3	5.597911	14.96052	17	2.411887	7.64714
4	5.656226	15.05054	18	4.886665	13.78025
5	5.342404	14.55474	19	5.445744	14.72115
6	4.976322	13.93817	20	5.516175	14.83276
7	3.843378	11.69426	21	5.174645	14.27761
8	3.854209	11.7187	22	4.401555	12.87212
9	2.184868	6.788505	23	4.537717	13.13675
10	4.557094	13.17376	24	5.002038	13.98294
11	4.024767	12.09482	25	4.069355	12.19051
12	4.920941	13.84096	26	4.468833	13.00388
13	5.668243	15.06897	27	3.832711	11.67012
14	4.951167	13.89415			

Maximum Overall utility and S/N ratio are 5.6682 and 15.06897

The obtained values of overall utility and S/N ratios for L_{27} are given in Table 50.7.

50.3.4 Confirmation Experiment

The confirmation experiment was performed using predicted value which was calculated from optimization using overall utility approach analysis. The best experiment was considered on the basis of maximum overall utility, not on the basis of particular single properties. The stress–strain curve obtained by this confirmation test which is target aim considering optimization using utility approach, as shown in Fig. 50.6.

The observed values of quality characteristics were correlated with experimental confirmation test results in the given Table 50.8.

The results ensure the existence and reliability of optimization coupling with utility approach for MPAW. It ensures that it is a cost-effective way to minimize the production cost and quality welds. It was observed from the analysis that enhancement in the overall utility is 0.28 and S/N ratio is 0.42, compared with the initial best setting. For the general requirements, UTS is considered dominant properties among all, so UTS properties have been highlighted in Table 50.9.

Hence, the model developed based on overall utility approach using Taguchi is well applicable and reliable to get desired weighted output for all properties of MPAW process.

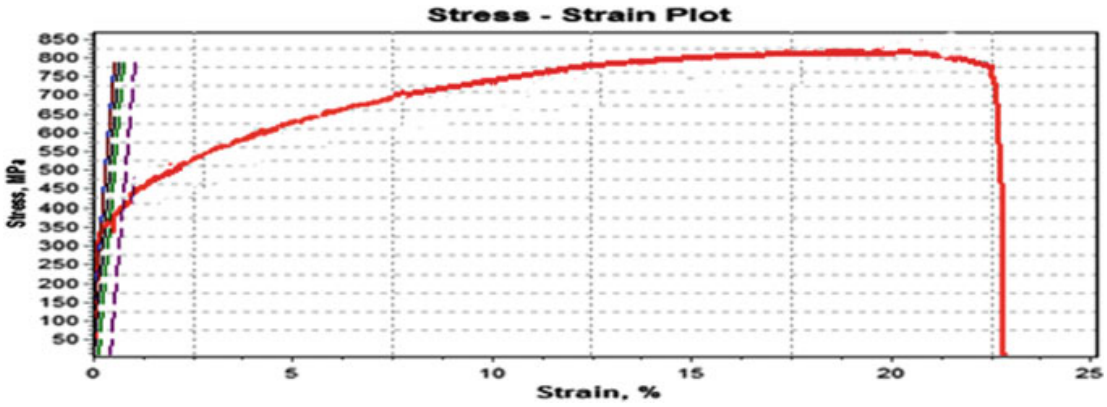


Fig. 50.6 Stress–strain curve of confirmation test

Table 50.8 Observed values of response characteristics for confirmation experiment

Response characteristics	Best factor setting for overall utility (Exp. 13)	Optimal process condition	
	IP_2, WS_2, $PG_{f/r3,}$ PT_1, PF_1, SOD_2	Prediction with optimum setting considering interactions (WS × SOD), (PT × SOD)	Confirmation experiment with optimal condition
UTS (MPa)	756.22	IP_3 WS_2 $PGFR_2$ PT_2 PF_1 SOD_1 (Based on mean responce graph)	823.00
YS (MPa)	360.67		552.93
YM (GPa)	108467.8		98956
%E	0.57		0.67
H_{FZ}	222		199
H_{HAZ}	210		212
Initial S/N ratio	15.06897 (Max)	15.9438	15.48
Overall utility for best setting	5.668243 (Max)	5.91000	5.94
Improvement in overall utility = (5.94 − 5.66) = 0.28 Improvement in S/N ratio = 15.48 − 15.06 = 0.42			

Table 50.9 UTS of confirmation test and some highlighted experiments

Sample name	UTS (MPa)
(a) Exp. 25	395.75 MPa (lowest value)
(b) Exp. 6	803.54 MPa (highest value for UTS)
(c) Exp. 13	756.21 MPa (UTS is moderate but overall utility as multi-output has highest significant value)
(c) Base metal	725 MPa (reference value)
(d) The enhanced value	823 MPa (enhanced UTS using utility approach)

50.4 Conclusion

1. The recognition of rolling direction of sheet to choose welding direction has a vital importance in welding, and it improves weld quality such as % elongation.
2. It was observed that enhancement in the overall utility is 0.28 and S/N ratio is 0.42 compared with the initial best setting. Hence, the model developed based on overall utility optimization approach is well applicable to get desired weighted output for the UTS, YS, %E, YM, $H_{FZ,}$ and H_{HAZ} properties of MPAW process.
3. ANOVA coupled with utility approach was performed to determine the significant parameters in this study, and it was found that welding speed with 49.01% and pulse time with 13.21% were most influential factors affecting the weld quality. Even though the contribution of other parameters was minimal but cannot be ignored in the study.
4. The confirmation test projects an enhancement in UTS which is 823 MPa. It is showing improvement in UTS from Exp. 6 and Exp. 13.
5. The optimization work using utility approach helps to enhance the productivity, reliability, quality, and cost reduction for an industrial application of the MPAW process.

References

1. Sanders, F.W., Hurst, G.S., Bortner, T.E.A.: Study of Alpha Particle in Argon Mixtures. Oak Ridge National Laboratory (1959)
2. Foest, R., Schmidt, M., Becker, K.: Microplasma an emerging field of low-temperature plasma science and technology. Int. J. Mass Spectrom. **248**, 87–102 (2006)
3. Shiv Prasad, K., Srinivasa, R.C., Nageswara, R.D.: Optimization of fusion zone grain size, hardness, and ultimate tensile strength of pulsed current microplasma arc welded AISI 304L sheets using genetic algorithm. Int. J. Manuf. Eng. 1–8 (2014)
4. Shiv Prasad, K., Srinivasa, R.C., Nageswara, R.D.: Optimizing pulsed current microplasma arc welding parameters to maximize ultimate tensile strength of AISI 304L sheets using response surface method. Int. J. Eng. Sci. Technol. 226–236 (2011)
5. Shiv Prasad, K., Srinivasa, R.C., Nageswara, R.D.: Optimizing pulsed current microplasma arc welding parameters to maximize ultimate tensile strength of SS304L sheets using Hooke and Jeeves Algorithm. J. Manuf. Prod. 39–48 (2011)
6. Alontsevadalontseva, D.L., Abilev, M.B., Zhilkashinova, A.M., Voinarovych, S.G., Kyslytsia, O.N., Ghassemieh, E., Russakova, A., Łatka, L.: Optimization of hydroxyapatite synthesis and microplasma spraying of porous coatings onto titanium implants, 79–95 (2018)
7. Yang, W.H., Tarng, Y.S.: Design optimization of cutting parameters for turning operations based on the Taguchi method. J. Mater. Process. Technol. **84**(1), 122–129 (1998)
8. Yang, J.L., Chen, J.C.: A systematic approach for identifying optimum surface roughness performance in end milling operations. J. Ind. Technol. **17**(2), 1–8 (2001)
9. Deng, C.S., Chin, J.H.: Hole roundness in deep hole drilling as analysed by Taguchi methods. Int. J. Adv. Manuf. Technol. **25**(5–16), 420–426 (2005)
10. Kumar, P., Creese, R.: Effect of Zircon Sand on Surface Finish of Al-7% Si Alloy V-Process Castings, pp. 95–98. AFS Transactions, USA (1999)

11. Creese, R., Kumar, P.: Effect of Fly Ash (As an Additive to Greensand) on the Surface Roughness of Al-7%Si Alloy Castings Using Taguchi Technique. AFS Transactions, USA (2001)
12. Lee, H.T., Yur, J.P.: Characteristic analysis of EDM surfaces using Taguchi approach. J. Mater. Process. Manuf. **15**(6), 781–806 (2000)
13. Kumar, P., Barua, P.B., Gaindher, J.L.: Quality of V-process method through the Taguchi technique. J. Qual. Reliab. Eng. **12**, 421–427 (1996)
14. Tong, L.I., Su, C.T.: Optimizing multi-response problems in the Taguchi method by fuzzy multiple attribute decision making. J. Reliab. Qual. Eng. Int. **13**, 25–34 (1997)
15. Antony, J.: Multi-response optimization in industrial experiments using Taguchi quality loss function and principle component analysis. J. Reliab. Qual. Eng. Int. **16**, 3–8 (2000)
16. Datta, S., Bandyopadhyay, A., Pal, Kumar P.: Grey-based Taguchi method for optimization of bead geometry in submerged arc bead-on-plate welding. Int. J. Adv. Manuf. Technol. **39**(11–12), 1136–1143 (2008)
17. Kumar, P., Barua, P.B., Gaindhar, J.L.: Quality optimization (multi characteristic) through Taguchi's technique and utility concept. Qual. Reliab. Eng. Int. **16**, 475–485 (2000)
18. Goyal, T., Walia, R.S., Sidhu, T.S.: Multi- response optimization of low-pressure cold-spread coatings through Taguchi method and utility concept. Int. J. Adv. Manuf. Technol. **64**, 903–914 (2013)

Chapter 51
Characterization of Geometrical Features of Ultra-Short Pulse Laser-drilled Microholes Using Computed Tomography

K. Kiran Kumar, G. L. Samuel and M. S. Shunmugam

Abstract High-aspect-ratio high-quality microholes are required in turbine blades to improve cooling performance. These cooling holes are drilled by pulsed laser and hence dimensional as well as geometrical tolerances like circularity and cylindricity are important. The measurement of geometrical features of the microholes is a very challenging task without destroying the components. In the present work, the microholes are produced on Ti6Al4V alloy by ultra-short pulse laser. The geometrical features of microholes are then captured using a non-destructive technique, namely computed tomography. CT-scanned 3D data is directly used for geometrical analysis using open-source software, GOM Inspect. Since algorithms used in the GOM Inspect are proprietary in nature, the extracted coordinate data are also analyzed using the computational methods developed by the authors based on least squares technique. The dimension, circularity, and cylindricity of microholes are compared with the results obtained from GOM Inspect software and a close match is found.

Keywords Microholes · Ultra-short pulse laser · Computed tomography · Geometrical analysis · Least squares technique

51.1 Introduction

The high-aspect-ratio microholes with tight tolerances are in high demand in automotive, aerospace, electronic, and biomedical industries. Specific applications include fuel injection nozzle, turbine blade and printed circuit board (PCB), inkjet printer nozzle and microneedles. The microholes for film cooling are drilled on turbine blade having complex geometry by ultra-short pulse laser. Unlike tool-based machining in which the geometry and axis of the hole are controlled by

K. Kiran Kumar · G. L. Samuel · M. S. Shunmugam (✉)
Indian Institute of Technology, Madras, Chennai 600036, India
e-mail: shun@iitm.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_51

the tool and its movement, the geometry of hole and its axis are determined by the ablation process in the laser drilling. Therefore, selection of parameters for proper control of laser ablation process is important to produce high-quality holes. It is imperative that the dimensional and form errors of hole produced by laser drilling are measured and characterized. Measurement of the dimensional and geometrical errors of microholes is a difficult task. The geometrical deviations of microhole are characterized as 2D and 3D errors, namely circularity and cylindricity. Different approaches have been reported in the literature to measure the dimensional and geometrical accuracy of microholes. The conventional methods reported in the literature mostly involve sectioning and polishing. These techniques require accurate axis alignment and they are also time-consuming.

It is important to characterize the microholes without destroying the machined holes. Kao and Shih [1] have investigated the form measurements of microholes using coordinate measuring machine (CMM) with a combination of optical and contact sensors. Non-destructive testing method based on optical system has been used for the inspection of cooling holes in turbine blade in aero-engine application [2]. Computed tomography (CT) is the only technology capable of measuring inner and outer geometry of components without destroying the machined surface. CT is able to detect pores or small delaminations within composites and measure small and hidden coolant bores of turbine blade [3]. CT technique has been used for industrial applications in the field of dimensional metrology [4]. Zhang et al., have reported measurement of features of the laser-drilled microholes using micro-CT [5]. 3D geometrical measurements of ablated crater such as volume, depth, and width and cone angle have been reported by Galmed et al. [6]. The characterization of geometrical features of microholes without destroying the holes still remains as a very big challenge.

In order to avoid destruction of machined component, non-destructive technique based on computed tomography is employed in the present work. CT is used to scan an array of 3×3 microholes drilled by ultra-short pulse laser micromachining. Further, the CT-scanned volume data is analyzed by using GOM Inspect available as open-source software. Dimensional and geometrical errors of laser-drilled microholes are also evaluated using computation technique developed by the authors based on least squares technique. The results are compared with those obtained from GOM Inspect software and discussed.

51.2 Microhole Drilling

The microhole drilling experiments are conducted by using ultra-short pulse laser. The Yb-doped fiber laser (make: Amplitude Systems; model: Satsuma HP2) operates at a central wavelength of 1030 nm with pulse duration in the range of 350 fs to 10 ps. The maximum average power is 20 W and pulse energy of 40 μJ is

delivered at a repetition rate of 200 kHz. The laser system is integrated with the ALPhNOV machine and laser beam is guided with trepanning head of GFH GmbH. Trepanning drilling technique is used to machine an array of 3 × 3 microholes in 1.5-mm-thick Ti6Al4V alloy plate.

51.3 Characterization of Microholes

The characterization of microholes is very challenging task. The conventional methods using sectioning and polishing are time-consuming and inaccurate due to axis misalignment. In order to overcome these problems, non-destructive technique like computed tomography (CT) is used. CT technology is capable of measuring the internal and external geometric features of complex-shaped components.

51.3.1 Computed Tomography

X-ray-based computed tomography technique is used for dimensional metrology of industrial components. The radiation emitted by a point-source X-ray passes through the measured object to a flat detector (sensor) and several 2D images are taken at various angular positions with the help of a rotating stage. 3D data reconstruction generates a 3D voxel model. Post-processing of voxel data using image processing technique detects the edges of the component for further analysis. The schematic diagram of computed tomography technique is shown in Fig. 51.1.

For precise measurement using CT, a high-quality tomogram is extracted from the detected surface. The CT scan image quality is mainly dependent on voxel size and focal spot size. The sharper images can be obtained with smaller X-ray spot size of high-power nanofocus tube. In the present work, nanotomphoenix|x-ray (make: GE, Germany) available at Centre for Non-destructive Evaluation (CNDE), Mechanical Engineering Department, IIT Madras, shown in Fig. 51.2 is used. The nanotom CT works at a tube voltage of 110 kV and current of 80 μA, with a minimum voxel size of <500 nm. The array of 3 × 3 microholes is scanned, and the scanning time is approximately 45 min.

From the generated volume data file, volume graphics (VG viewer) open-source software creates 3D and 2D images for visualization [7]. The 2D and 3D views of the component are shown in Fig. 51.3. In the present work, GOM Inspect open-source software [8] is used in the next step to generate the geometric model as shown in Fig. 51.4.

Computed tomography volume data (.vgl format) file is directly imported in GOM Inspect software, without converting into STL format. This is an easy

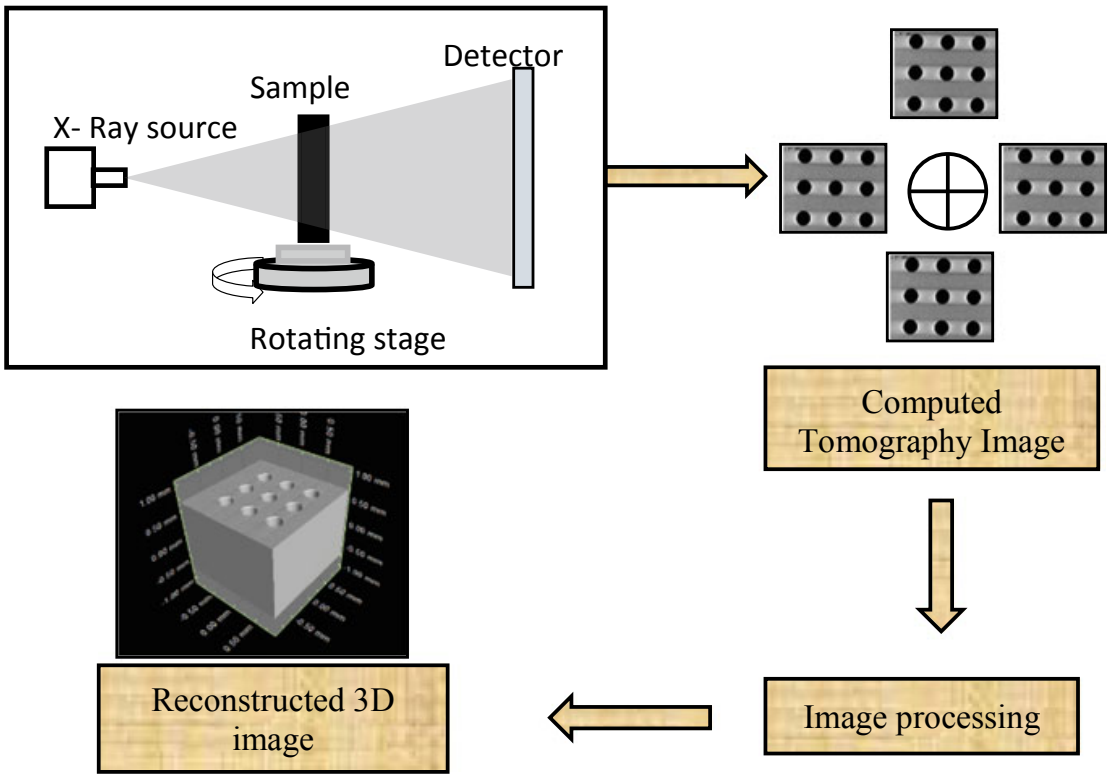


Fig. 51.1 Schematic diagram of computed tomography technique

Fig. 51.2 Computed tomography (CT) system at CNDE, IIT Madras

approach for the dimension analysis of CT-scanned complex shape and size. After importing CT volume data in GOM inspect software, it converts into mesh elements. Cylinder surfaces are constructed by selecting the holes. The coordinate transformation is done by using axis alignment with respect to z-axis in vertical direction. From the geometric model, dimensional and form errors of microholes are assessed as outlined in Sect. 51.4.

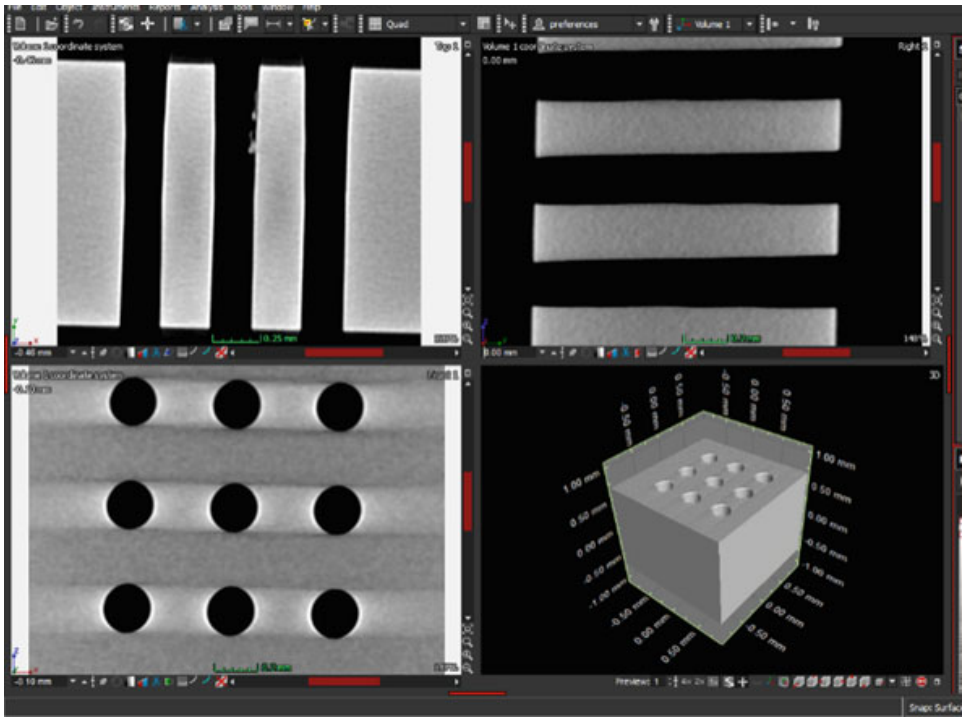


Fig. 51.3 Screenshot of reconstruction of CT-scanned images using volume graphics module

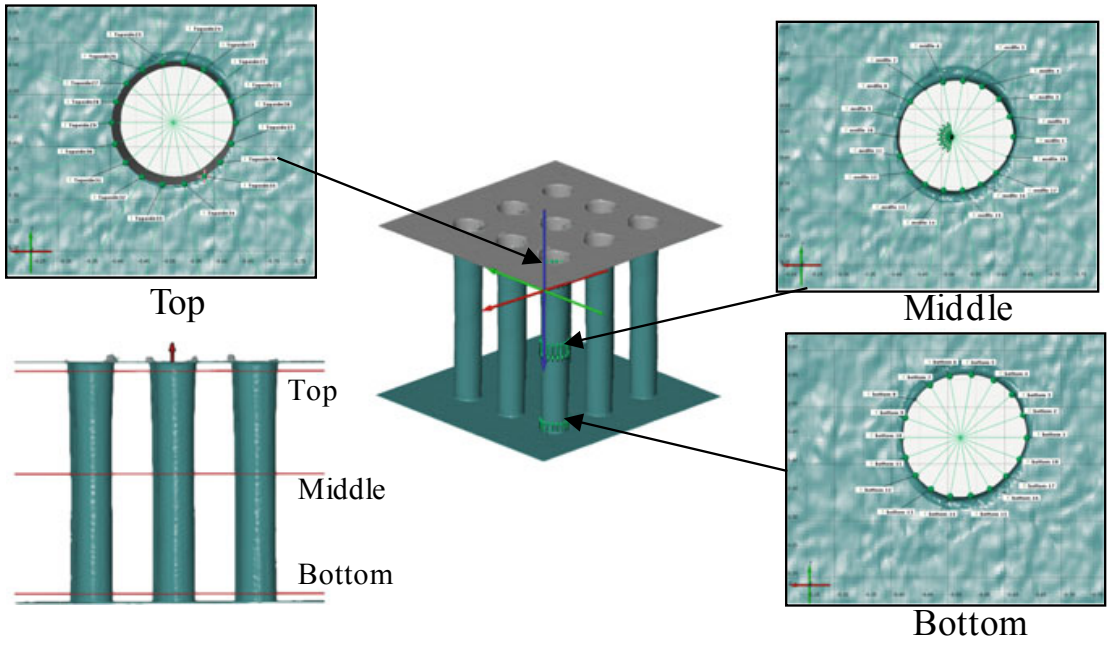


Fig. 51.4 Extracting the xyz coordinates from CAD geometry at three sections (top, middle, and bottom)

51.4 Computational Methods

To compare with the results obtained from GOM Inspect software, least squares technique developed by the authors is used for evaluation of form errors. The coordinates of microholes are extracted using GOM Inspect at selected sections as shown in Fig. 51.4. The *xyz* coordinates are extracted at top, middle, and bottom sections, respectively. The *xyz* coordinates at top surface of one microhole are given in Table 51.1.

The extracted coordinate data of microholes contains both size and form information, and a suitable method has to be implemented to separate the form data and size information. Samuel and Shunmugam [9] have proposed a transformation technique for evaluation of form error from the coordinate data. Accordingly, a reference circle is chosen passing through three selected points and the deviations are computed using Eq. (51.1).

$$d_i = \left[(x_i - x_R)^2 + (y_i - y_R)^2\right]^{1/2} - r_R \tag{51.1}$$

Table 51.1 *xyz* coordinate data (top, middle, and bottom sections)

Sl. no.	z' = 0.056 mm		z' = 0.681 mm		z'= 1.306 mm	
	x mm	y mm	x mm	y mm	x mm	y mm
1	−0.643	0.446	−0.637	0.446	−0.640	0.446
2	−0.633	0.487	−0.630	0.485	−0.633	0.487
3	−0.612	0.522	−0.609	0.519	−0.613	0.523
4	−0.581	0.548	−0.578	0.543	−0.581	0.548
5	−0.543	0.562	−0.541	0.554	−0.542	0.558
6	−0.502	0.562	−0.504	0.552	−0.503	0.555
7	−0.463	0.548	−0.470	0.538	−0.469	0.539
8	−0.432	0.522	−0.442	0.514	−0.442	0.514
9	−0.412	0.486	−0.423	0.482	−0.425	0.482
10	−0.403	0.446	−0.417	0.446	−0.417	0.446
11	−0.410	0.405	−0.423	0.410	−0.425	0.411
12	−0.430	0.369	−0.442	0.379	−0.443	0.380
13	−0.462	0.342	−0.471	0.357	−0.470	0.356
14	−0.501	0.328	−0.504	0.345	−0.504	0.343
15	−0.543	0.329	−0.540	0.344	−0.540	0.344
16	−0.582	0.343	−0.575	0.356	−0.574	0.356
17	−0.613	0.370	−0.604	0.378	−0.603	0.379
18	−0.634	0.406	−0.627	0.408	−0.627	0.408

Points for establishing reference circle (1, 4, 7) at all three sections

where; (x_i, y_i) are coordinates of extracted data. (x_R, y_R) is center of reference circle and r_R is its radius. The form data in the form of (r_i, θ_i) is computed as

$$r_i = d_i - d_{\min} + d_o \tag{51.2}$$

$$\theta_i = \tan^{-1}\left[\frac{(y_i - y_R)}{(x_i - x_R)}\right] \tag{51.3}$$

where; $d_{\min}$ is minimum value of d_i and d_o is taken as 0.001 mm.

In the present work, transformation approach is implemented for evaluating circularity and cylindricity of microholes and the results are discussed in Sect. 51.5. 2D and 3D form errors (circularity, cylindricity) of the microholes are evaluated by using least squares technique [10]. The deviations from the limacon and limacon cylinder are expressed in the linear form as follows.

$$e_i = r_i - (R_o + x_o \cos\theta_i + y_o \sin\theta_i) \tag{51.4}$$

$$e_i = r_i - [R_o + (x_o + l_o z_i)\cos\theta_i + (y_o + m_o z_i)\sin\theta_i] \tag{51.5}$$

where; e_i is linear deviation at ith point, R_o is estimated radius for limacon/limacon cylinder, (r_i, θ_i) are polar coordinates at ith point, (x_o, y_o) are estimated center coordinates and (l_o, m_o) are estimated slope values for limacon cylinder.

The least squares parameter for the limacon and limacon cylinder are estimated from the following equations.

$$\begin{bmatrix} \sum\cos^2\theta_i & \sum\sin\theta_i\cos\theta_i & \sum\cos\theta_i \\ \sum\sin\theta_i\cos\theta & \sum\sin^2\theta_i & \sum\sin\theta_i \\ \sum\cos\theta_i & \sum\sin\theta_i & \sum 1 \end{bmatrix} \begin{bmatrix} x_o \\ y_o \\ R_o \end{bmatrix} = \begin{bmatrix} \sum r_i\cos\theta_i \\ \sum r_i\sin\theta_i \\ \sum r_i \end{bmatrix} \tag{51.6}$$

$$\begin{bmatrix} \sum\cos^2\theta_i & \sum\sin\theta_i\cos\theta_i & \sum z_i\cos^2\theta_i & \sum z_i\sin\theta_i\cos\theta_i & \sum\cos\theta_i \\ \sum\sin\theta_i\cos\theta & \sum\sin^2\theta_i & \sum z_i\sin\theta_i\cos\theta & \sum z_i\sin^2\theta_i & \sum\sin\theta_i \\ \sum z_i\cos^2\theta_i & \sum z_i\sin\theta_i\cos\theta_i & \sum z_i^2\cos^2\theta_i & \sum z_i^2\sin\theta_i\cos\theta_i & \sum z_i\sin\theta_i \\ \sum z_i\sin\theta_i\cos\theta_i & \sum z_i\sin^2\theta_i & \sum z_i^2\sin\theta_i\cos\theta_i & \sum z_i^2\sin\theta_i & \sum z_i\sin\theta_i \\ \sum\cos\theta_i & \sum\sin\theta_i & \sum z_i\cos\theta_i & \sum z_i\sin\theta_i & \sum 1 \end{bmatrix}$$
$$\begin{bmatrix} x_o \\ y_o \\ l_o \\ m_o \\ R_o \end{bmatrix} = \begin{bmatrix} \sum r_i\cos\theta_i \\ \sum r_i\sin\theta_i \\ \sum r_i z_i\cos\theta_i \\ \sum r_i z_i\sin\theta_i \\ \sum r_i \end{bmatrix} \tag{51.7}$$

51.5 Results and Discussion

The microhole's dimensions are measured at top, middle, and bottom sections by fitting three-point circle at the respective sections in the GOM software. The diameters of microhole are obtained as 240, 220, and 220 μm at top, middle, and bottom sections, respectively, with height of 0.056, 0.681, and 1.306 mm, respectively, as shown in Fig. 51.5. If the practice of quantifying taper based on entry (top) and exit (bottom) diameters is followed, there is a general taper in the microhole with the size narrowing with the depth.

Similarly, the cylindricity error of microhole is measured as 11.2 μm by using Gaussian fit method, which is available in GOM Inspect software.

51.5.1 Comparison of Results with Least Squares Limacon Technique

The microhole geometrical xyz coordinate data given Table 51.1 is transformed to polar coordinates with respect to a reference circle. The reference circle is established by using a circle fitted through three selected points from given data. For example, center coordinates and radius of reference of circle passing through points 1, 4, and 7 are obtained as $x_R = -0.522$ mm; $y_R = 0.443$ mm and radius $r_R = 0.120$ mm, respectively. The deviations of each point with respect to reference circle are computed using Eq. (51.1). From the form data (r_i, θ_i), the form errors are evaluated with reference to a limacon. This is because, the center (x_R, y_R) chosen for transformation is different from the actual center of the circular feature. It has been shown in the literature that, it has been shown that the offset introduces certain distortion in the transformed profile and hence a limacon must be used as a reference for evaluating the circularity error form from data [11].

The transformed data (r_i, θ_i) is not listed in this paper due to space limitation. The data exhibits nonuniform angular spacing, and hence the general solution given by Eq. (51.6) is used to arrive at the limacon parameters. The results are plotted for top, middle, and bottom sections in polar chart as shown in Fig. 51.6. The parameters are given in Table 51.2, along with circularity errors and diameters. The circularity error is obtained as 3.182, 5.670, and 7.699 μm at top, middle, and bottom sections, respectively. From the above results, it observed that the quality of laser-drilled microhole changes with respect to depth. There is a good match between the diameter values obtained.

The cylindricity of microhole is evaluated by considering the height (z-coordinates). The xyz coordinates are extracted from geometric model for top, middle, and bottom sections as shown in Fig. 51.5. Table 51.3 gives value parameters obtained in the cylindricity evaluation. The cylindricity error of microhole is evaluated as

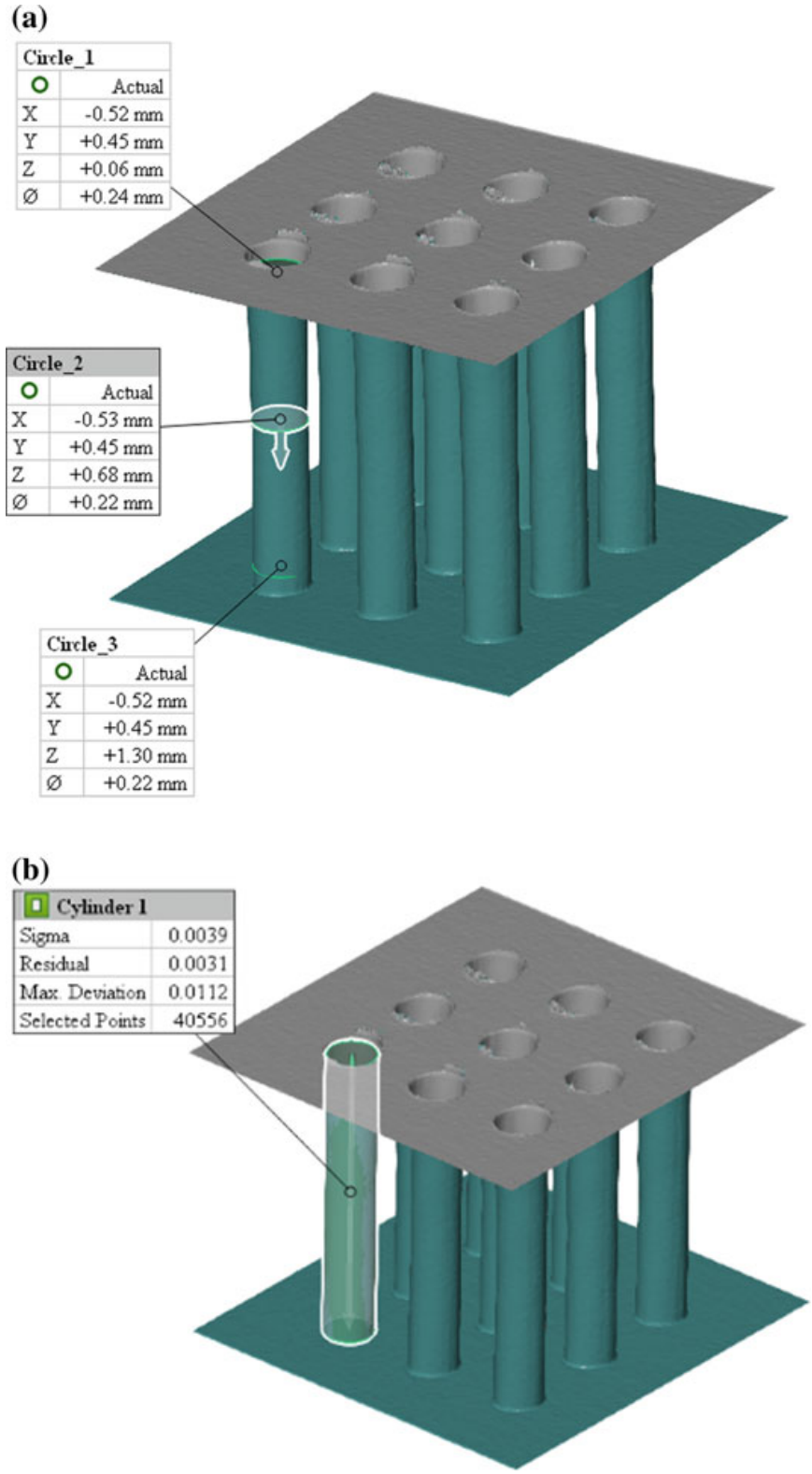


Fig. 51.5 Screen shot of microhole geometry measurement **a** diameter, **b** cylindricity

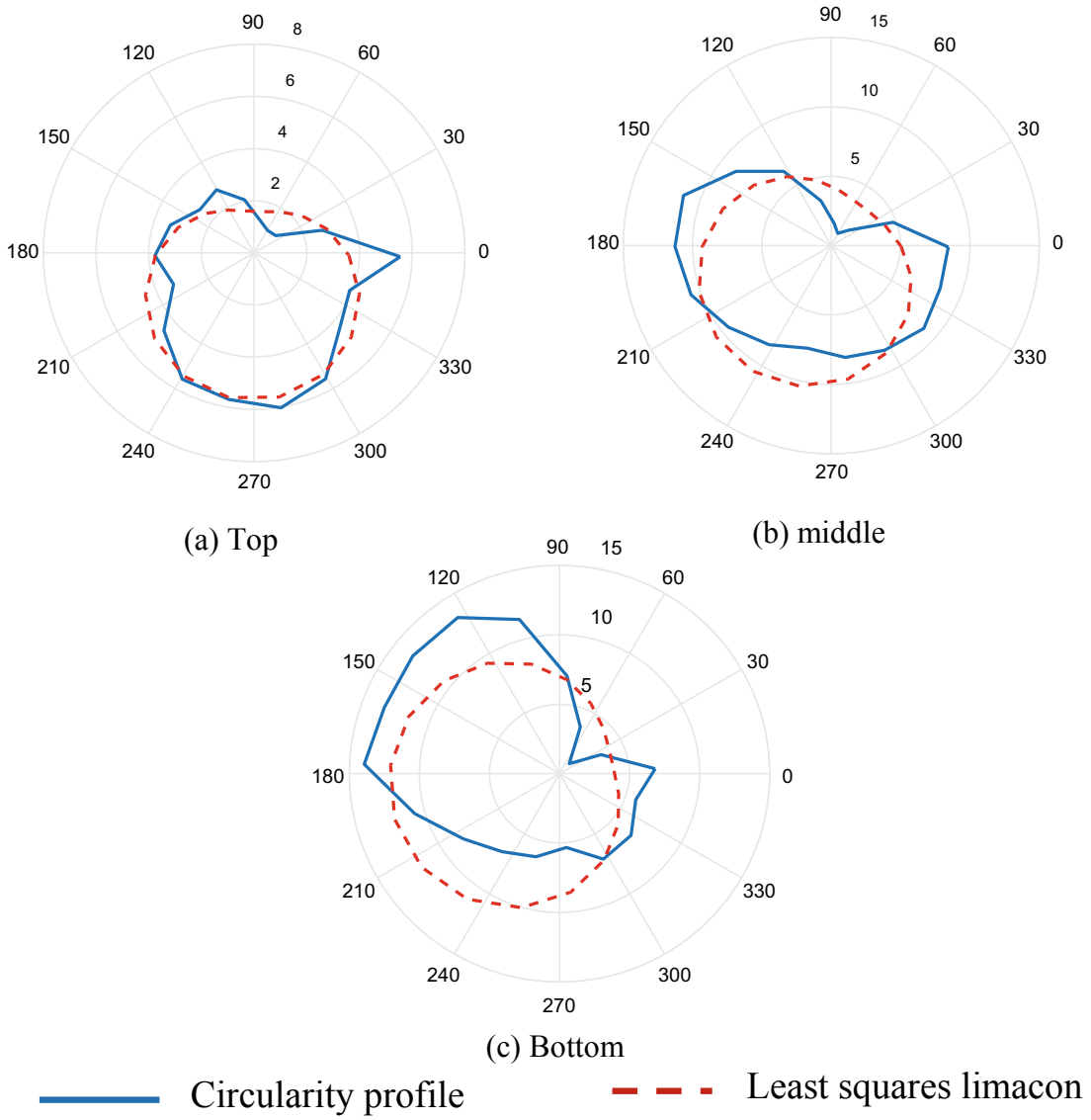


Fig. 51.6 Circularity profile along with least squares limacon at different sections

16.317 µm by using least squares technique. The cylindricity value of 11.2 µm is obtained from Gaussian fit method in GOM Inspect software. The value is lower than that computed from least squares limacon technique. Without the details of Gaussian fit method used in GOM Inspect, it is difficult to give reasons for the difference in the cylindricity errors. Interestingly, the diameter values agree with each other well.

In the above table, subscripts R and o refer to reference cylinder and limacon cylinder, respectively, while x, y represent the intersection of axis with top plane; l, m are the slope of axis; r is the radius of the cylinder.

Table 51.2 Circularity evaluation

	Top	Middle	Bottom
Initial values of reference circle			
x_R (mm)	−0.522	−0.528	−0.532
y_R (mm)	0.443	0.446	0.451
r_R (mm)	0.120	0.108	0.107
After fitting limacon			
x_o (μm)	−0.582	−2.096	−4.042
y_o (μm)	−2.045	−2.466	0.972
r_o (μm)	3.620	6.749	7.751
Circularity error (μm)	3.182	5.670	7.699
Diameter (mm)[a]	0.245	0.227	0.227
GOM inspect			
Circularity error (μm)	Not available in free version		
Diameter (mm)	0.24	0.22	0.22

[a]Diameter = $2\{r_R + [(r_0/1000) - 0.001]\}$

Table 51.3 Cylindricity evaluation (top section taken as $z = 0$ mm)

Initial values of reference cylinder	
x_R (mm)	−0.522
y_R (mm)	0.443
l_R (μm/mm)	−8.0
m_R (μm/mm)	6.4
r_R (mm)	0.107
Values of limacon cylinder	
x_o (μm)	0.321
y_o (μm)	−2.576
l_o (μm/mm)	−3.005
m_o (μm/mm)	2.399
r_o (μm)	10.721
Cylindricity error (μm)	16.317
Diameter (mm)[a]	0.233
GOM inspect	
Cylindricity error (μm)	11.2
Diameter (mm)	0.22

[a]Diameter = $2\{r_R + [(r_0/1000) - 0.001]\}$

51.6 Conclusion

New approach is established in the present work to characterize the geometrical features of the ultra-short pulse laser-drilled microholes. Using computed tomography, it is feasible to characterize the laser-drilled microholes without destroying the sample. The dimensional accuracy of computed tomography volume data is directly analyzed in GOM Inspect software. The diameters of microhole are 240,

220, and 220 µm at top, middle, and bottom sections, respectively, showing a taper in the microhole.

The form errors of microholes are established by least squares limacon technique. The circularity errors are 3.188, 5.670, and 7.699 µm at top, middle, and bottom sections, respectively. Cylindricity of microhole is evaluated as 16.317 µm using least squares limacon technique, which is higher than the result obtained from Gaussian fit in GOM Inspect software. Future work will be extended to evaluate the conicity error of microholes using computational technique.

Acknowledgements Authors would like to thank M/s Amplitude Systems, France, for providing the ultra-short pulse laser-drilled test sample. Authors also thank Aeronautical Research and Development Board (ARDB), Government of India (Project number: ARDB/01/2031768/M/I dated 10 Aug 2015) and CNDE, Mechanical Engineering Department, Indian Institute of Technology, Madras.

References

1. Kao, C.C., Shih, A.J.: Form measurements of micro-holes. Meas. Sci. Technol. **18**(11), 3603–3611 (2007)
2. Shetty, D., Eppes, T., Campana, C., Filburn, T., Nazaryan, N.: New approach to the inspection of cooling holes in aero-engines. Opt. Lasers Eng. **47**, 686–694 (2009). https://doi.org/10.1016/j.optlaseng.2008.12.001
3. Hanke, R., Fuchs, T., Uhlmann, N.: X-ray based methods for non-destructive testing and material characterization. Nucl. Instrum. Methods Phys. Res., Sect. A Accel. Spectrom. Detect. Assoc. Equip. **591**, 14–18 (2008). https://doi.org/10.1016/j.nima.2008.03.016
4. Kruth, J.P., Bartscher, M., Carmignato, S., Schmitt, R., De Chiffre, L., Weckenmann, A.: Computed tomography for dimensional metrology. CIRP Ann. - Manuf. Technol. **60**, 821–842 (2011). https://doi.org/10.1016/j.cirp.2011.05.006
5. Zhang, Y., Wang, Y., Zhang, J., Liu, Y., Yang, X., Zhang, Q.: Micromachining features of TiC ceramic by femtosecond pulsed laser. Ceram. Int. **41**, 6525–6533 (2015). https://doi.org/10.1016/j.ceramint.2015.01.095
6. Galmed, A.H., du Plessis, A., le Roux, S.G., Hartnick, E., Von Bergmann, H., Maaza, M.: Three dimensional characterization of laser ablation craters using high resolution X-ray computed tomography. Spectrochim. Acta - Part B At. Spectrosc. **139**, 75–82 (2018). https://doi.org/10.1016/j.sab.2017.11.011
7. Volume Graphics, https://www.volumegraphics.com/
8. GOM mbH, GOM Inspect Software. http://www.gom.com (2018)
9. Samuel, G.L., Shunmugam, M.S.: Evaluation of circularity and sphericity from coordinate measurement data. J. Mater. Process. Technol. **139**, 90–95 (2003). https://doi.org/10.1016/S0924-0136(03)00187-0
10. Shunmugam, M.S.: On assessment of geometric errors. Int. J. Prod. Res. **24**, 413–425 (1986)
11. Samuel, G.L., Shunmugam, M.S.: Evaluation of circularity from coordinate and form data using computational geometric techniques. Precis. Eng. **24**(3), 251–263 (2000)

Chapter 52
Investigation on Rotary Electrical Discharge Machining (EDM) of EN-25 Material Using Copper Electrode for Improving Geometrical Errors

L. Selvarajan, R. Rajavel, F. Leo Princely, R. Aravind and T. P. Habeeb Masood

Abstract In this analysis, an investigation and testing work were carried out on the spark EDM of EN-25 material using copper electrodes. Considering the machining parameters such as pulse on time, pulse off time, current, dielectric pressure for machining the effect of these following input parameters on output characteristics like metal removal rate (MRR), tool wear rate (EWR), wear ratio (WR), cylindricity (CYL), circularity (CIR), and perpendicularity (PER). The investigation was carried on with L_9 orthogonal array. The effect of each machining parameters on response characteristics was studied independently using trend analysis. From results, circularity has been minimized to decrease of current and increase of dielectric pressure. Cylindricity has been minimized to decrease of current and increase of dielectric pressure. Perpendicularity has been minimized to increase of current and decrease of dielectric pressure.

Keywords EDM · Geometrical tolerances · Circularity · Cylindricity · Perpendicularity

L. Selvarajan (✉) · R. Rajavel · T. P. Habeeb Masood
Department of Mechanical Engineering, Mahendra Institute of Engineering and Technology, Salem, India
e-mail: selvalakshmanan86@gmail.com

R. Rajavel
e-mail: catchrajavel@gmail.com

T. P. Habeeb Masood
e-mail: habi1@yahoo.com

F. Leo Princely
Department of Production Engineering, National Institute of Technology, Tiruchirappalli, India
e-mail: leoprincely@gmail.com

R. Aravind
Alstom Transport India Limited, Bengaluru, India
e-mail: aravindrajagopalan96@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_52

52.1 Introduction

As the world of manufacturing is growing at an exponential rate, the focus on surface finish and accuracy of parts produced is becoming highly important. This importance to finish and accuracy is evolving due to modern unconventional machining processes such as electric discharge machining. The shift to nontraditional machining is due to the exogenous nature of machining in case of conventional machining process. Owing to these current industry trends, an experimental study on the EDM of EN-25 is mandatory rather than that of a choice. It is used in automobile components and aerospace components. It is used for Adapters, Axles, Die Holders, Connecting Rods, Ejector Rods and Drill Shanks. Thus, in this investigation, EN-25 steel was used as the workpiece. The experiments were performed with EDM oil as dielectric fluid and 2-mm-copper tool. The input factors considered are current (A), pulse off time (μs) and pulse on time (μs) and dielectric fluid (kg/cm^2). This experimental investigation will increase the depth of knowledge of EDM and hence increase the propensity to success in terms of accuracy and surface finish.

52.2 Literature Review

Selvarajan et al. [1] in this work deal with the review of various papers related to optimization of EDM process. The electrode is chosen from regularly used Cu, CuW, and AgW as copper, which is based on the study by Muttamara et al. [2]. Uno et al. [3] spark EDM has emerged as an outstanding approach for cutting conductive metals that are otherwise highly difficult to machine. Lin et al. [4] in research introduce the multiresponse performance index is used to solve the spark EDM process with multiple responses or output characteristics. The machining parameters (the workpiece polarity, duty factor, pulse on time, and current, dielectric fluid and discharge voltage) are optimized with examinations of the multiple performance characteristics (electrode wear ratio and metal removal rate). Luis et al. [5] in this work, a comparative study of the spark EDM of various conductive ceramics had been carried out. Based on the experience acquired in the field of EDM for other materials, as, when the pulse time is reduced, which is equivalent to improving duty cycle, the MRR value tends to gradually increase. Arivazhaghan and Rajaduri [6] in this paper, illustration of the influence of input machining parameters on the metal removal rate is determined along with performance measurement analysis. The objective of this paper is to study the influence of operating technique parameters of copper electrode on the metal removal rate of EN-19 material followed by optimization. Soni and Chakraverti [7] in this paper describe optimum parameters for Titanium Grade (Ti-6Al-4V) using rotary spark machining process. The response of volume metal removal rate (MRR) and surface roughness (Ra) is considered for enhancing or increasing the machining efficiency.

Selvarajan et al. [8] EN-19 is tensile strength steel so it is widely used in different machinery parts such as shafts, racks, studs, bolts, nuts, rollers, etc. Wire EDM seems a good choice for machining the irregular shapes on medium steel Material. In the research, optimization of EDM parameters to improve geometrical errors by Selvarajan et al. [9] and the various methods to improve tolerance in $MoSi_2$–SiC was done. In another study by Selvarajan et al. [10], the optimization of machining parameters for improving circularity of Si_3N_4–TiN composite was done. Optimization of spark EDM Parameters in machining of $MoSi_2$–SiC composites for enhancing geometrical error was done by Selvarajan et al. [11]. It is evident that in all the above research studies conducted on either different materials or with different parameters. There is a noticeable research gap in terms of research on analyzing the machining of input parameters on geometrical output characteristics such as circularity, cylindricity, and perpendicularity. Thus, the research objective was chosen to be analyzing the individual impact of input parameters on each of these geometrical characteristics. The output parameters such as perpendicularity, cylindricity, and circularity are measured by using CMM, they are found applications in high accuracy.

52.3 Research Plan

Figure 52.1 shows that the research plan at EN-25 material with copper electrode.

52.4 Experimental Methodology

The experimental setup of EDM drilling machine is shown in Fig. 52.2. The workpiece material is shown in Fig. 52.3. The chemical composition of EN-25 steel was shown in Table 52.1

The properties of EN-25 steel are furnished in Tables 52.2 and 52.3 determine the working condition of the EDM drilling used for machining EN-25 steel.

The number of experiments to be conducted is nine and thus the L_9 orthogonal array was found to be suitable as instead of conducting 81 experiments based on factorial setup; this orthogonal array helps to finish the study within nine experiments. The values for each factor at various levels are shown below (Table 52.4).

This is the Taguchi's L_9 orthogonal array table where the experiment is done according to this procedure. The outputs are measured using a CMM whose image is provided in Fig. 52.4.

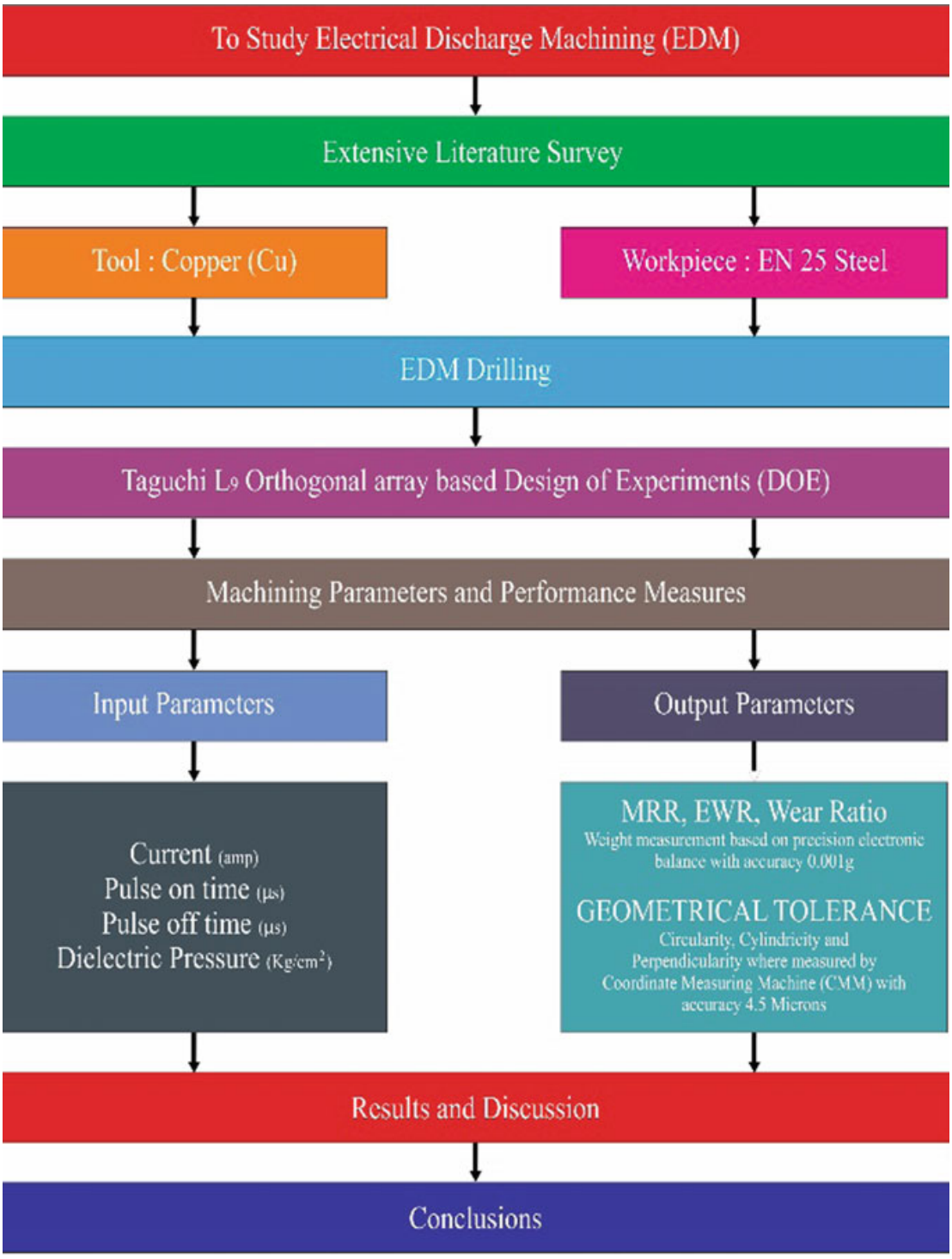


Fig. 52.1 Research plan

Fig. 52.2 Experimental setup of EDM drilling machine

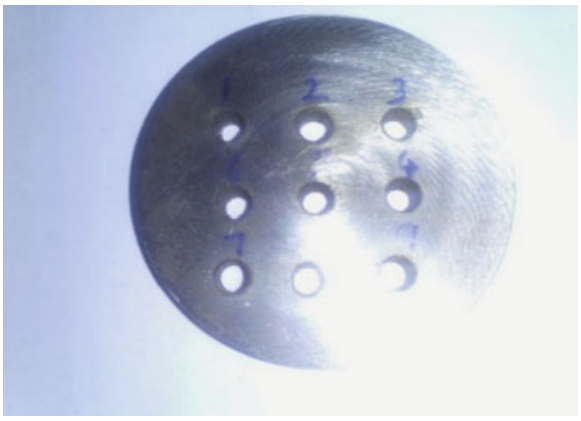


Fig. 52.3 EN-25 workpiece material

Table 52.1 Chemical composition of EN-25

C	Si	Mn	Ni	Cr	Mo	S	P
0.35	0.20	0.62	2.60	0.67	0.58	0.03	0.03

Table 52.2 Properties of EN-25 steel

Properties	EN-25 Steel
Density	7.59 (kg/m^3)
Thermal expansion	12.5 (10^{-6} k)
Melting point	1400 (°C)
Thermal conductivity	48.6 (W/m k)
Specific heat	1500 (J/Kg k)
Electrical resistivity	1200 (10^{-9} W m)
Tensile strength	980 (MPa)
Brinell hardness	627 (kg)

Table 52.3 EDM operating conditions for EN-25

Working conditions	Description
Electrode material	Copper
Diameter of drilling hole	3 mm
Depth of drilling hole	2.5 mm
Electrode polarity	Positive
Workpiece polarity	Negative
Specimen material	EN-25 steel
Type of current	DC power supply
Discharge current (I, A)	4–8
Pulse on time (t_{on}, μs)	20–30
Pulse off time (t_{off}, μs)	4–6
Dielectric fluid	EDM oil
Fluid pressure (kg/cm^2)	24–28

Table 52.4 Machining parameters and levels of EDM drilling

Factor	Parameter	Units	Level 1	Level 2	Level 3
A	Current	A	4	6	8
B	Pulse on time	μs	20	25	30
C	Pulse off time	μs	4	5	6
D	Dielectric pressure	kg/cm^2	24	26	28

52.4.1 MRR and TWR Formula

Measurement of metal rmoval rate:

- MRR = Machined workpiece weight loss (g)/machining time (min)

Fig. 52.4 CMM testing machine

Measurement of tool wear rate:

- TWR = Electrode weight loss (g)/machining time (min)

Measurement of percentage wear rate:

- Wear ratio is defined as the ratio of the tool wear rate to the material removal rate.
  %WR = EWR/MRR

52.5 Results and Discussion

Table 52.5 depicts the input parameters, namely current, pulse on time, pulse off time, and dielectric pressure. The weight of the electrode and workpiece before and after machining is taken to calculate MRR and EWR as per formula in Sect. 52.4.1

The machining time, material removal rate, electrode wear rate, wear ratio, circularity, perpendicularity, and cylindricity are shown in Table 52.6.

Now, the relationship between output parameters and each input parameter can be deeply studied. Firstly relationship with respect to current is studied in Fig. 52.5a–c.

The material and electrode wear rate are increased with increasing current. The geometrical tolerances like circularity, perpendicularity, and cylindricity increase with respect to current at a very gradual manner. We can also observe that the wear ratio is decreasing with current, and thus by decreasing current we can decrease wear ratio. Maximum metal removal rate (0.743 g/min) is obtained on the condition of current 7 A, pulse on time 30 µs, pulse off time 5 µs, and dielectric pressure 24 kg/cm^2.

Table 52.5 Input processes parameters at EN-25 material

Exp. no.	Input parameters				Workpiece weight		Electrode weight		Machining time
	A	*B*	*C*	*D*	Before machining	After machining	Before machining	After machining	
	A	µs	µs	kg/cm^2	g	g	g	g	min
1	4	20	4	24	25.168	24.992	19.795	19.758	1′ 12″
2	4	25	5	26	24.992	24.815	19.758	19.721	1′ 01″
3	4	30	6	28	24.815	24.631	19.721	19.687	48″
4	6	20	5	28	24.631	24.455	19.687	19.621	36″
5	6	25	6	24	24.455	24.273	19.621	19.573	42″
6	6	30	4	26	24.273	24.105	19.573	19.522	30″
7	7	20	6	26	24.105	23.947	19.522	19.417	35″
8	7	25	4	28	23.947	23.747	19.417	19.367	24″
9	7	30	5	24	23.747	23.576	19.367	19.314	14″

Table 52.6 Response parameters of EN-25 material use copper electrode

Output parameters						
S. no.	MRR g/min	EWR g/min	WR %	CIR mm	PER mm	CYL mm
1	0.146	0.030	20.55	0.0653	0.0292	0.1134
2	0.422	0.036	8.53	0.0361	0.0384	0.1201
3	0.230	0.042	18.26	0.0365	0.0155	0.0907
4	0.293	0.110	37.54	0.0595	0.0149	0.1274
5	0.260	0.068	26.15	0.0512	0.0359	0.1548
6	0.336	0.102	30.36	0.0497	0.0390	0.1490
7	0.272	0.181	66.54	0.0546	0.0452	0.1505
8	0.500	0.125	25.00	0.0429	0.0642	0.1246
9	0.743	0.230	30.96	0.0603	0.0140	0.1011

The relationship of performance characteristics with respect to pulse on time is shown in Fig. 52.6a–c.

The material and electrode wear rate are high when pulse on time is high and they are at maximum pulse on time of 30 µs. Cylindricity is decreasing with pulse on time, but there is only negligible change in perpendicularity and circularity with pulse on time. The wear ratio of the whole shows an increasing trend with respect to pulse on time. Minimum EWR (0.030 g/min) is acquired at the condition of current 4 A, pulse on time 20 µs, pulse off time 4 µs, and dielectric pressure 24 kg/cm^2.

The relationship between output characteristics and pulse off time is shown in Fig. 52.7a–c.

The material removal rate is showing a decreasing trend with pulse off time, and electrode wear rate has a negligible variation with respect to pulse off time. All geometrical characteristics, namely perpendicularity, cylindricity, and circularity

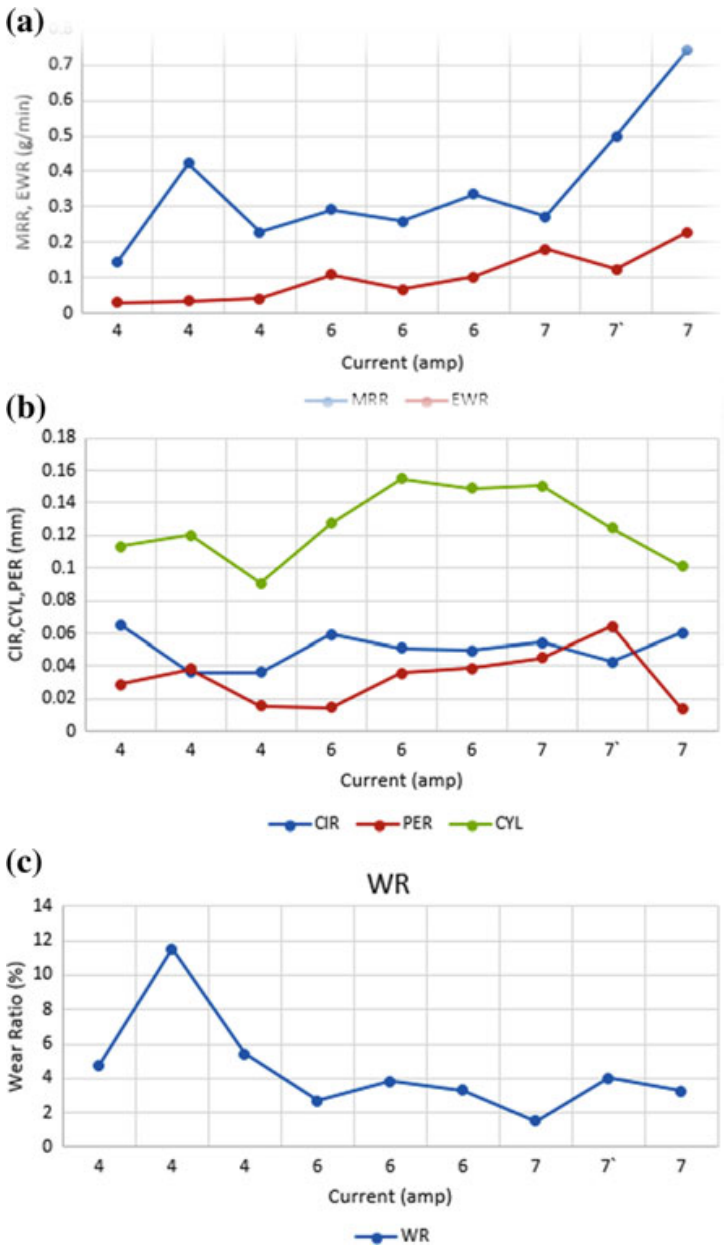


Fig. 52.5 **a** Relationship between current and MRR and EWR. **b** Relationship between current and CIR, PER, and CYL. **c** Relationship between current and wear ratio

show only a negligible change with respect to pulse off time. Whereas a gradual increasing trend is observed with wear ratio and pulse off time. Minimum circularity 0.0361 mm is acquired at the condition of current 4 A, pulse on time 25 µs, pulse off time 5 µs, and dielectric pressure 26 kg/cm^2. Minimum perpendicularity 0.0140 mm is acquired at the condition of current 7 A, pulse on time 30 µs, pulse off time 5 µs, and dielectric pressure 24 kg/cm^2. Minimum cylindricity 0.0907 mm is acquired at the condition of current 4 A, pulse on time 30 µs, pulse off time 6 µs, and dielectric pressure 28 kg/cm^2.

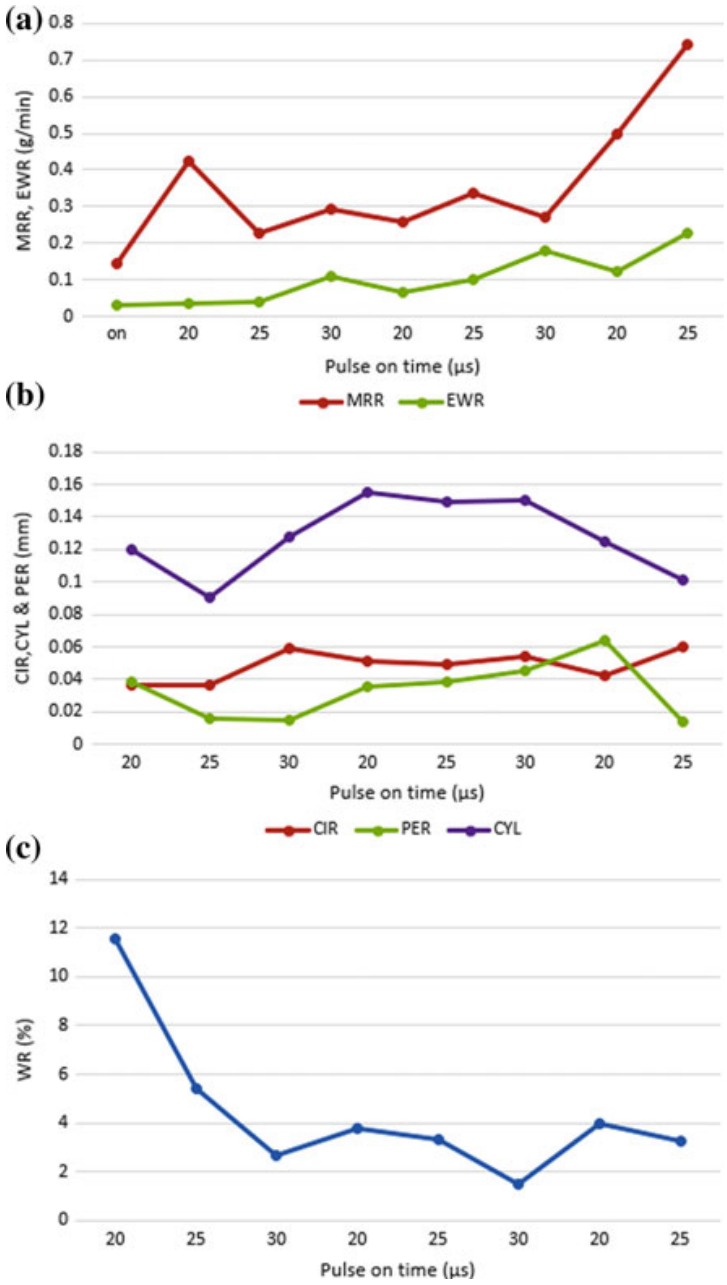


Fig. 52.6 **a** Relationship between pulse on time and MRR and EWR. **b** Relationship between pulse on time and CIR, PER, and CYL. **c** Relationship between pulse on time and wear ratio

The relationship of output characteristics with dielectric pressure is expressed in Fig. 52.8a–c.

There is only negligible change in both MRR and EWR, but at minimum dielectric pressure of (24 kg/cm^2), both are at their maximum. Cylindricity shows the gradual decrease with respect to dielectric pressure, circularity, and perpendicularity have not witnessed only negligible change with respect to dielectric pressure. There is a gradual decrease in wear ratio with respect to dielectric pressure. Minimum wear ratio 8.53% is obtained on the effect of current 4 A, pulse on time 25 µs, pulse off time 5 µs, and dielectric pressure 26 kg/cm^2.

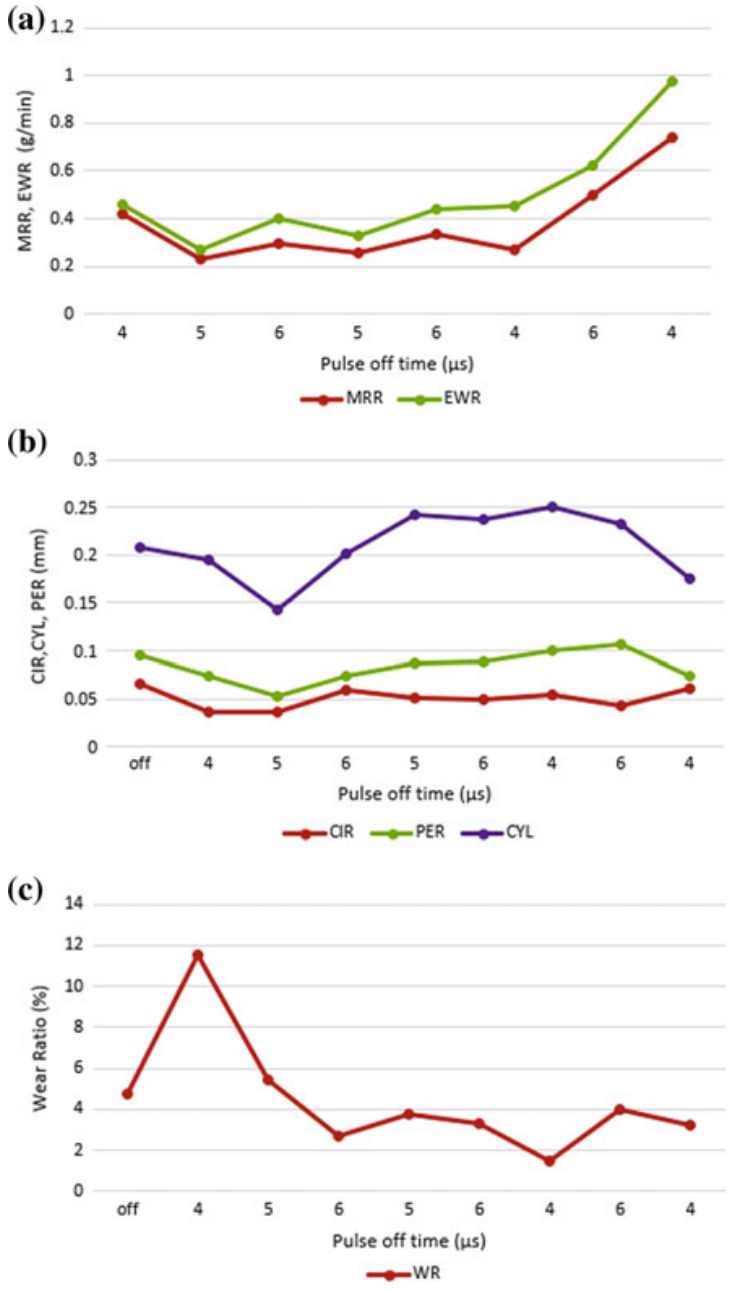


Fig. 52.7 **a** Relationship between pulse off time to MRR and EWR. **b** Relationship between pulse off time to CIR, PER, and CYL. **c** Relationship between pulse off time and wear ratio

52.6 Conclusions

- The three significant factors influencing the value of the MRR are the current (A), pulse on time (µs), and the pulse off time (µs). From results, MRR has been increased with increase current (A), pulse on time (µs), and decrease of pulse off time (µs). Metal removal enhancing increases with increase of pulse off time up to 6 µs then it's tended to decreases.
- The three main outstanding factors affecting the value of the EWR are the current (A), pulse on time, and the pulse off time. From results, EWR has been decreasing with decrease current (A), dielectric pressure (kg/cm^2). The value of

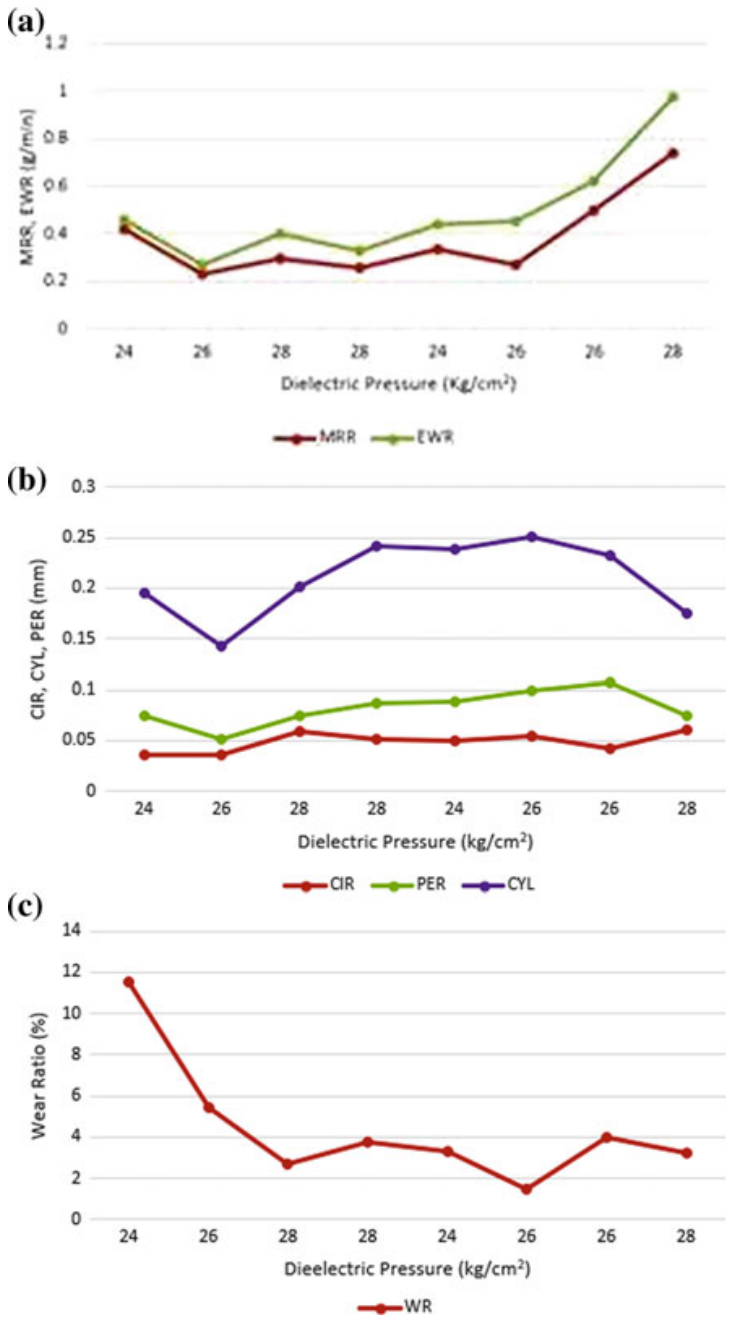


Fig. 52.8 **a** Relationship between dielectric pressure to MRR and EWR. **b** Relationship between dielectric pressure to CIR, PER, and CYL. **c** Relationship between dielectric pressure to wear ratio

EWR decreases with the decrease of pulse on time. However, EWR decreases with increase the pulse off time up to 6 μs.

- The two main notable factors affecting the geometric tolerance of circularity are the current (A) and dielectric pressure (kg/cm^2). From results, circularity has been minimized with decrease of current (A) and increase of dielectric pressure (kg/cm^2).
- The two main important factors affecting the geometric tolerance of perpendicularity are the current (A) and dielectric pressure (kg/cm^2). From results, perpendicularity has been minimized with increase of current (A) and decrease of dielectric pressure (kg/cm^2).

- The two main notable factors affecting the geometric tolerance of cylindricity are the current (A) and dielectric pressure (kg/cm^2). From results, cylindricity has been minimized with decrease of current (A) and increase of dielectric pressure (kg/cm^2).
- From the results of response table and response graph, the optimal level setting of four machining parameters is *A*, *B*, *C,* and *D* for minimizing the electrode wear rate and maximizing material removal rate among the nine experiments by using the Taguchi L_9 orthogonal array.
- The most important factors affecting the EWR and MRR of the spark EDM process have been identified as factor *A* (current), factor *B* (spark on time), factor *C* (spark off time), and factor *D* (dielectric pressure).

References

1. Selvarajan, L., Sathiya Narayanan, C., Jeyapaul, R.: Optimization of machining characteristics in EDM of Si_3N_4-TiN composites by Taguchi grey relational analysis. Appl. Mech. Mater. **592**, 600–604 (2014)
2. Muttamara, A., Fukuzawa, Y., Mohri, N., Tani, T.: Effect of electrode material on electrical discharge machining of alumina. J. Mater. Process. Technol. **209**(5), 2545–2552 (2009)
3. Uno, Y., Okado, A., Itoh, M., Yamaguchi, T.: Electrical discharge machining of a groove with a rotating disk electrode. Int. J. Electr. Mach. **1**, 13–20 (1996)
4. Lin, C.L., Lin, J.L., Ko, T.C.: Optimisation of the EDM process based on the orthogonal array with fuzzy logic and grey relational analysis method. Int. J. Adv. Manuf. Technol. **19**(4), 271–277 (2002)
5. Luis, C.J., Puertas, I., Villa, G.: Material removal rate and electrode wear study on the EDM of silicon carbide. J. Mater. Process. Technol. **164**, 889–896 (2005)
6. Arivazhaghan, P., Rajaduri, A.: Electronic discharge machining of Al/SiC carbide MMCs. In: Proceeding of the International Conference on advanced Mechanics and Industrial Engineering, University of Roorkee, p. 709 (1997)
7. Soni, J.S., Chakraverti, G.: Machining characteristics of titanium with rotary electro-discharge machining. Wear **171**, 51–58 (1994)
8. Selvarajan, L., Katherasan, D., Srivijai, B., Rajavel, R., Ramamoorthi, M.: Experimental analysis of EN 19 alloy material on EDM for improving geometrical errors using copper pentagon shaped electrode. In: International Conference of Materials Processing and Characterization, Materials Today Proceedings, vol. 5, pp. 4508–4514 (2018)
9. Selvarajan, L., Sathiya Narayanan, C., Jeyapaul, R.: Optimization of process parameters to improve form and orientation tolerances in EDM of MoSi2-SiC composites. Mater. Manuf. Processes **30**(8), 954–960 (2015)
10. Selvarajan, L., Sathiya Narayanan, C.S., Jeyapaul, R.: Optimization of EDM parameters on machining Si3N4–TiN composite for improving circularity, cylindricity, and perpendicularity. Mater. Manuf. Processes **31**(4), 405–412 (2016)
11. Selvarajan, L., Sathiya Narayanan, C., Jeyapaul, R.: Optimization of EDM hole drilling parameters in machining of MoSi2-SiC intermetallic/composites for improving geometrical tolerances. J. Adv. Manuf. Syst. **14**(04), 259–272 (2015)

Chapter 53
Computation of End-Cutting-Edge Wear of Single-Point Cutting Tool Using Image Processing

Vishal K. Singh and Shrikrishna N. Joshi

Abstract This paper presents a simple and efficient method for computation of cutting-edge wear of a single-point cutting tool in a plain turning operation. It comprises development of an efficient and accurate image processing-based computer algorithm which compares two images of the cutting tool, i.e., before and after its use in the machining operation. The first image has been used as the reference image. Location of the tool in the image was found out using the developed image processing algorithm. Vertical and horizontal magnifications of the tool image were adjusted such that both the images of the tool have same tool width in terms of pixels. The portion of the images containing tools was extracted and used for the wear computation. The performance of the algorithm was tested by conducting systematic experiments and it was found to be efficient and stable. Only 28 s was noted for the computation of tool wear based on the images captured by a simple digital camera. The algorithm efficiently removes the noise present in the images of worn-out cutting tool used in real-life conditions.

Keywords Plain turning · Tool wear · Cutting-edge wear · Computer vision · Image processing

53.1 Introduction

In the rapidly changing global marketplace, cost-cutting and boosting up productivity are the main goals of the manufacturing industry. With the introduction of manufacturing concepts like JIT, high-speed machining and precision manufacturing have gained more importance. Predicting exact tool life has been a challenge for manufacturing engineers for a long time. Thus, the tool rooms prefer preventive maintenance of the machine tools and cutting tools. However, often the capacity of

V. K. Singh · S. N. Joshi (✉)
Department of Mechanical Engineering, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India
e-mail: snj@iitg.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_53

the tool remains unused and thus increases the cost of production. Therefore, predicting tool life with a fair degree of accuracy is essential to carry out an efficient metal cutting process. Further, the realization of unattended machining processes is a prerequisite for ensuring successful operation of intelligent manufacturing systems (IMSs), flexible manufacturing systems (FMSs), and flexible manufacturing cells (FMCs) [1].

To analyze the performance of cutting tool life, dimensional tool life can be used as a parameter. Astakhov [2] defined the term 'dimensional tool life' as the time period within which the cutting tool assures the required dimensional tolerance and required surface integrity of the workpiece. Authors derived the mathematical equations based on the tribology of metal cutting. Literature reports eminent attempts on development of multisensor models to predict the tool failure [3–7]. Most of them are based on the indirect approach of measurement/monitoring and predict the time of catastrophic failures only [4]. Use of multisensor models is expensive and costly for being employed in real-world daily machining work. On the importance scale, computation of gradual wear ranks higher than catastrophic failure, as it controls the dimensional tolerance of the tool. Direct methods of measurement, particularly computer vision are by nature more suitable for monitoring gradual wear of the tool; but tedious [8]. Therefore, researchers worldwide focused their attention on measuring tool wear between the machining cycles and *ex situ* tool wear measurement.

Use of image processing in the field of tool wear monitoring dates back to the early 90s. Kurada and Bradley in their work proposed a tool wear monitoring system to measure the flank wear. This system was dependent on a He–Ne laser source for illumination [9]. Fadare and Oni proposed a system, which uses two light sources inclined at 45° to horizontal. In addition, two perpendicular views of the tool are captured [10]. Shahabi and Ratnam used workpiece roughness profile, obtained using computer vision for assessment of flank wear and nose radius wear [11].

Recent work in the field of tool wear computation and prediction involves computer vision; use of high-quality charge-coupled device (CCD) sensors and artificial neural network (ANN). ANN accounts for the unknown variables and parameters involved. Most of these research works have been focused on measurement of flank wear and crater wear. Flank wear is the major criterion for measurement of tool life but capturing images during machining is quite a challenging task; a lot of noise in form of chips, dirt, and cutting fluid is present in the background [8]. Tan et al. in their work used sensor placement and optimum fusion of sensors for real-time machine condition monitoring [12].

In the present work, a simple and efficient method for computation of cutting-edge wear of the single-point cutting tool in a plain turning operation is developed. It comprises of development of an efficient and accurate image processing-based computer algorithm which compares two images of the cutting tool, i.e., before and after its use in the machining operation. In what follows, the details of development of algorithm and its experimental validation are presented in the following sections.

53.2 Methodology

A true-color digital image is a 3-dimensional matrix arrangement of primary color pixels with different intensities, whereas a grayscale image is a 2-dimensional arrangement of pixel intensities. Since every digital image or video (series of images varying with time) is a matrix; therefore, two images can be differentiated using the contents of their matrix using mathematical operations. This is the basic idea behind using image processing to determine the tool wear. Use of computer algorithms to detect, improve, or compute features of the image, or the objects in the image is known as image processing. Feature extraction from image requires various steps to be followed which are presented in Fig. 53.1.

53.2.1 Image Acquisition

It is the first step of image processing. It comprises of camera calibration and image capturing.

Camera Calibration. In this step, the parameters of camera's lens and image sensor are estimated. These parameters help to correct the lens distortion, to measure the size of an object in world units, and to determine the location of the camera in the scene. Camera parameters include intrinsic, extrinsic, and distortion coefficients. To estimate the camera parameters, 3-D world points and their corresponding 2-D image points are required. Multiple images of a calibration pattern such as a checkerboard are used. Using the correspondences, camera parameters are estimated. After calibration of the camera, the accuracy of the estimated parameters could be evaluated using reprojection errors and parameter estimation errors. Figure 53.2 shows the developed image acquisition setup for capturing the images of worn-out cutting tool.

Image Capturing Setup. Image capturing step comprises of a click-type camera and LED light source (a set of three numbers of 10 W LED bulbs) placed at the bottom of a transparent and hollow box-shaped tool holder. A white paper sheet placed

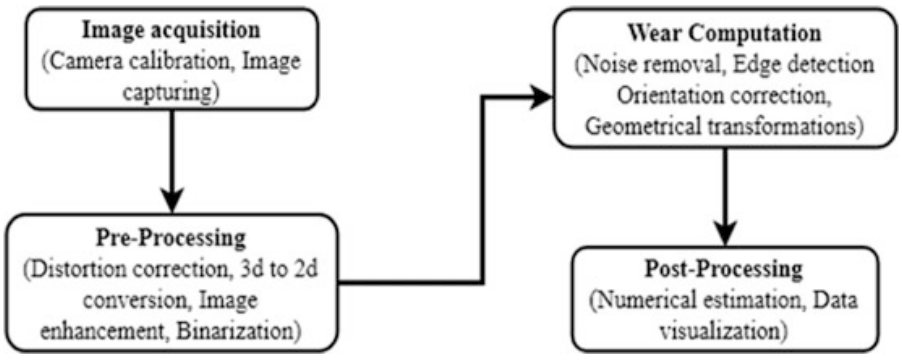


Fig. 53.1 Image processing flow diagram

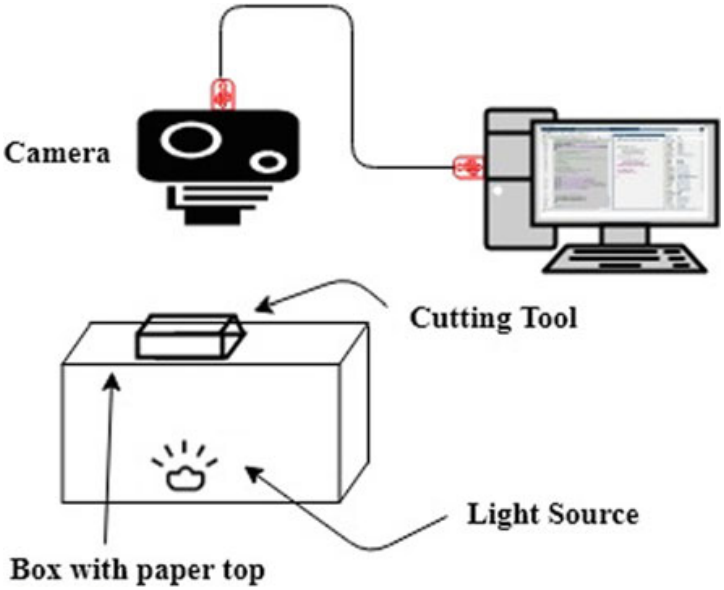


Fig. 53.2 Image acquisition setup

between the tool holder and tool produces a diffused light effect. The diffused light effect helps to minimize the light blobs appearing in the captured images due to the unbalanced distribution of the light from the illumination source. The cutting tool is placed on the marked space on tool holder. An image is captured on the camera which is placed directly above the tool holder. Two images of the tool are captured —one before the start of a machining run carried out for a chosen set of process conditions, and the second image captured after completion of the machining run. The Sony DSC-W570 camera captures a true-color image of size 4608 × 3456 and then, the images are imported into MATLAB™-based image processing utility.

53.2.2 Preprocessing

Distortion Correction. The environmental factors, namely temperature and humidity, affect the quality of images. This causes faulty representation of real-world objects in the image world, which reduces the precision in the measurement of parameters of an image. All the regions in an image cannot be assigned a proportionate number of pixels per world unit.

This problem was solved by using MATLAB™ undistorting function. It was used during camera calibration. Figure 53.3 shows the possible radial camera distortions.

3D to 2D Conversion. In this work, the true-color image [stacking of red (*R*), green (*G*), and blue (*B*) intensity distribution] is converted to a grayscale image. Equation (53.1) is used for conversion of true-color image to grayscale image, *I*.

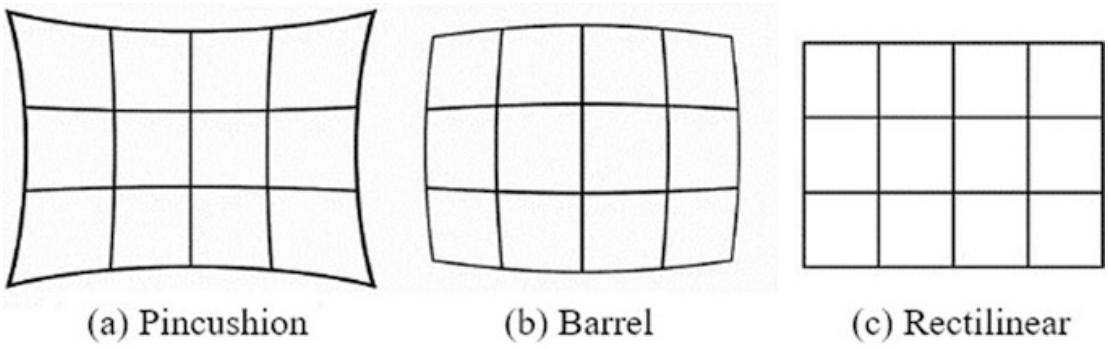


Fig. 53.3 Types of camera distortion

$$I = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B \tag{53.1}$$

Image filtering. Some of the captured images have inherited noise, namely Gaussian noise and salt and pepper noise. These noises are removed by Gaussian filter and median filter, respectively. Gaussian filtering is achieved by convolving the 2-D Gaussian distribution function with the image matrix.

$$G(x, y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \tag{53.2}$$

where G is the output of Gaussian operation on the image; x and y are the directions in which the function is operating. The standard deviation σ is the most important of all; it controls the amount of blurring, and its higher value causes more blurring and might lead to the loss or distortion of image features. A median filter is more effective than convolution as it reduces the noise and preserves the edges. Use of smoothening filters causes blurred edges. For proper detection of features and edges, unsharp masking (subtraction of a blurred version of the image from itself) sharpens the images. Image taken has poor contrast between the region of interest (ROI) and background; to tackle this problem 'imadjust,' a MATLAB™ function is employed. It uses the histogram stretching/shrinking to increase or decrease the contrast.

Binarization. The operations performed on the image are combinations of addition, subtraction, multiplication, and matrix inversion. It is convenient to perform these operations on matrix ones (1s) and zeros (0s). Binarization is the conversion of pixel intensities to ones and zeros. Pixels above the threshold are assigned ones and those below the threshold are assigned zeros. Threshold calculated using Otsu's method is more suitable for separation of classes in gray level [11]. In the current work, the effectiveness of the threshold varied from 0.9348 to one during the experiments. If the binarization assigns ones to the background, then inversion (reversing assigned values to pixels) of the image is required to increase the efficiency of computation.

53.2.3 Wear Computation

Noise Removal. Noise occurring in the images is of two types, namely disconnected noise and connected noise. Figure 53.4 shows the various types of noise that can be associated with a camera image. Disconnected noise is present in the image but stays at a distance apart from ROI. Connected noise is attached to ROI or in the vicinity of ROI. In addition, due to their random nature, connected noise cannot be separated using predefined steps. Connected noise occurs due to the shade of the tool as well. The width of such noise happens to be constant if the image acquisition setup does not change, and thus, making it possible to clear them. When the image is binarized, all the disconnected noises are present; there is no difference in the noise blocks and tool in terms of pixels. In the present algorithm, the binarized image was labeled, i.e., provided pixel numbers to each object in the image. Figure 53.5 shows typical binary image matrices. The camera was placed in such a manner that tool gets the maximum area in captured images. This helps in detecting the objects related to cut-chips, as the size of disconnected noise can be in the form of chips, which will not be larger than the tool itself. Then, the identity of the largest object is detected, and the rest of the objects in the image are discarded (Fig. 53.5).

Edge Detection. Orientation correction of the image is resource-consuming operation; it works faster with the edge image of the tool. The Canny operator is more suitable to find the edges in the image by using the local maxima of the image gradient. This method uses two thresholds to detect strong edges (high magnitude

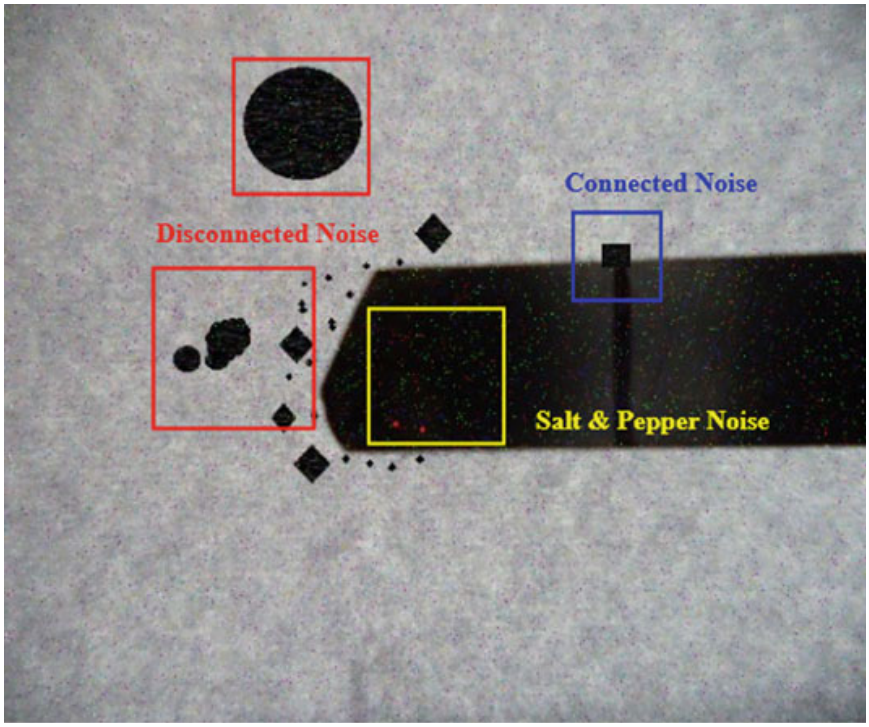


Fig. 53.4 Types of noise, disconnected noise, connected noise, and salt and pepper noise

of gradient) and weak edges (low magnitude of gradient) while preserving the weak edges connected to strong edges. However, the use of two thresholds makes the Canny method less susceptible to noise and helps in faster detection of true weak edges.

Use of automatic anisotropic Gaussian filters has shown improved results for edge detection process which removes the inherent deficiency of missing edges due to nonmaxima suppression associated with the Canny's algorithm [13].

Orientation Correction. The tool images were captured after removing the cutting tool from the lathe machine; and therefore, they are vulnerable to a slight deviation in tool orientation in the horizontal plane. In addition, the amount of deviation is different for any two images. This leads to false area detection. The orientation of images is corrected using Radon transform and is given by

$$\acute{f} = Rf = \int_L f(x, y)\mathrm{d}s \tag{53.3}$$

$$x = p\,\cos\phi - s\,\sin\phi \tag{53.4}$$

$$y = p\,\sin\phi + s\,\cos\phi \tag{53.5}$$

where ds is the increment in length along L (Eq. 53.3), With reference to the image, the Radon transform is the projection of the image intensity along a radial line oriented at a specific angle, ϕ; and ϕ varies from $\pi/2$ to $\pi/2$. If ϕ is a scalar, $\acute{f}$ is a column vector containing the Radon transform for theta degrees. The angle ϕ corresponding to the maximum value of Radon transform is the orientation angle β for the object in the image. In addition, the accuracy of orientation angle can be tuned to suit the need.

Geometrical Transformations. The cutting tool shown in Fig. 53.4 can have four possible orientations, horizontal with tool base on left or right, vertical with base on top or bottom. By using the geometrical properties, the tip, and the base of the tool are identified. After obtaining the orientation angle of the tool in an image, the image is rotated in a manner (i.e., using rotation angle as $\pm\beta$ or $90° \pm \beta$ or $180° \pm \beta$) that every tool image has the same orientation. Rotation of the image matrix in MATLAB^{TM} increases its size. It is due to image padding (addition of ones and zeros based on user's choice) action of rotation function. In addition, an increase in the size of an image is different for different angles of rotation. To subtract one image from another, the sizes of two images must be same. Therefore, the image cropping accompanied translation is necessary. Figure 53.6 shows the captured images of a typical cutting tool and their corresponding binary images. The centroid of binary image of the worn tool is aligned to the centroid of healthy tool. Then, the subtraction operation is performed to obtain the worn-out area, as shown in Fig. 53.7.

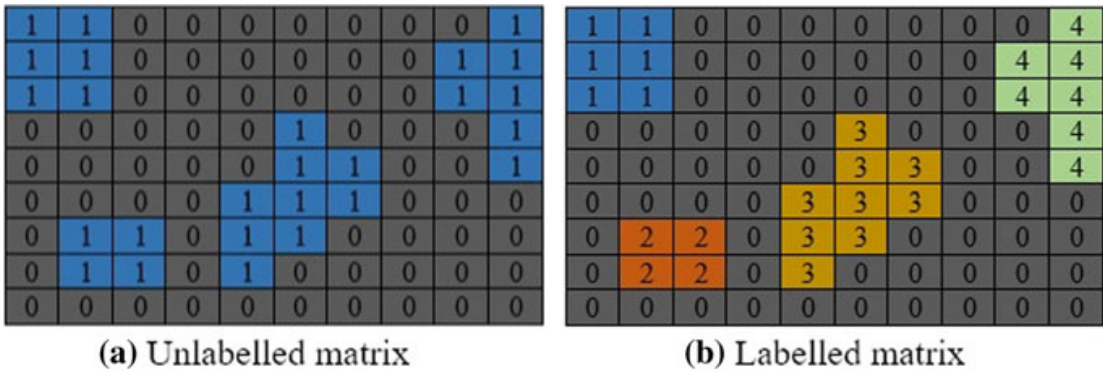


Fig. 53.5 Binary image matrices

Numerical Estimation. Tool wear data is computed in terms of pixels, and size of pixel is a function of camera's intrinsic properties, standoff distance between camera and objects. Therefore, this data is converted to world units, millimeters. It may happen that worn tool image and healthy tool image are assigned different number of pixels for same actual width. This discrepancy is taken care by resizing both the images such that 700 pixels are assigned to the width of the tool in each image; the number of pixels assigned to tool width for the experimental setup varied around 700.

53.3 Experimental Setup

To validate the performance of the developed algorithm, in-house experiments on plain turning operation were carried out on cylindrical workpiece made up of H13 steel, measuring 25 mm in diameter and 110 mm length. The experiments were performed on 11 kW, high-speed precision lathe, HMT NH26. At start, an image of the cutting tool was captured just after it was ground and the second image was captured after its failure. Machining parameters used for this experiment are listed in Table 53.1. The image for cutting tool was taken in the lab in presence of proper illumination source and away from machining noises like chips, dirt, cutting fluid,

Table 53.1 Machining parameters

Sr. no.	Parameter	Value
1	Tool material	High-speed steel (HSS) (*Miranda Tools make*)
2	Tool dimension	1/4″ × 4″
3	Work material	H13 steel (cylindrical)
4	Spindle speed	598 rpm
5	Feed	0.058 mm/rev
6	Depth of cut	0.5 mm
7	Machining time	2.1 min (stopped due to tool failure)

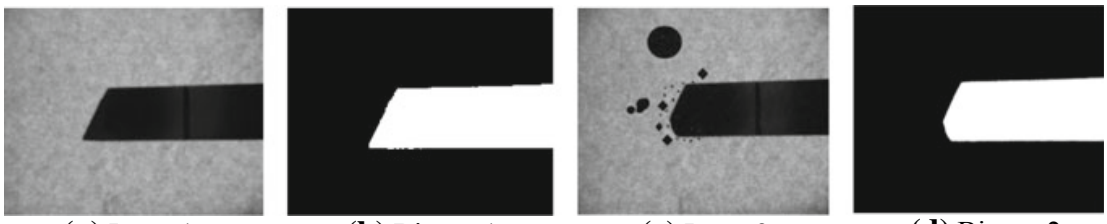
etc. The details regarding the image acquisition setup are already presented in Sect. 53.2.1. A schematic representation of image acquisition setup is shown in Fig. 53.2.

53.4 Results and Discussion

53.4.1 Image processing

Figure 53.6a, c shows the input true-color images and Fig. 53.6b, d shows the noise removed in corresponding binary images with scale factor one. Tables 53.2 and 53.3 show the properties of tools in corresponding binary images. Table 53.4 shows the numerical value of wear area.

The number of pixels assigned for tools in the image is 687 for the healthy tool and 731 for the worn tool image. Both the images were resized to bring the tool width to 700 pixels. Table 53.2 shows the properties of both the images.



(a) Input 1 **(b)** Binary 1 **(c)** Input 2 **(d)** Binary 2

Fig. 53.6 Captured images of cutting tool and their corresponding binary images

Table 53.2 Properties of binary images of tools

Properties	Values before machining	Values after machining
Area	2,286,601	2,271,953
Perimeter	72,483.3	70,788.2
Equivalent diameter	1706.3	1700.8
Centroid (x, y)	2851.1, 1609.6	2869.8, 1551.5
Euler number	1	1
Initial orientation	−0.5°	−0.3°

Table 53.3 Properties of the binary image of worn-out area in pixels

Sr. no.	Properties	Values
1	Area	24,739
2	Perimeter	815.57
3	Equivalent diameter	177.48
4	Centroid (x, y)	1464.5, 1963.8

Table 53.4 Results: worn-out area

Properties	Values in pixels	Values in world units
Area	24,739	2.03 mm^2
Perimeter	815.57	7.39 mm
Equivalent diameter	177.48	1.41 mm

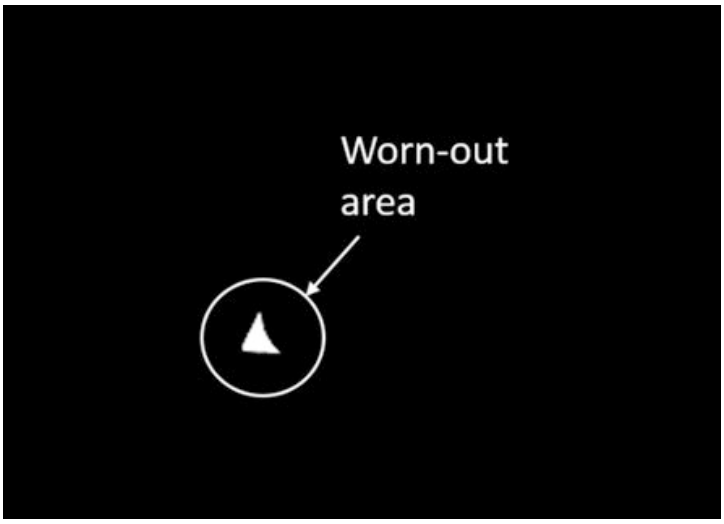


Fig. 53.7 Worn-out area image

Area mentioned in Table 53.2 is in terms of number of pixels. It can clearly be noted that there is reduction in the area of captured image after the machining run. Equivalent diameters were computed by treating the area to be equal to the area of a circle. A difference in centroid coordinates of both images occurs because of two reasons; first, actual position of tool while image capturing and second, due to wear of a portion of the tool.

Figure 53.7 shows the worn-out area image. The original worn-out area image contains some disconnected noise. This disconnected noise occurs, as the alignment of the image tools is not exactly 100%, but very close to it. The size of noise is very small and can be removed by using disconnected noise removal method as mentioned in Sect. 53.2.3.

The width of the tool is 6.35 mm and 700 in terms of pixels. Based on this, the conversion scale was computed as 0.0091 mm/pixel. Since the pixels are square of the data obtained using the current experimental setup, it was assumed that area conversion scale is square of conversion scale. It is equal to 8.23×10^{-5} mm^2/ area in terms of pixels. Table 53.4 shows the worn-out area in terms of pixels and in world units.

In this work, image processing toolbox available in MATLAB™ is used. It has a plethora of predefined functions, some of which can be customized to suit the need of the work. In addition, MATLAB™ provides a seamless integration of its various predeveloped modules which found to be very helpful during the development of algorithm of the present work.

Fig. 53.8 Variation of time and percentage error with different scale factors

53.4.2 Speed Test Results

During the algorithm development, it was noticed that a trial on a large size image, i.e., 4608 × 3456 (original image) consumed a considerable amount of time, 87 s; therefore, a speed test was conducted to test the response of the developed algorithm. In this test, the dimensions of the original image were scaled and the effect of reduction in size of the image was studied. The reduction in size of the image was performed manually using Windows Paint application. The time for single run of the developed algorithm was collected from MATLAB™ profiler. First, the scale factor (by what factor the image size has been reduced) was varied from 1 to 2.304. Then, the errors in wear areas were determined. It was noted that by using the scaled images, the speed and accuracy of the computation were improved. Figure 53.8 shows that a scale factor (reduction) of two performed well and can be used for the further computations.

Overall, this work demonstrates a simple and easy method to compute the tool wear, in which only a simple digital camera is sufficient for computation of tool wear. Camera can be mounted in CNC machines, which captures the images in intervals when tool is not interacting with workpiece. By computing the end-cutting-edge wear, dimensional accuracy of workpiece could easily be improved by applying in-process tool compensation. The present work may also help in monitoring the in-process tool wear.

53.5 Conclusion

In the present work, a simple and efficient image processing algorithm has been developed for computation of wear in cutting edges of single-point cutting tool used in plain turning operation. An in-house image acquisition setup has been developed to capture the images of cutting tools (before and after machining). The captured images were preprocessed using various advanced image processing functions such as distortion correction, conversion from 3D to 2D, filtering using Gaussian and

median filters, and binarization. Canny operator was used to properly orient the captured images. The orientation of the images was corrected using Radon transform. Finally, the worn-out area was computed in terms of pixels by using the geometric transformations of captured images. The size of pixel is a function of camera's intrinsic properties, standoff distance between camera and objects. The performance of the developed algorithm has been validated by conducting in-house experiments on plain turning operation, and it was found satisfactory. During the speed test, it was noticed that processing of the original image (4608 × 3456) consumed a considerable amount of time, 87 s. However, the scaled image with reduction factor of two improved the speed of computation as well as accuracy.

The work presented hereby is useful for tool monitoring process and for tool change decision-making process. The presented algorithm has potential and flexibility of being used with fully automatic tool changing process. The whole process involves use of a simple digital camera, which reduces the cost of the operation. This work would be useful for all tool rooms, machine tool manufacturers, and small-scale industry. However, reduction in computation time would further enhance the potential of this work. Integration of artificial neural networks (ANN) would reduce the frequency of image acquisition required for the process.

References

1. Cho, D.W., Jung, M.Y.: On-line monitoring of tool breakage in face milling using a self-organized neural network. J. Manuf. Syst. **14**(2), 80–90 (1995)
2. Astakhov, V.P.: Tribology of Metal Cutting. Elsevier Ltd., Amsterdam (2006)
3. Bahr, B., Motavalli, S., Arfi, T.: Sensor fusion for monitoring machine tool condition. Int. J. Comput. Integr. Manuf. **10**, 314–323 (1997)
4. Balasmo, V., Caggiano, A., Jemielniak, K.: Multi sensor signal processing for catastrophic tool failure detection in turning. Procedia CIRP **41**, 939–944 (2016)
5. Brooks, R.R., Iyengar, S.S.: Multi Sensor Fusion: Fundamentals and Applications with Software. Prentice Hall PTR (1998)
6. Chirssolouris, G., Domroese, M., Beaulieu, P.: Sensor synthesis for control of manufacturing processes. J. Eng. Ind. Trans. ASME **114**, 158–174 (1992)
7. Dimla, D.E., Lister, P.M., Leighton, N.J.: A multi sensor integration method of signals in a metal cutting operation via application of multi-layer perceptron neural networks. In: 5th International Conference on Artificial Neural Networks, pp. 306–311. Cambridge (1997)
8. Dutta, S., Pal, S.K., Mukhopadhyay, S., Sen, R.: Application of digital image processing in tool condition monitoring: a review. CIRP J. Manufact. Sci. Technol. **6**(3), 212–232 (2013)
9. Kurada, S., Bradley, C.: A machine vision system for tool wear assessment. Tribol. Int. **30**(4), 295–304 (1997)
10. Fadare, D.A., Oni, A.O.: Development and application of a machine vision system for measurement of tool wear. ARPN J. Eng. Appl. Sci. **4**(4) (2009)
11. Shahabi, H.H., Ratnam, M.M.: Assessment of flank wear and nose radius wear from workpiece roughness profile in turning operation using machine vision. Int. J. Adv. Manuf. Technol. **43**, 11–21 (2009)

12. Tan, K.K., Er, P.V., Yang, R., Teo, C.S.: Selective precision motion control using weighted sensor fusion approach. In: IEEE International Conference on Mechatronics and Automation, IEEE ICMA 2013, pp. 179–184 (2013)
13. Zhang, W., Zhao, Y., Breckon, T.P., Chen, L.: Noise robust image edge detection based upon the automatic anisotropic Gaussian kernels. Pattern Recogn. **63**, 193–205 (2017)

Chapter 54
Modelling and Analysis of Multi-agent Approach for an IoT-Enabled Autonomous Manufacturing System

K. B. Badri Narayanan and M. Sreekumar

Abstract Industrial Internet of Things (IIoT) is a digital platform to perform machine-to-machine communication, acquire sensor data, analyse, and deliver comprehensive information and also facilitate the manufacturing system to implement smart decisions faster. This paper presents modelling and analysis of an IoT-enabled autonomous manufacturing system based on the performance metrics. Multi-agent interaction has been defined through a generalised control architecture, considering as an autonomous manufacturing system. The number of jobs entering and leaving the machines computes time-dependent performance metrics such as cycle time and work in process using Little's law. The mean arrival rate of each machine is calculated by evaluating the amount of work and a utilisation factor. The squared coefficient of inter-arrival rates based on successive substitution of job flow from machine x to machine y has also been obtained. This is preliminary work in the process of designing an IoT-enabled autonomous manufacturing system capable of self-learning and executing accurate decision faster.

Keywords Internet of Things · Cyber-physical systems · Factory modelling and simulation · Autonomous manufacturing systems

54.1 Introduction

According to Moore's law, the exponential growth of technology is forcing constant up-gradation in manufacturing the environment. Current manufacturing scenario lacks in adopting the changes, so an intelligent manufacturing system with a decentralised multi-agent approach shall be of a solution. In manufacturing systems, downtime can happen due to various disorders such as tool failure, workstation

K. B. Badri Narayanan · M. Sreekumar (✉)
Centre for AI, IoT, and Robotics, Department of Mechanical Engineering, Indian Institute of Information Technology, Design and Manufacturing, Kancheepuram, Vandalur–Kelambakkam Road, Chennai 600127, India
e-mail: msk@iiitdm.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_54

failure, the breakdown of material handling devices like a robot, conveyor, and transfers system. This downtime issue generally increases the lead time and production cost. Immediate rescheduling of the process flow of the respective workpiece offers a solution for reducing systems downtime wisely and efficiently [1]. Also, reconfiguration changes the structure of the manufacturing layout that requires new scheduling [2–4]. Even state-of-the-art modern manufacturing facilities cannot be able to provide a complete solution for machine downtime.

For agent breakdown cases, present automation systems use a defined protocol. Hence for defining a new protocol, the system should be interrupted in the middle. This aspect increases the rigidness of the system. This can be avoided by using the machine-to-machine communication protocol called Process Field Net (PROFINET). Similarly, a cognitive approach can be implemented to adopt rapid changes in the manufacturing systems environment. Several unique paradigms of Intelligent Manufacturing Systems (IMS) such as holonomic systems [5, 6], cognitive systems [7], biological systems [8, 9], and reconfigurable manufacturing systems [10, 11] have already been established.

A manufacturing Petri Net (MPN)-based scheduling model for IoT-enabled hybrid flow shop manufacturing is discussed in [12]. IoT in the manufacturing system digitises the real-time production system and provides a faster smart decision. In the present work, an IoT enabled autonomous manufacturing system, as shown in Fig. 54.1, is designed by considering every machine as agents. Jobs are considered as encapsulated with radio-frequency identification (RFID) tags and represented as stable, smart objects (SSO). Every machine is attached with RFID readers and represented as dynamic smart objects (DSO). Similarly, every worker is

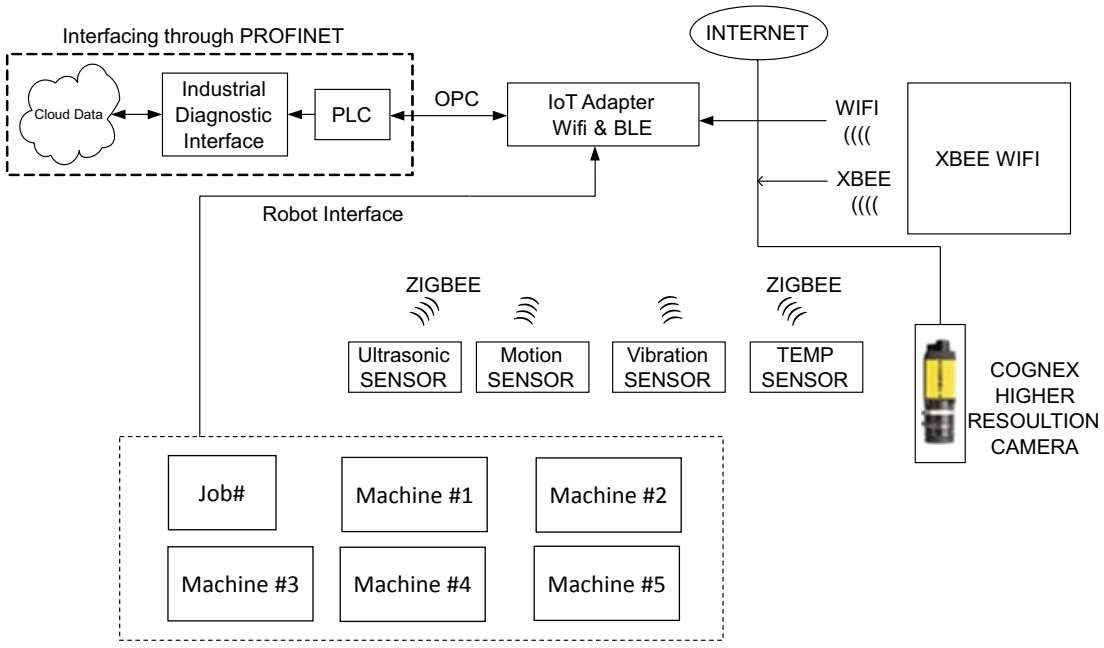


Fig. 54.1 Schematic arrangement of wireless communication in an IoT-enabled manufacturing system

provided with RFID smart ID card. Machining and buffering are recorded in the database by interactions between SSO, DSO, and smart ID card.

The rest of this article is constructed as follows. Section 54.2 discusses the core technologies such as swarm intelligence and control architecture of autonomous manufacturing systems. Section 54.3 is focused on modelling and analysis of the performance metrics of manufacturing systems. The conclusion is drawn in Sect. 54.4, followed by references.

54.2 Core Technologies

54.2.1 Swarm Intelligence

Swarm intelligence is a bio-inspired technique that mimics the localised interaction behaviour that exists amongst various species such as ant, insects, and other animals. Ant colony optimisation (ACO) is a swarm technique which mimics the food searching behaviour of ants. The interaction between ants is carried out basically by deposition of pheromone by a preceding ant. ACO can be implemented in manufacturing systems by considering machines as a group of autonomous and cooperative agents like ants. Such systems could be self-organised by a local interaction between SSO, DSO, and smart ID card. Every machine can calculate its pheromone value for a particular process. During machine failure, the agent interaction can be defined by comparing the machine specification and process information that are available in the database. The formulation of pheromone value can be accomplished by

$$P_X = \left[\alpha_{\mathrm{WIP}}(\mathrm{WIP}_O/\mathrm{WIP}) + \alpha_{\mathrm{CT}}(\mathrm{CT}_O/\mathrm{CT}) + \alpha_S\left(\frac{C_{SO}^2}{C_S^2}\right) + \alpha_I\left(\frac{C_{AO}^2}{C_A^2}\right) + \alpha_A\left(\frac{\delta_{xo}}{\delta_x}\right)\right]^{-1} \tag{54.1}$$

where P_X—pheromone value, WIP_O—maximum work in process, CT_O—maximum cycle time, C_{SO}^2 represents the maximum squared coefficient of variation for service time, C_{AO}^2—maximum squared coefficient of variation for the inter-arrival rate of jobs, δ_{xo}—maximum arrival rate of jobs to machine x. The factors $\alpha_{\mathrm{WIP}}, \alpha_{\mathrm{CT}}, \alpha_S, \alpha_I$ and α_A represent the weight of WIP, CT, C_{SO}^2, C_{AO}^2 and δ_{xo}, respectively. After generating the pheromone values of each process for all machines, the machine failures can be identified with the following cases

Case 1: If $\forall P_X < 0.2$
Then send a message from MES to Employee staff RFID /*all machines are good*/

Case 2: If $\exists! P_X \geq 0.2$
Then send a message from MES to Employee staff RFID /*the machine x is selected for rectifying the failure*/
Case 3: If there are more than one $P_X \geq 0.2$
Then send a message from MES to Employee staff RFID /*the machine which is having the highest pheromone value selected for rectifying the first and followed by the subsequent highest value*/
Case 4: If $\forall P_X \geq 0.2$
Then

(i) send a message from MES to staff RFID /*All machines had pheromone value above 0.2*/
(ii) In response to the above message, send a message from staff to MES /*requesting the MES for rescheduling*/.

54.2.2 Control Architecture

The generalised control architecture of autonomous manufacturing systems is shown in Fig. 54.2. The core elements of the control architecture are agent interaction, agent behaviour, and database. The interaction between agents and manufacturing execution system (MES) occurs through Extensible Markup Language (XML). XML provides readable formats to both humans and machines. The industrial control system protocol Open Platform Communication (OPC) interfaces the machine with Programmable Logic Controller (PLC) and connects the physical devices like sensors integrated with machining centres. The database stores the information like job address, machine address, process details about jobs, sensor data, and tool failures. The agent cooperation is incorporated by calculating the pheromone value of each agent. Then the agents compare the pheromone value by interacting with nearest agents and finally, the agent which has the highest pheromone value is chosen to rectify the failure.

54.3 Performance Metrics

The mean of performance metrics considered for modelling and analysis of an IoT-enabled autonomous manufacturing system is cycle time CT, work in process WIP, arrival rate, squared coefficient of variation for inter-arrival, and service time. Using Little's law, the probability of seven different jobs flow across the manufacturing systems is formulated by equating the number of jobs entering and leaving the machine. This also computes the time-dependent variables such as cycle time and work in process. The mean arrival rate of each machine is calculated by

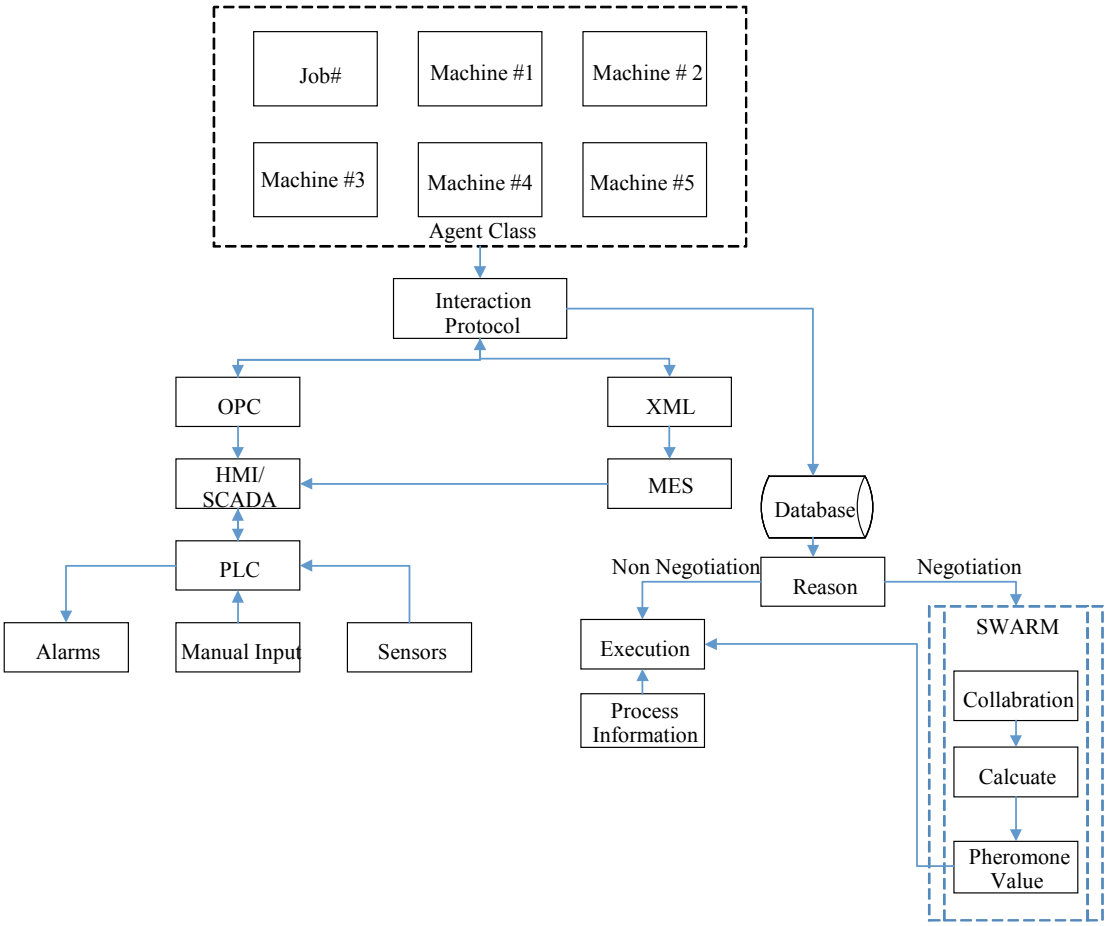


Fig. 54.2 A generalised control architecture of an autonomous manufacturing system

evaluating the amount of work and a utilisation factor. The squared coefficient of inter-arrival rates is obtained by successive substitution of job flow from machine x to machine y.

54.3.1 Workflow Modelling

A discrete workflow network, defining the arrival rate, is more complicated than the continuous systems. Consider machines, as shown in Fig. 54.3, the job is entering from an external source into the first machine with a mean rate of o. However, machine 1 also got rejected job with probability d. Hence, the flow into the machine 1 is not clear. Sequentially, the probability for machine 2 is also not clear since the job arrives from machine 1 has a mean rate of p_1. This quandary is a regular due to the discrete movement of jobs. These situations demand to calculate flow rates concurrently. For example, δ_x for $x = 1, 2 \ldots n$ is used to indicate the total arrival of jobs for machine x. The mean flow rates are modelled mathematically using steady-state conditions as

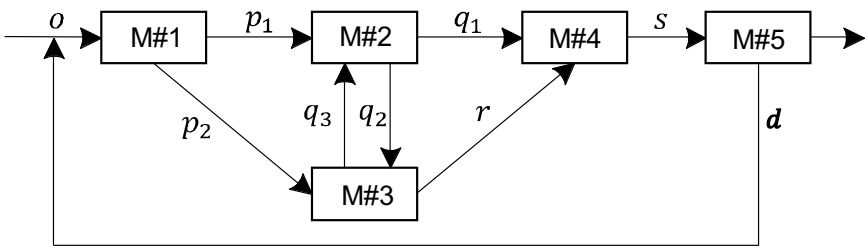


Fig. 54.3 Autonomous manufacturing system topology

$$\left.\begin{aligned}\delta_1 &= o + d\delta_5\\ \delta_2 &= p_1\delta_1\\ \delta_3 &= p_2\delta_1 + q_1\delta_2\\ \delta_4 &= \delta_3 + q_2\delta_2\\ \delta_5 &= s\delta_4\end{aligned}\right\} \tag{54.2}$$

The steps involved in the evaluation of autonomous manufacturing systems are detailed below.

Step 1. Compute the amount of work for each machine by considering n machines in a manufacturing system with i type jobs and each type flows is defined by transmitting matrix R_i and by assuming that summation of any row should be less than one. Let o_i denote mean arrival rate of type i jobs, and now the vector δ_i is given by [13]

$$\delta_i = \left[I - (R_i)^T\right]^{-1} o_i \tag{54.3}$$

Amount of work for each machine is calculated by

$$\mathrm{AW}_x = \sum_{i=1}^{m} \delta_{i,x} E[T_s(i,x)] \tag{54.4}$$

where m—maximum number of products,
$\delta_{i,x}$—the rate of product type i entering into machine x as shown in Table 54.1 and $E[T_s(i,x)]$ is mean service time of type i job at machine x.

The utilisation factor u_x, for machine x, is,

$$u_x = \frac{\mathrm{AW}_x}{c_x} = \frac{\sum_{i=1}^{m} E[T_s(i,x)]}{c_x} \tag{54.5}$$

Table 54.1 Arrival rate of each machine

Arrival rate of each machine x	Job types						
	J# 1	J# 2	J# 3	J# 4	J# 5	J# 6	J# 7
$\delta 1$	5.0813	5.4585	5.102	5.4348	5.102	5.0201	5.004
$\delta 2$	4.065	4.3668	0	0	0	0.2008	0.2002
$\delta 3$	0	3.7118	1.0204	0	1.0204	1.004	1.0008
$\delta 4$	0.813	4.5852	1.0204	4.3478	0	0	0.04
$\delta 5$	0.813	4.5852	1.0204	4.3478	1.0204	0.2008	0.04

Step 2. The meantime, squared coefficient of variation for the service time, and squared coefficient of variation for the arrival times at machine x are calculated by

$$E[(T_s(x)] = \sum_{i=1}^{m} \frac{\delta_{ix}}{\delta_x} E[T_s(i,x)] = \frac{\mathrm{AW}_x}{\delta_x} \tag{54.6}$$

$$C_s^2(x) = \frac{\left\{\sum_{i=1}^{m} \frac{\delta_{i,x}}{\delta_x} E[T_s(i,x)]^2 [1 + C_s^2(i,x)]\right\}}{\left\{\sum_{i=1}^{m} \left(\frac{\delta_{i,x}}{\delta_x}\right) E[T_s(i,x)]\right\}^2} - 1 \tag{54.7}$$

Step 3. The combined transmitting matrix $R = (r_{x,y})$ gives the transferring possibilities of a subjective job and is determined by

$$r_{x,y} = \frac{\sum_{i=1}^{m} \delta_{ix} r_{x,y}^{i}}{\delta_x} \quad \text{for } x, y = 1, \ldots, n \tag{54.8}$$

The flow of jobs from machine x that are routed directly to machine y will be called as x to y stream and the squared coefficient of variation of inter-arrival times to y from x is obtained by

$$C_a^2(y) = \frac{o_y}{\delta_y} C_a^2(0,y) + \sum_{x=1}^{n} \frac{\delta_x r_{x,y}}{\delta_y} \begin{bmatrix} r_{x,y}(1 - u_x^2) C_a^2(x) \\ + r_{x,y} u_x^2 \left(\frac{C_s^2(x) + \sqrt{c_x} - 1}{\sqrt{c_x}}\right) + 1 - r_{x,y} \end{bmatrix} \tag{54.9}$$

Then, the mean cycle time spent within the manufacturing system by a type i job is given by

$$CT^i = \frac{\sum_{i=1}^{m} \delta_{i,x}\{CT_q(x) + E[T_s(i,x)]\}}{\sum_{y=1}^{n} o_{iy}} \tag{54.10}$$

54.3.2 Performance Analytics

The flow sequence of each product is shown in Table 54.2, and the processing time is shown in Table 54.3. The flow rate of each machine is assumed as [13] $o = 5$, $p_1 = 0.8$, $p_2 = 0.2$, $q_1 = 0.6$, $q_2 = 0.2$, $q_3 = 0.2$, $r = 1$, $s = 1$ and $d = 0.1$. The amount of work done is calculated by Eq. (54.4). The meantime and squared coefficient of variation for the service time at machine x can be calculated using the Eqs. (54.6) and (54.7).

Table 54.2 Sequence of each product type

Job sequence J	Job Types						
	J#1	J#2	J#3	J#4	J#5	J#6	J#7
Sequence 1	M# 1	M# 1	M# 1	M# 1	M# 1	M# 1	M# 1
Sequence 2	M# 2	M# 2	M# 3	M# 4	M# 3	M# 3	M# 3
Sequence 3	M# 4	M# 3	M# 4	M# 5	M# 5	M# 2	M# 2
Sequence 4	M# 5	M# 4	M# 5			M# 5	M# 4
Sequence 5		M# 5					M# 5

Table 54.3 Processing data of each machine for seven products

Machine x		M# 1	M# 2	M# 3	M# 4	M# 5
J# 1	$E[T_s(1,x)]$	0.071429	0.1	–	0.066667	0.06
	$C_s^2(1,x)$	0.8	1.2	–	1.5	0.75
J# 2	$E[T_s(2,x)]$	0.066667	0.055556	0.083333	0.06	0.0625
	$C_s^2(2,x)$	1.33	2	1.5	0.75	1.7
J# 3	$E[T_s(3,x)]$	0.065574	–	0.081633	0.070175	0.061538
	$C_s^2(3,x)$	1.43	–	1.56	1.5	1.72
J# 4	$E[T_s(4,x)]$	0.064725	–	–	0.069204	0.06079
	$C_s^2(4,x)$	1.53	–	–	1.6	1.8
J# 5	$E[T_s(5,x)]$	0.068259	–	0.079051	–	0.06006
	$C_s^2(5,x)$	1.527047	–	1.425532	–	1.76555
J# 6	$E[T_s(6,x)]$	0.06734	0.05305	0.077821	–	0.059347
	$C_s^2(6,x)$	1.526055	2.45	1.425532	–	1.751174
J# 7	$E[T_s(7,x)]$	0.066445	0.052493	0.076628	0.066445	0.058651
	$C_s^2(7,x)$	1.525533	2.65	1.377358	1.487805	1.737327
c_x		19	4	4	4	6

The multiple job study problems now shrink to a single job study. The transferring possibilities of a subjective job and is determined by Eq. (54.8). The mean time spent with the manufacturing system by a type i job is calculated using Eq. (54.10). The shop floor parameters like cycle time CT, work in process WIP, arrival rate, squared coefficient of variation for inter-arrival, and service time are shown in Fig. 54.4.

In Fig. 54.4, the performance metrics of each machine is monitored remotely. The database provides dynamics of real-time data for intelligent decision-making. Pheromone value of each machine is evaluated based on real-time data of performance metrics. Each machine has a different threshold value for different performance metrics. When the performance metrics getting deviated beyond the threshold value, messages are sent to employee's smart ID to overcome the deviation. The significance of this paper provides a solution to overcome the downtime of each machine using the cognitive ability of agents. Swarm intelligence improves the

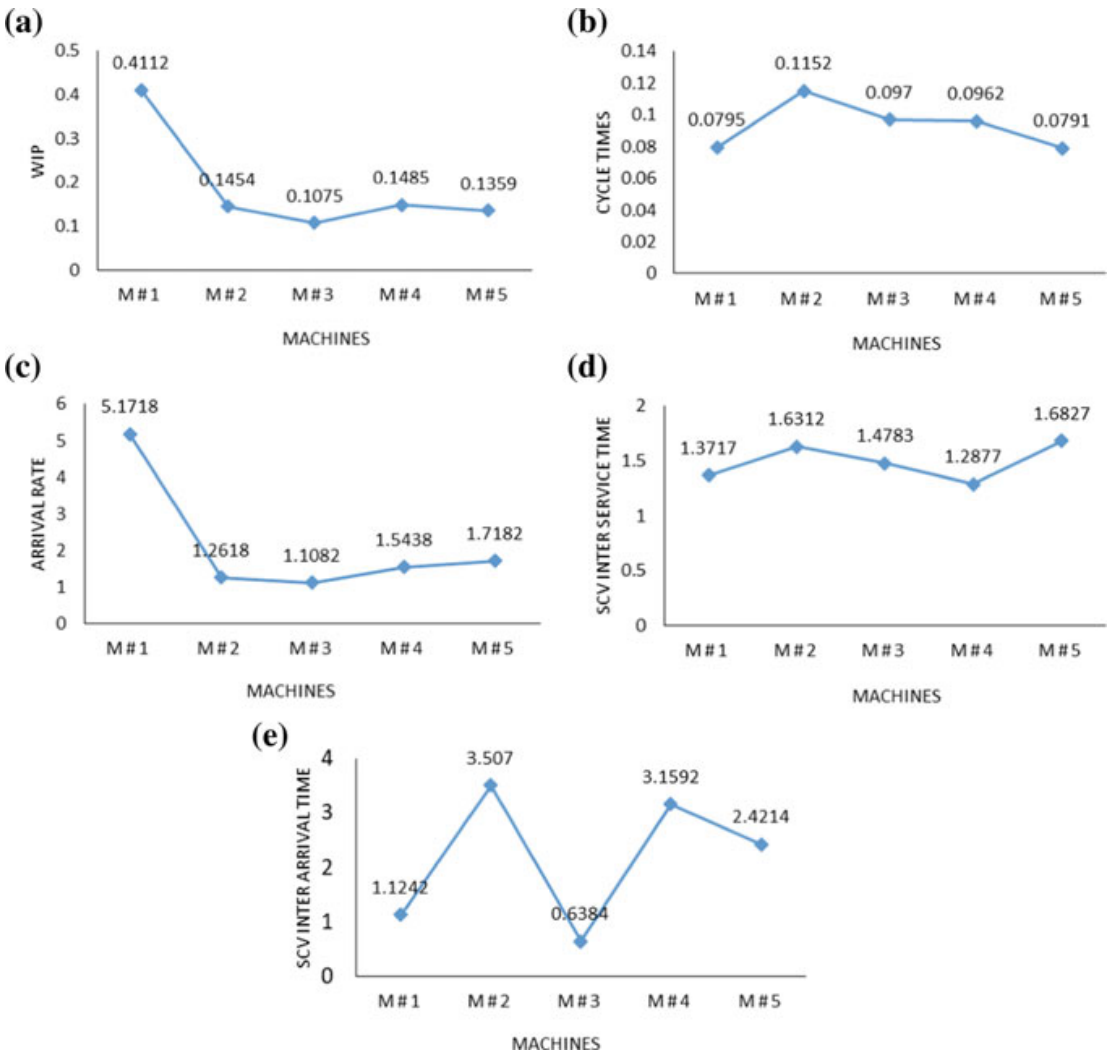


Fig. 54.4 Performance parameters **a** cycle time, **b** work in process, **c** arrival rate, **d** SCV inter-arrival time, and **e** SCV service time

autonomous behaviour of each agent to defend the disturbances without any upper-level aid and enhances the reliability of the system. The dashboard is transparently showing the dynamic status of machine health and the solution based on pheromone value priorities and addresses the failure immediately.

54.4 Conclusion

In this work, an attempt has been made to derive mathematical model and analyse a manufacturing system integrated with IoT based on the mean of performance metrics such as cycle time, work in process, arrival rate, squared coefficient of variation for service time, and inter-arrival rate. This model reflects the real-time operations of machining and buffering in a smart manufacturing system. The swarm intelligence of machining agents improves system robustness to face machine failures by multi-agent interaction, adaptable to changes, and reliable in product quality. The manufacturing system becomes rigid by decentralised control and autonomous behaviour in decision-making. The simulation results of all performance metrics provide dynamic states of each machine. The results also provide continual process adjustment in the manufacturing system to avoid failure of machines by evaluating the pheromone value. The manufacturing systems become intelligent to handle the disturbances in priority sequence using pheromone value and to reduce the downtime of the machine. Future research line extends to the development of the finite state of a machine for more precise modelling.

Acknowledgements Authors sincerely thank the Ministry of Electronics and Information Technology (MeitY), Government of India, for providing financial assistance under Visvesvaraya Ph.D. Scheme to carry out this research work at IIITDM Kancheepuram, Chennai, India.

References

1. Vieira, G.E., Hermann, J.W., Lin, E.: Rescheduling manufacturing systems: a framework of strategies, policies, and methods. J. Sched. 39–62 (2003)
2. Park, H.S., Choi, H.W.: Development of a modular structure-based changeable manufacturing system with high adaptability. Int. J. Precis. Eng. Manuf. 7–12 (2008)
3. Wang, Y.F., et al.: An integrated approach to reactive scheduling subject to machine breakdown. In: Proceedings of the IEEE International Conference on Automation and Logistics, pp. 542–547 (2008)
4. Shea, K., et al.: Design-to-fabrication automation for the cognitive machine shop. Adv. Eng. Inform. 251–268 (2010)
5. Leitão, P., Restivo, F.: ADACOR: a holonic architecture for agile and adaptive manufacturing control. Comput. Ind. 121–130 (2006)
6. Park, H.S., et al.: Agent-based shop control system under holonic manufacturing concept. In: Proceedings of the 4th Korea-Russia International Symposium, pp. 116–121 (2000)
7. Zäh, et al.: The cognitive factory. In: Changeable and Reconfigurable Manufacturing Systems, pp. 355–371. Springer, London (2009)

8. Bannat, A., et al.: Artificial cognition in production systems. IEEE Trans. Autom. Sci. Eng. 148–174 (2011)
9. Monostori, L.: AI and machine learning techniques for managing complexity, changes and uncertainties in manufacturing. Eng. Appl. Artif. Intell. 277–291 (2003)
10. Matsuda, M., Ishikawa, Y., Utsumi, S.: Configuration of machine tool agents for flexible manufacturing. In: 39th CIRP International Conference on Manufacturing Systems, pp. 351–357 (2006)
11. Wang, S., et al.: Towards smart factory for industry 4.0: a self-organized multi-agent system with big data based feedback and coordination. Comput. Netw. 158–168 (2016)
12. Wang, M., et al.: A MPN-based scheduling model for IoT-enabled hybrid flow shop manufacturing. Adv. Eng. Inform. 728–736 (2016)
13. Curry, G.L., Feldman, R.M.: Manufacturing Systems Modeling and Analytics. Springer Science & Business (2010)

Part VI
Material Fabrication

Chapter 55
Estimation of Distortion for CFRP Antenna Reflector During Autoclave Processing

Shrey Patel, Dhaval Shah, Shashikant Joshi and Kaushik Patel

Abstract Carbon fibre-reinforced polymer (CFRP) composites are broadly utilized in space applications due to their very high specific properties over typical metals and alloys. Autoclave curing is one of the finest composite producing techniques to achieve great dimensional accuracy of the advanced shaped component. The cure shrinkage, which is commonly referred to as warpage or spring-back distortion, is induced as a result of material thermo-elastic anisotropy during composite laminates manufacturing. In this paper, the distortion of CFRP reflectors is estimated using experimental and numerical approaches for autoclave processing. The antenna reflectors have been fabricated using HCU200/A45 prepreg and stainless steel AISI 430 as a part and tool material, respectively. The geometric modelling and finite element analysis (FEA) have been performed for parabolic reflectors using ABAQUS software along with COMPRO plug-in. Thermochemical and stress-deformation analysis has been performed to obtain the deformed shape of the antenna reflectors. The laminate thickness has been considered as a process parameter. The experimental and numerical results have been compared and located in sensible agreement with one another.

Keywords CFRP · Distortion · Finite element analysis · Autoclave

S. Patel · D. Shah (✉) · S. Joshi · K. Patel
Mechanical Engineering Department, Institute of Technology, Nirma University, Ahmedabad 382481, Gujarat, India
e-mail: dbshah@nirmauni.ac.in

S. Patel
e-mail: 17mmed05@nirmauni.ac.in

S. Joshi
e-mail: s.j.joshi@nirmauni.ac.in

K. Patel
e-mail: kaushik.patel@nirmauni.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_55

55.1 Introduction

The use of CFRP prepreg composites in aerospace, automotive, and defence industries is growing mainly due to their high strength and stiffness to weight ratios. Although there is a wide scope of application of these materials, still there is a limited use because of quality reliability issues, high raw material, manufacturing cost, and lengthy manufacturing cycle. Autoclave process is the most popular for getting high-quality composites components compared to other composite manufacturing processes. This manufacturing process gives the minimum void content in manufactured composite components as compared with other vacuum infusion process, resin transfer moulding process, filament winding process, and hand layup process. The antenna reflectors used for transmitting the radio frequency signals between two devices in space applications, and therefore, the profile of the reflectors should be accurate. The COMPRO, a special purpose subroutine, can be used for simulation of autoclave manufacturing process along with ABAQUS software. The COMPRO plug-in consists of the material database for prepreg, fibre and resin system, and tool materials. The thermochemical, flow-compaction, and stress-deformation analysis modules are available for the simulation of composites.

Many researchers worked on the prediction of warpage and spring-back deformations for composite components. During the process of curing, Hahn and Pagano [1] found out that some stresses which developed during that process varied with respect to temperature development. Darrow and Smith [2] established three process parameters, i.e., mould expansion, thickness cure shrinkage, and fibre volume fraction gradients that effects on spring-in deformations. Twigg et al. [3] predicted warpage of flat laminates with considering the part length, tool–part interaction, the effect of release film, and release agent as process parameters. Stefniak et al. [4] performed experimental work to quantify the various mechanisms, i.e., property gradient mechanism, stress gradient mechanisms with respect to their relative contribution to composite part distortions. Kaushik et al. [5] studied the result of various process parameters, i.e., pressure, ramp rate, and the degree of cure on the coefficient of friction in the autoclave process. Fernlund and Poursartip [6] demonstrated the effect of the cure cycle and tool surface condition on spring-back deformation. Zeng and Raghavan [7] implemented three-dimensional process model through user subroutine interfaced with finite element software. Abusafieh et al. [8] described the effect of temperature and degree of cure by using experimental setup and FEA. Dong [9] derived the mathematical formula to calculate spring-in of flat laminates using the effective coefficient of thermal expansion. Mezeix et al. [10] simulated spring-back deformation of unidirectional carbon/epoxy flat laminates manufactured through the autoclave process.

In this research paper, experimental and numerical approaches have been used to estimate distortion for CFRP antenna reflectors. The stainless steel AISI 430 and HCU200/A45 prepreg tape have been identified as a tool and part material, respectively. The laminate thickness, i.e., 0.8, 1.6, 2, and 2.4 mm has been considered as a process parameter. In the experimental approach, CFRP antenna

reflectors have been fabricated using the autoclave process. The geometrical modelling as well as FE analysis has been performed with the help of COMPRO plug-in using ABAQUS software. The thermochemical and stress-deformation analysis has been carried out for CFRP antenna reflectors to predict distortion. The comparison of distortions for fabricated reflectors using autoclave process and simulation has been carried out.

55.2 Experimental Approach

55.2.1 Material Selection

The AISI 430 grade of stainless steel has been selected as a tool material due to the ease of availability and machinability by comparing other grades of steel. The HCU200/A45 prepreg tape (tensile strength = 3450 MPa, fibre density = 1.8 g/cm^3) has been identified, which consists of excellent surface finish and mechanical properties. The consolidation materials, i.e., release agent, vacuum bag, peel ply, and breather cloth have been required in the manufacturing of composite parts using the autoclave process. These consolidation materials have been sustained temperature up to 200 °C. The laboratory-scale autoclave machine of overall diameter 900 mm and length of 1000 mm has been used for the manufacturing process.

55.2.2 Fabrication of Antenna Reflectors

The 300 × 300 × 70 mm size block of stainless steel AISI 430 material has been procured. The 3D CAD model has been prepared using solid modelling software for paraboloid profile with a circumferential diameter of 240 mm at one face with the depth of 50 mm. The machining of stainless steel tool has been performed using NC code in the 5-axis VMC machine with the great surface finish as shown in Fig. 55.1. The three coats of release agent have been applied on the parabolic tool surface. The numbers of prepreg layers as per the required thickness of reflector have been cut into the circular shape of 300 mm diameter. Peel ply, breather cloth, and vacuum bag have been kept on layers of prepreg. The sealant tape has been applied over the entire assembly to prevent leakage of air and resin from the assembly. The vacuum valve has been placed at the centre of a vacuum bag to maintain pressure. The one end of the vacuum hose has been connected to a valve whereas another end at the outlet of the vacuum pump. The entire setup has been kept in the autoclave for curing as shown in Fig. 55.2. The manufacturer recommended curing cycle has been given to the autoclave. The demoulding process has been performed at atmospheric condition after the completion of the curing process. The finish cured and demoulded reflector is shown in Fig. 55.3.

Fig. 55.1 Stainless steel AISI 430

Fig. 55.2 Autoclave machine

Fig. 55.3 Manufactured CFRP antenna reflector

55.3 Numerical Approach

The geometry of the tool and part has been created as per actual dimensions in ABAQUS software. The thermochemical and stress-deformation analysis has been performed with the help of COMPRO plug-in for antenna reflectors to predict distortion [11]. The methodology followed in the numerical approach has been described as a flow chart as shown in Fig. 55.4.

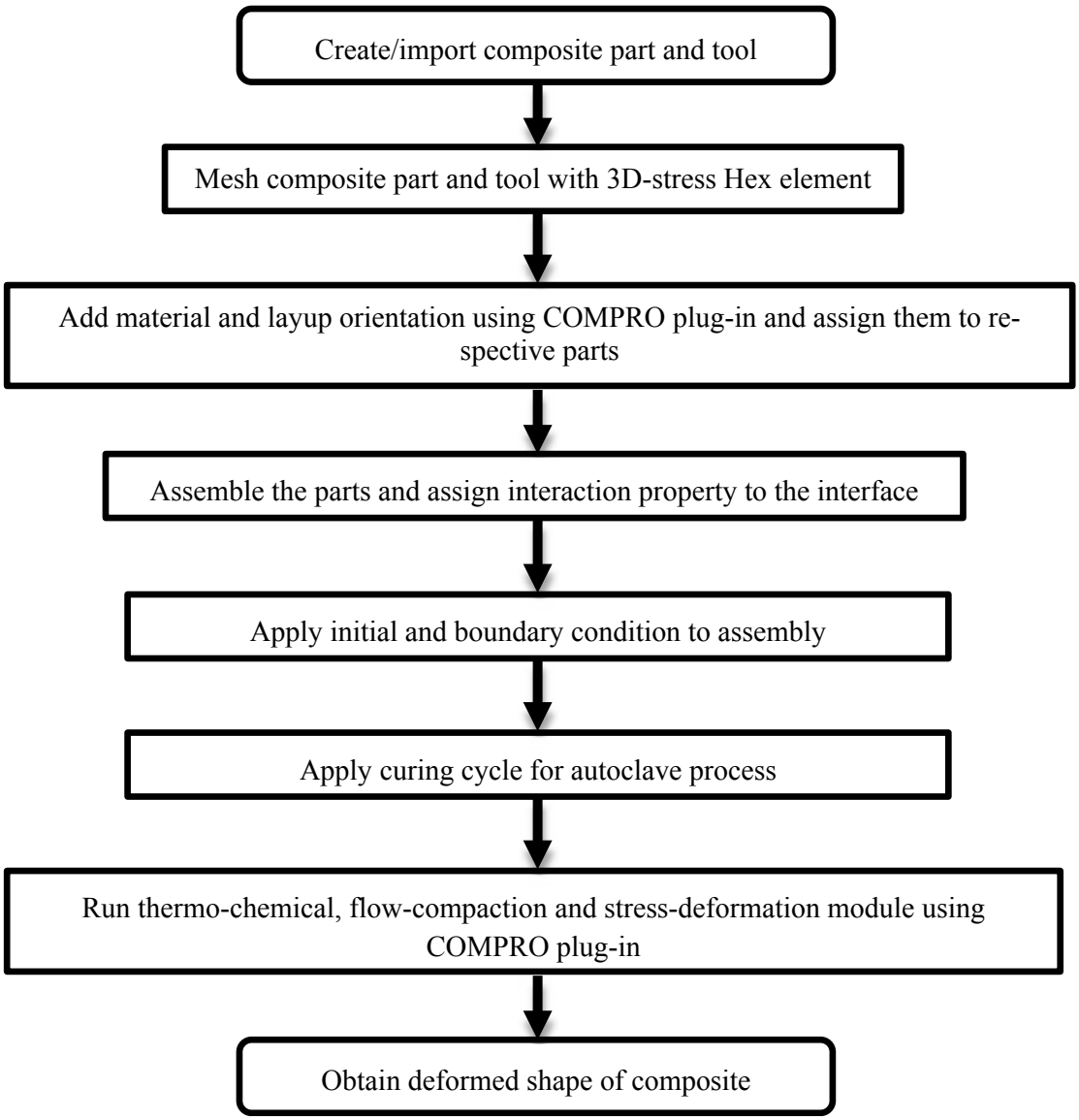


Fig. 55.4 Methodology for numerical simulation using ABAQUS with COMPRO plug-in

55.3.1 Simulation Procedure

The curing process of CFRP reflector has been simulated using ABAQUS along with COMPRO plug-in [12]. The paraboloid shape tool and reflector have been modelled using solid modelling software and imported in ABAQUS. The tool has been partitioned along two perpendicular planes passing through the centre of the model. The total of eight partitions has been created for the reflector in ABAQUS modeller. Each edge has been divided into five segments at a radius of 25, 50, 75, and 100 mm for the reflector. This will be facilitated for node generation at each junction and is further used to measure distortion at these locations. The thickness of the reflector should be modelled according to a number of prepreg layers as per thickness of each layer. The tool and reflector have been discretized using 3D Stress

20 node quadratic C3D20 elements. The meshed model for tool and reflector is shown in Fig. 55.5a, b, respectively. The stainless steel and carbon fibre prepreg material have been assigned for tool and reflector, respectively. These both materials have been imported using COMPRO plug-in. The through-thickness direction and in-plane direction has been assigned to the reflector as per stacking direction and fiber orientation to perform simulation using COMPRO. Assembly of tool and reflector has been performed by creating individual instances in the assembly module in ABAQUS. Frictional properties have been assigned to the interaction surfaces of tool and reflector. The tool top and reflector bottom surface have been selected as master and slave surface, respectively. The friction coefficient and limiting shear stress have been considered as 0.15 and 140,000 $N/mm^{2,}$ respectively. Assembly of tool and reflector has been assigned an initial temperature of 20 °C. The fix boundary constraints have been applied for tool and reflector to restrict their free expansion during the cure cycle in *X*-, *Y*-, and *Z*-directions.

Thermochemical analysis has been performed using COMPRO plug-in. The heat transfer coefficient of 80 W/m^2k has been applied on all exterior surfaces of reflector and tool. The manufacturer recommended cure cycle has been given as temperature-time data for autoclave processing. The temperature profile and degree of cure plot have been obtained as a result of thermochemical analysis. The demoulding step has been defined to get the deformed shape of the reflector. The tool has been deactivated using model change command in the interaction module in tool removal step. The proper boundary conditions have been applied to the reflector during tool removal step. The residual stress profile and deformed shape have been obtained as a result of the stress-deformation analysis.

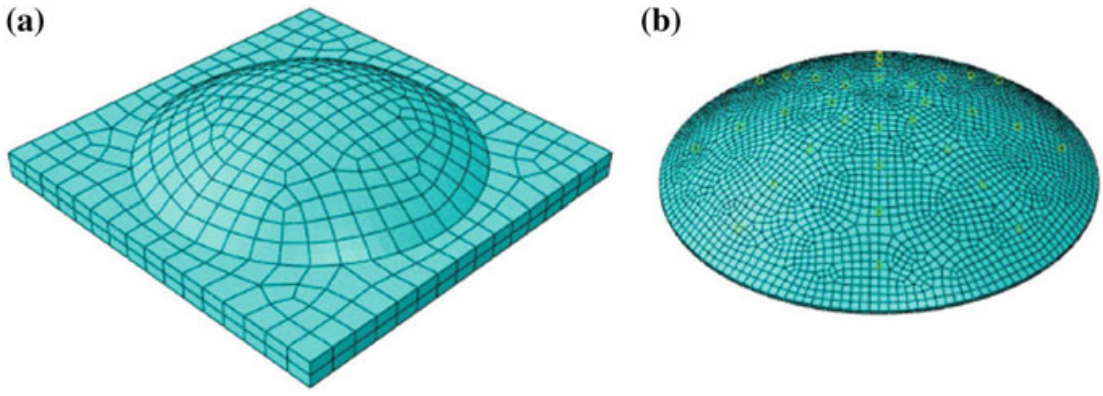


Fig. 55.5 Meshed mould and reflector

55.4 Results and Discussion

55.4.1 Distortion Measurement in Experimental Approach

The measurement of the specified coordinate points on the reflector surface has been performed using 3D scanning to confirm the accuracy of the parabolic shape obtained after curing. This has been accomplished by comparing the theoretical CAD model with 3D scan data. The 200 coordinate points have been obtained by projecting circles of diameter 50, 100, 150, 200, and 240 mm on the parabolic surface of the reflector. The deviation for scan data of reflector and CAD model has been measured for each coordinate point, and the same has been shown as a contour plot in Fig. 55.6. The distortion for each fabricated reflector has been calculated using root mean square (RMS) of deviations for all coordinate points. The predefined coordinate points on the CAD model and scanned data points on reflector along with deviation are given in Table 55.1. The final distortion, i.e., RMS value has been obtained as 0.37 mm.

55.4.2 Distortion Measurement in Numerical Approach

The deformation for total 41 coordinate points at the junction of division and partition on edge has been obtained as shown in Fig. 55.7. The original and deformed shape after demoulding for reflector is shown in Fig. 55.8. The deviation between coordinate points of original and deformed shape reflector are given in Table 55.2. The distortion for each fabricated reflector has been calculated using root mean square (RMS) of deviations for all coordinate points. The final distortion, i.e., RMS value has been obtained as 0.43 mm.

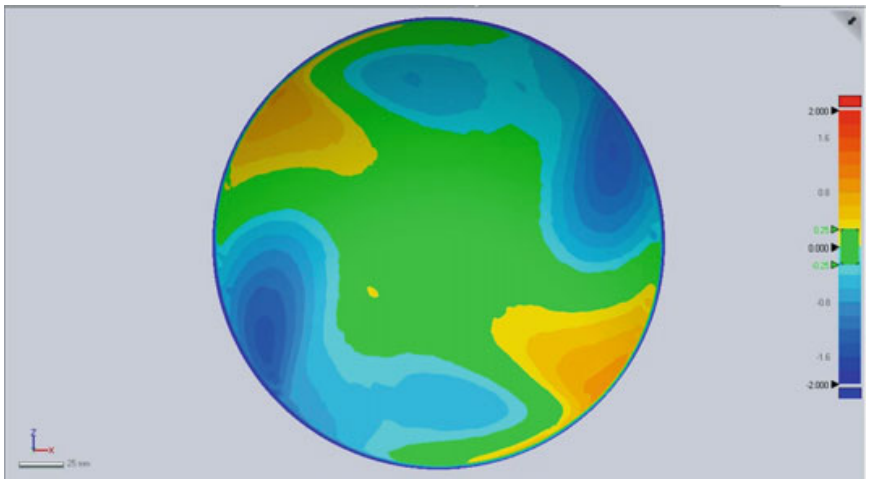


Fig. 55.6 Deviation of individual predefine coordinate

Table 55.1 Deviation of each coordinate point in autoclave curing process

Sr. no.	Predefined coordinate (Nominal) points (mm)			Scanned coordinate points (mm)			Deviation (mm)
1	0	−2.003	25	0	−2.051	24.992	0.04866
2	−3.911	−2.003	24.692	−3.911	−2.007	24.691	0.00412
3	−7.725	−2.003	23.776	−7.726	−1.978	23.78	0.02534
4	−11.35	−2.003	22.275	−11.353	−1.957	22.281	0.04649
5	−14.694	−2.003	20.225	−14.7	−1.947	20.232	0.05675
6	−17.677	−2.003	17.677	−17.683	−1.948	17.683	0.05565
.							
.							
.							
200	18.772	−46.083	118.521	18.728	−46.436	118.253	0.44524
Min	−120	−46.084	−120	−120.303	−47.054	−119.51	0.001
Max	120	0	120	120.454	0.0057	119.784	1.225
RMS							0.37

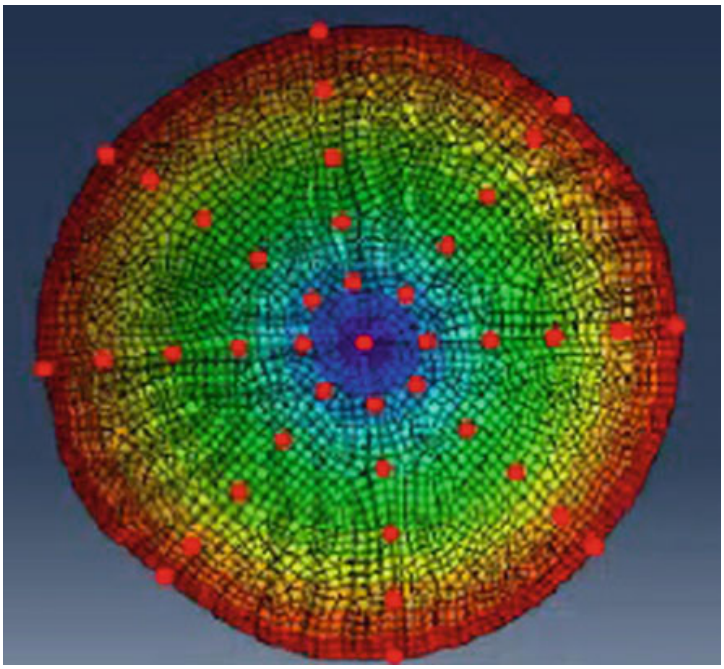


Fig. 55.7 Deformed coordinates of reflector

55.4.3 Comparison of Experimental and Numerical Results

Distortion of CFRP antenna reflectors has been determined using experimental and numerical approaches as given in Tables 55.1 and 55.2, respectively. Numerical results also compare quite well with experimental results. It has been shown that the difference between them in the range of 5–14%. The similar procedure has been repeated for different laminate thicknesses, i.e., 0.8, 1.6, 2, and 2.4 mm using both

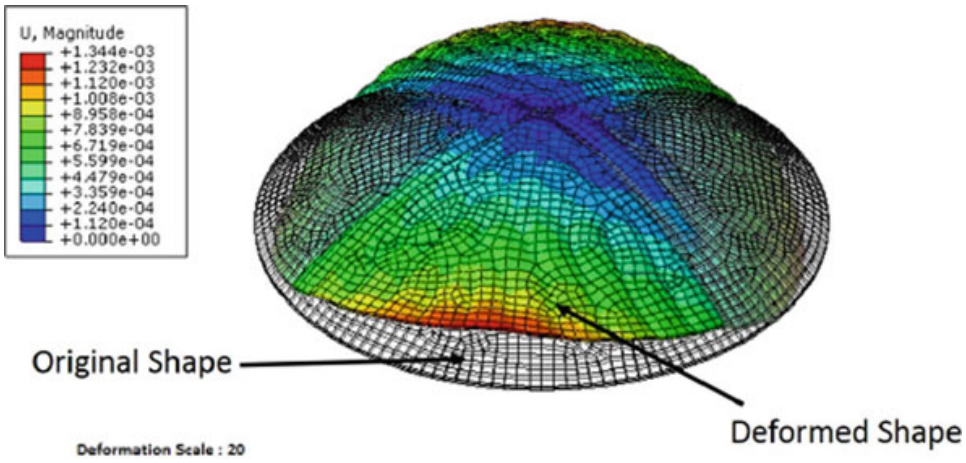


Fig. 55.8 Original and deformed shape of reflector

Table 55.2 Deviation of each coordinate point in the simulation process

Sr. no.	Predefined coordinate (Nominal) points (mm)			Deformed coordinate points (mm)			Deviation (mm)
1	0	0	0	0	0	0	0
2	25	−2	0	25.007	1.91	0.082	0.121
3	50	−8	0	50.010	7.91	0.160	0.187
4	75	−18	0	74.978	17.98	0.234	0.235
5	100	−32	0	99.885	32.15	0.302	0.356
6	120	−48	0	119.759	46.41	0.354	0.539
.							
.							
.							
41	84.8528	−48	−84.853	85.211	45.50	−84.290	0.886
Min	−120	−48	−120	−119.398	−47.39	−119.76	0
Max	120	0	120	119.759	0	119.648	1.225
RMS							0.43

approaches. The measured deformation using experimental and numerical approaches has been found as a good agreement with each other. However, for lower thickness components, the percentage difference is higher compared to high thickness components. The distortion for the corresponding reflector has been determined and compared with each other as shown in Fig. 55.9. The distortion decreases as laminate thickness increases due to the transverse contraction in the composite part. The close examination of the predicted final reflectors shape reveals that non-uniform thermal and cure shrinkage strains due to the coefficient of thermal expansion and coefficient of cure shrinkage anisotropies would result in changes in the shape of geometry. A sensitivity analysis revealed that the part thickness variation is the most important source of the predicted distortion.

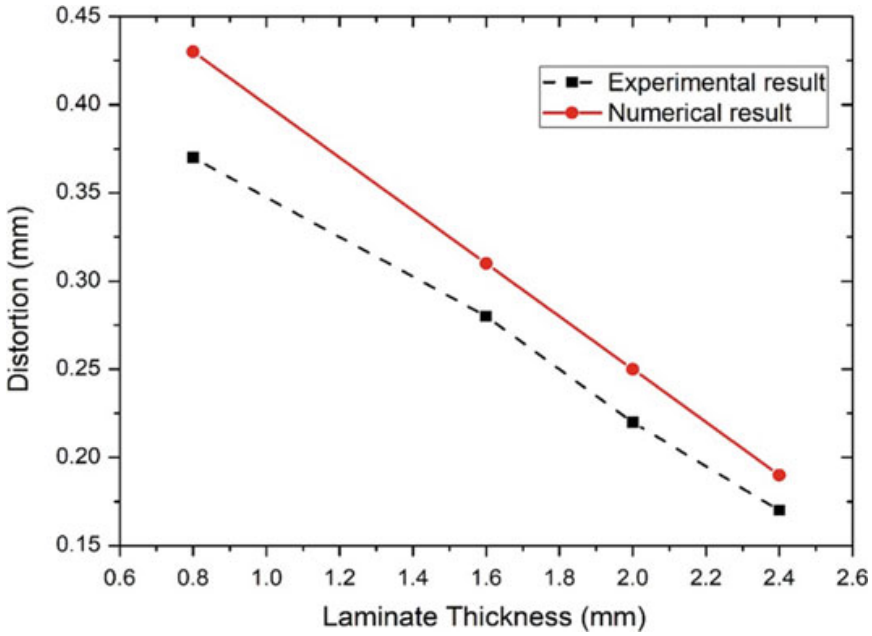


Fig. 55.9 Comparison of distortion for experimental and numerical result

55.5 Conclusions

Carbon fibre-reinforced polymer (CFRP) antenna reflectors have been fabricated using autoclave processing. The HCU200/A45 prepreg and stainless steel AISI 430 as a part and tool material, respectively, have been used in this study. The deviation in reflectors has been measured using 3D scanning technique. The geometric modelling and finite element simulation have been performed using ABAQUS software along with COMPRO plug-in. Thermochemical and stress-deformation analysis has been executed to obtain the distortion of the antenna reflectors. The distortion for reflectors has been determined using both approaches. The laminate thickness has been considered as a process parameter. The distortion decreases as laminate thickness increases. The measured deformation using experimental and numerical approach has been compared and found within 5–15% range.

Acknowledgements The authors express their gratefulness to the Convergent Manufacturing Technologies, CANADA for providing COMPRO software as well as for their technical support and help.

References

1. Hahn, H.T., Pagano, N.J.: Curing stresses in composite laminates. J. Compos. Mater. (1975)
2. Darrow, D.A., Smith, L.V.: Isolating components of processing induced warpage in laminated composites. J. Compos. Mater. (2002)

3. Twigg, G., Poursartip, A., Fernlund, G.: Tool–part interaction in composites processing. Part I: experimental investigation and analytical model. Compos. A **35**, 121–133 (2004)
4. Stefaniak, D., Kappel, E., Spröwitz, T., Hühne, C.: Experimental identification of process parameters inducing warpage of autoclave-processed CFRP parts. Compos. A **43**, 1081–1091 (2012)
5. Kaushik, V., Raghavan, J., Zeng, X.: Experimental study of tool–part interaction during autoclave processing of thermoset polymer composite structures. Compos. A **41**, 1210–1218 (2010)
6. Fernlund, G., Poursartip, A.: The effect of tooling material, cure cycle, and tool surface finish on spring-in of autoclave processed curved composite parts
7. Zeng, X., Raghavan, J.: Role of tool-part interaction in process-induced warpage of autoclave-manufactured composite structures. Compos. A **41**, 1174–1183 (2010)
8. Abusafieh, A., Federico, D., Connell, S., Cohen, E.: Dimensional stability of CFRP composite for space based reflectors
9. Dong, C.: Modeling the dimensional variations of composites using effective coefficients of thermal expansion. J. Compos. Mater. **43**(22) (2009)
10. Mezeix, L., Seman, A., Nasir, M.N., Aminanda, M., Rivai, Y., Castanie, A., Olivier, P., Ali, K.M.: Spring-back simulation of unidirectional carbon/epoxy flat laminate composite manufactured through autoclave process. Compos. Struct. **124**, 196–205 (2015)
11. Compro Model Documentation, Release 1.3. Convergent Manufacturing Technologies Inc. (2013)
12. COMPRO for ABAQUS Modelling Guidelines. Convergent Manufacturing Technologies Inc. (2014)

Chapter 56
Comparative Study of Hydroxyapatite Nanocomposites Reinforced with Zirconia and Titania Produced by Using HEBM

Vemulapalli Ajay Kumar, Penmetsa Rama Murty Raju, Nallu Ramanaiah and Siriyala Rajesh

Abstract Hydroxyapatite ($Ca_{10}(PO_4)_6(OH)_2$–HAp) conglomerate with various weight fractions (5, 10, 15, 20 and 25 wt%) is fabricated by reinforcing TiO_2 and ZrO_2 + TiO_2 nanoparticles (20–40 nm). In order to obtain homogeneous mixing of HAp, particles are reinforced using high-energy ball milling (HEBM) at 300 rpm for 1 h, compacted at 100 bar with holding time of 150 s and sintered at 1200 °C. X-ray diffraction (XRD) analysis, energy-dispersive spectroscopy (EDX) and microstructural analysis using field emission scanning electron microscopy (FESEM) have been carried out on the obtained samples. The mechanical characterization of the composites has been evaluated through flexural strength, compression strength, hardness, young modulus and fracture toughness. The mechanical properties are found to be improved with increasing content of TiO_2; however, with increasing content of TiO_2 + ZrO_2, mechanical properties are found to be improved significantly up to the addition of 20 wt % of reinforcement content. When 25 wt% of TiO_2 + ZrO_2 reinforcement is added to HAp, the mechanical properties are found to be decreased. The reduction could be due to the increase in dominant smaller particles of ZrO_2 and grain size. The improved mechanical properties are correlated with the observed microstructural features.

Keywords High-energy ball milling · ZrO_2 · TiO_2 · Mechanical properties

V. Ajay Kumar · P. Rama Murty Raju · S. Rajesh (✉)
Department of Mechanical Engineering, S.R.K.R. Engineering College, China Amiram, Bhimavaram 534204, India
e-mail: rajeshsiri.mech@gmail.com

V. Ajay Kumar
e-mail: ajayvijaya02@gmail.com

P. Rama Murty Raju
e-mail: dr.prmraju@gmail.com

N. Ramanaiah
Department of Mechanical Engineering, AU College of Engineering (A), Andhra University, Visakhapatnam 530003, India
e-mail: n.rchetty1@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_56

56.1 Introduction

The chemical composition or the structure similar to human hard tissues such as bone and teeth exhibits biocompatibility and bioactivity of hydroxyapatite (HAp). HAp, after implantation into human body as artificial material, can provide scaffolds for the formation of new bone tissues. It plays a role of osteoconduction because its calcium and phosphorous can be free from the material surface and absorption of the new tissues. In the field of medical sciences, it is considered to be a vital aspect with respect to human health, and the bioactive ceramics such as HAp are attractive candidates for body's hard tissues replacement [1–4] among different categories. The application of HAp has excellent biocompatibility properties, but the application is limited due to its low strength and brittle nature, but it has excellent biocompatibility [5]. Prokopiev et al. [6] reported microstructure and mechanical properties of sintered HAp on the sintering temperature at fixed compaction pressure (20 MPa) and also observed that Young's modulus and shear modulus (isotropic) increase with sintering temperature between 1200 and 1280 °C. Lee et al. [7] studied about Al_2O_3 and ZrO_2. Due to their excellent oxidation resistance, good biocompatibility and wear resistance, they are considered as a matrix as well as reinforcement phases. Que et al. [8] reported the major effect on the HAp structure and enhanced HAp properties on addition of titania to HAp. Due to the introduction of secondary phases, the phase changes in the composites at higher sintering temperatures are observed. Viswanath and Ravi Shankar [9] studied the interfacial reactions in hydroxyapatite/alumina nanocomposite in which alumina reacted completely with hydroxyapatite. Xihua et al. [10] reported the fabricating of hydroxyapatite matrix ceramic composites with low cost; alumina is introduced in hydroxyapatite.

Yu et al. [11] reported the implication of nano-ZrO_2 on the mechanical properties of reinforced HAp bone cement. Hongbo et al. [12] investigated first of its kind the fabrication of laminated and functionally graded HAp/Y-TZP composites by means of the spark plasma sintering (SPS). The functionally graded HAp/Y-TZP and laminated HAp/Y-TZP are designed strong Y-TZP core-bioactive HAp layer structure, a non-symmetric Y-TZP layer to HAp later structure or a bioactive HAp core-strong Y-TZ player structure. Aminzare et al. [13] studied the nanocomposite by HEBM in order to obtain highly dense objects with desired mechanical properties. The phase decomposition and mechanical properties have been investigated with the effects of alumina and titanianoparticles on microstructure. Ahmed et al. [14] reported that the titanianano/hydroxyapatite composites produced by HEBM. The effect of addition of titania, sintering temperature, formed phases time and their influence regarding compressive strength of HAp were investigated. Khalil et al. [15] studied the mechanical and microstructure properties of hydroxyapatite and hydroxyapatite-yttria stabilized zirconia (3YSZ) with 20 vol%-(ZrO_2 + 3 mol% Y_2O_3) nanopowders developed by High Frequency Induction Heat Sintering (HFIHS) at different temperatures. The main objective of their study is to synthesize and then sinter the HAp–3YSZ powders with higher hardness, higher toughness,

grain size in fine scale and constituent distribution homogeneously. Oktar [16] observed the influence of sintering process of the composite materials, made of hydroxyapatite, that were derived from the bovine bone, with 5 and 10 wt% of TiO_2, on their mechanical properties. The mechanical properties of the produced composite materials revealed the developed microstructure and the phases formed during sintering effecting the densification. Mobasherpour et al. [17] adopted a precipitation method to synthesize nanocrystalline HAp and investigated the effects of ZrO_2–Al_2O_3 on mechanical properties of nanocrystalline HAp powder. Improved bending and the tensile strength of HAp by using La_2O_3, ZrO_2, and TiO_2 as binders have been observed by Shin-Ike et al. [18].

In the present study, the application of HEBM technique is to produce HAp/TiO_2 and HAp/ZrO_2 + TiO_2 nanocomposites and the effect of the addition of TiO_2 and ZrO_2 + TiO_2 to HAp on the microstructure and mechanical strength of the composites was investigated. The correlation between the mechanical and microstructural properties was systematically established.

56.2 Materials and Methods

56.2.1 Sample Preparation

Hydroxyapatite (HAp) and its composite powders of particle size 20–40 nm with various wt% of TiO_2 and ZrO_2 + TiO_2 (0, 5, 10, 15, 20 and 25 wt%) were combined with HAp by HEBM. Powders were milled under dry condition using tungsten carbide hardened steel balls for 1 h with rotational speed 300 rpm. The balls to powders ratio considered were 20:1. The milling powder samples were compacted under pressure of 100 bar at room temperature for dwell time 150 s. The equipment used for compacting has two channels which can operate from 0.005 to 200 bars and can go from room temperature to 500 °C. These compacted powders are made into cylindrical specimens (compression strength, hardness) which comprises of 20 mm diameter and 5 mm height and a plain specimen (flexural strength and fracture toughness) about 40 mm length, 10 mm width and 3 mm thickness. To sum, sintering was performed in normal atmosphere using electric resistance box furnace at 1200 °C and the heating time 1 h. The specimens were fed into the furnace and heated up to desired temperature. The heating time was calculated from the point at which the desired temperature has been reached. All the sintered samples were furnace cooled.

56.2.2 Experimental Procedure

X-ray diffractometer (XRD) equipped with Cu–K_α radiation (wavelength λ = 1.5406 Å°) is used to distinguish HAp and its composites. This test is operated between 6 and 90° using a step size of 0.01° which follows the width (β) of their

diffraction pattern; the average crystallite size (D) was estimated using Scherer's formula [14]. The microstructural information was known by field emission scanning electron microscopy (FESEM). The compositions of the elements Ca, P, O and Zr are evaluated using energy-dispersive analysis of X-rays (EDAX) attached with FESEM. The compressive strength of the nanocomposite ceramics of size 20 mm (diameter) and 5 mm (height) was conducted on UTM. The density of sinter samples was measured using distilled water by the Archimedes' method (relative density). This test was executed under a quasi-static loading range, where the initial strain rate (ε*) is fixed at 0.001 s^{-1}; i.e., the machine crosshead speed ($\vartheta = \varepsilon * h_o$) is 0.005 mm/s. And h_o is the specimen's initial height, where the test is conducted on specimens which are loaded up to fracture [14].

Hardness test (Knoop hardness indenter) is evaluated on the basis of polished HAp composites which were analyzed using the Shimadzu HMV-2000 microhardness tester. The load on the specimen 4.905 N and holding time of 15 s are used in the measurements as per ASTM E384. The width-to-length ratio of the indenter is 1:7.11. The concavity or indention depth is around (1/30)th of the longer dimension, and its face angle for the rectilinear edge is 172° and for the breadth is 130°.

Knoop hardness '*H*' can be determined by Eq. (56.1) as below [12]:

$$H = 14229 \frac{L}{d^2} \tag{56.1}$$

The load refers to '*L*' (Newton), and the longer diagonal of the indentation refers to '*d*' (mm).

Young's modulus along the composites was determined through the following Eq. (56.2) as given below [12]:

$$\frac{B'}{A'} = \frac{B}{A} - \alpha \frac{E}{H} \tag{56.2}$$

(B^1/A^1) gives the indent diagonal ratio, and (B/A) gives the ratio of the indenter dimensions. The term 'α' relates to a constant of a value of 0.45.

The three-point bending test is used to determine the flexural strength of the nanocomposite. The HAp composite specimens (40 mm length, 10 mm width and 3 mm thickness) were first prepared and refined to smooth out surface flaws before setting them on a three-point bending testing setup. The load (frame crosshead speed of 0.05 mm/min) was put from the top at middle of the bar. Through the indentation technique, the fracture toughness K_c (Eq. 56.3) was calculated, and the values were confirmed through the equation below [15]:

$$K_C = \xi \left[\frac{E}{H}\right]^{2/5} \frac{P}{d \times l^{1/2}} \tag{56.3}$$

'K_c' gives fracture toughness, 'E' refers Young's modulus (GPa), and 'H' gives the hardness (GPa). 'P' gives the applied load (N); 'ξ' shows the constant which is based on the geometry of the indenter ($\xi = 0.008$).

56.3 Results and Discussion

56.3.1 Microscopic Analysis

The morphologies of HAp and its composites are observed by a FESEM. HAp forms a crosslink with one and other during the growth process. This provides a conducive condition for the adhesion and growth of biological tissues. Figure 56.1 gives the XRD patterns of HAp and standard HAp (PdF card no. 74-0566). The diffraction pattern of HAp powder used for the current study is in good agreement with standard HAp pattern, and it is also observed that there are no reflections from impurities.

XRD patterns were used for the analysis of specimens which had undergone the process of pressing and then sintering at a temperature of 1200 °C. Figure 56.2 shows XRD patterns of HAp and its composites (HAp/TiO_2). When the TiO_2 powder was added, stronger peaks are detected along with HAp and TiO_2 when the compacting and the sintering processes were carried out under the same condition. With increase in TiO_2 (indicated with '$') content along with the HAp (indicated with '#'), TiO_2 phase reflections are present. Using Scherer's formula, all the XRD

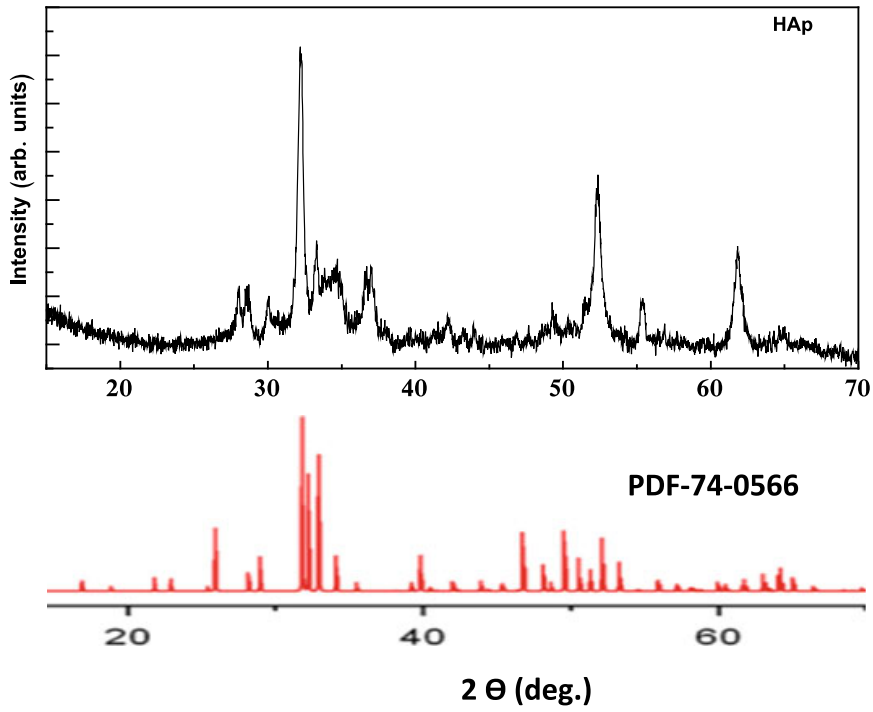


Fig. 56.1 XRD pattern of HAp sample and standard HAp

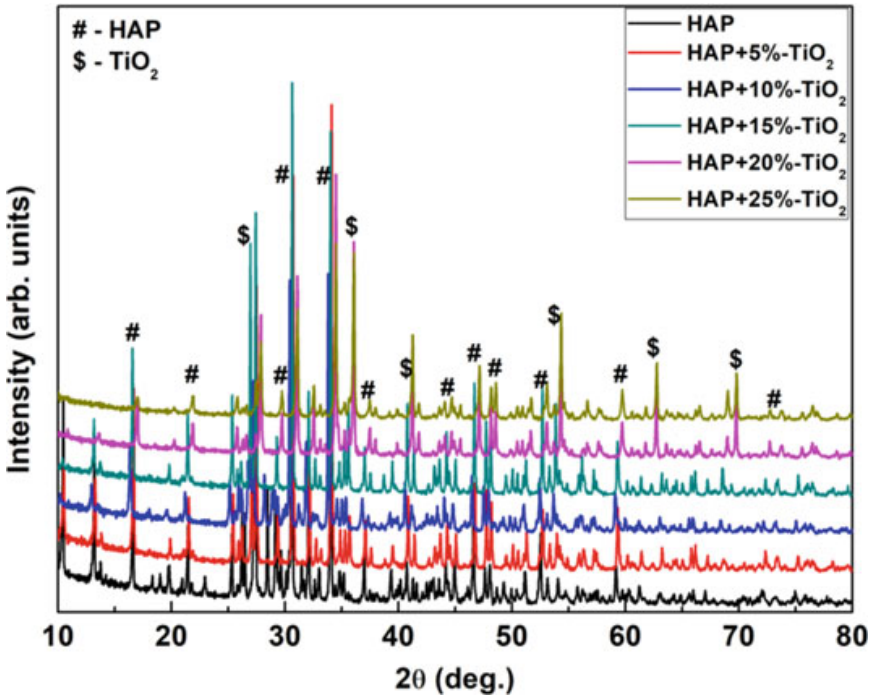


Fig. 56.2 XRD patterns of HAP/ZrO_2 composites after sintering

patterns crystallite sizes are calculated and observed, and they are in the range of 40.01–16.11 nm and 47.39–42.39 nm for HAp/TiO_2 and HAp/(TiO_2 + ZrO_2) composites, respectively. This shows clearly that there is a gradual decrease in the crystallite size when there is addition of TiO_2 in the process. It is analyzed with the increase of TiO_2 content; there will be increase in the reflections of TiO_2 phase. TiO_2 structural stability was studied with the improved TiO_2 content which in turn increased by intensities of diffraction patterns. There were no additional reflections due to the impurities.

Figure 56.3 depicts XRD patterns of HAp and its composites (TiO_2 + ZrO_2). The reflections observed were matched with the parent phase (HAp-marked with '#') and its composites (TiO_2 marked with '$' + ZrO_2 marked with '@'). It is observed that due to the impurities, there were no additional reflections. It is observed that ZrO_2 + TiO_2 attained more structural stability with increase in the diffraction patterns intensities due to the increasing of ZrO_2 + TiO_2 content. At 25 wt% of composite (TiO_2 + ZrO_2), strong tricalcium phosphate (TCP) peaks are detected along with zirconia and HAp. This formation of TCP may be attributed due to excessive reaction between the HAp and ZrO_2 with the improved wt% of ZrO_2 particles. The TCP structure is a typical biomaterial with excellent biocompatibility and osteoconductivity, by which the formation of new bone tissue is accelerated when these materials are implanted into the body which is observed by the Colon et al. [19]. TCP is fragile bioceramic and consequently cannot be used on their own as major load-bearing implants in the human body. From the identified phases as shown in Fig. 56.3, it can be seen that ZrO_2 has shown a dominant phase

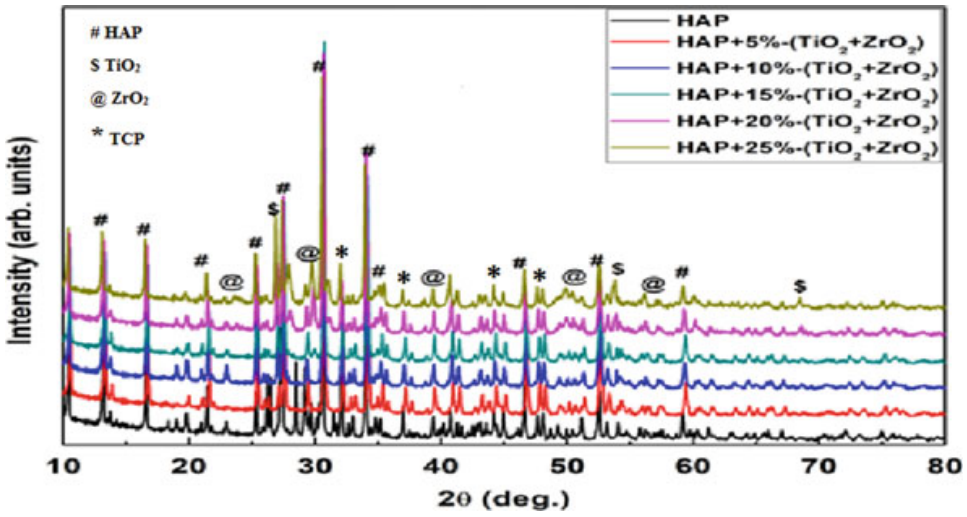


Fig. 56.3 XRD patterns of HAp/ZrO_2 + TiO_2 composites after sintering

over TiO_2 on addition of ZrO_2 + TiO_2 which is in good agreement with the studies conducted by Miao et al. [20].

Field emission scanning electron microscopy (FESEM) is used to analyze the morphology of the powders (HAp/5, 10, 15, 20 and 25 wt% TiO_2 and TiO_2 + ZrO_2) after HEBM. The morphology of the initial large particles is significantly changed due to fracture and de-agglomeration processes. After 1 h of HEB milling of the HAp/TiO_2 powders, sporadic-shaped morphologies are observed. These small particles were also observed after the cold welding phenomenon during ball milling.

FESEM micrographs of the HAp/TiO_2 composites which had undergone compacting and sintering at a temperature of 1200 °C with different wt% are shown in Fig. 56.4a–e. Based on the micrographs, it is observed that there is a considerable change in morphology due to the addition of TiO_2. This addition of different wt% may induce the nucleation which tends to the formation of leaf-like structures. The average size of leaf-like structure is observed around 18 μm, whereas a needle- and irregular shaped morphology of TiO_2 is observed for 25 wt% HAp/TiO_2 composite. Hence, it may be concluded that as the wt% of TiO_2 increases in the composite, the agglomerated irregular shape morphology is observed.

FESEM micrographs of the HAp/TiO_2 + ZrO_2 composites with different wt% are shown in Fig. 56.5a–e. A unique growth of rounded particle with small in size was exhibited by the composites with ZrO_2 + TiO_2 contents at various wt%. In Fig. 56.5a–e, the matrix is shown with smaller particles of ZrO_2 + TiO_2. From the micrographs, it is evident that there may be alteration in the composites microstructural properties due to the increase of ZrO_2 + TiO_2 content and homogeneous distribution of particles inside the grains which enhances mechanical properties. This may be attributed due to the interlocking of particles with the grain boundaries uniformly which can be observed from Fig. 56.5d and e.

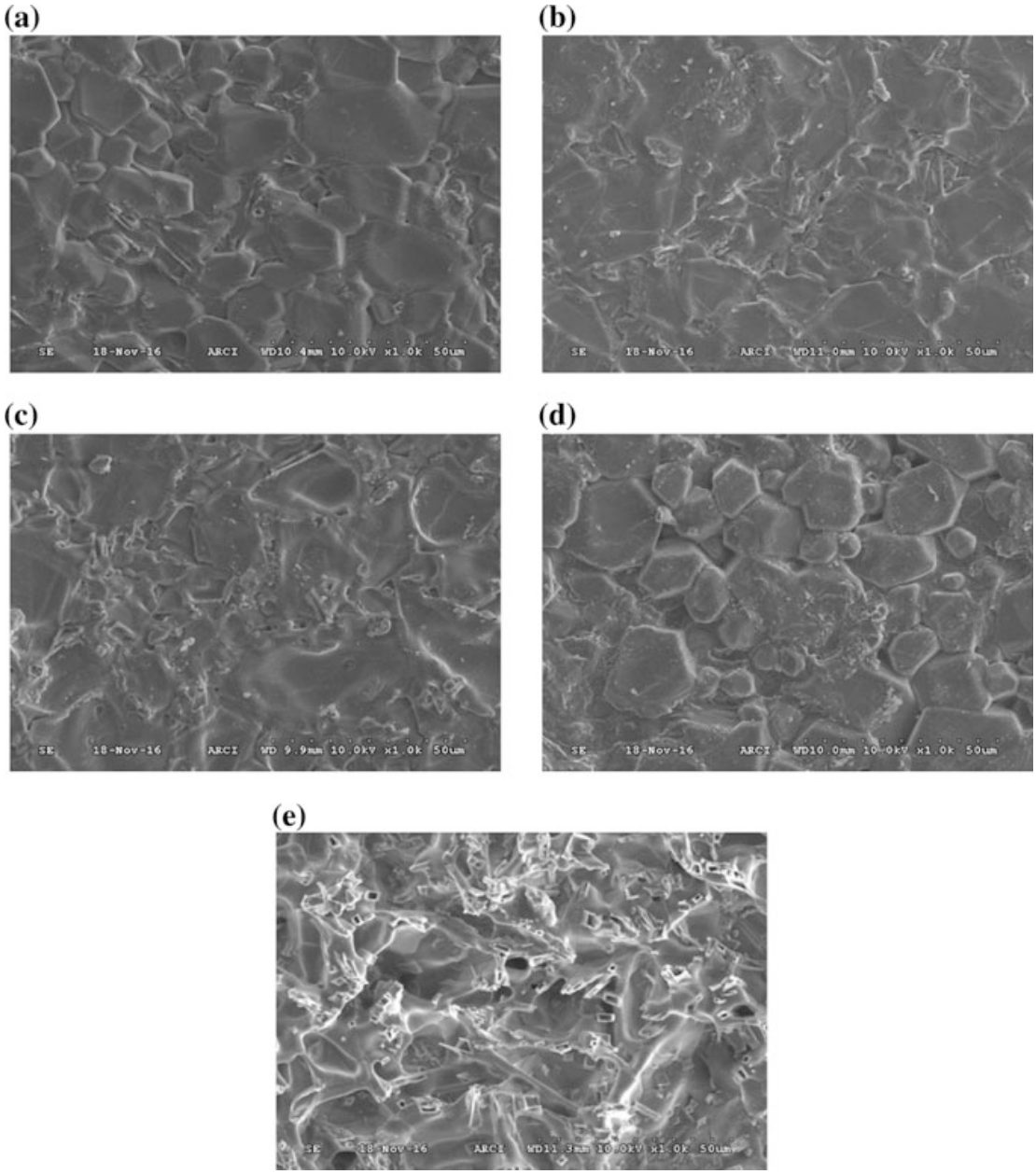


Fig. 56.4 FESEM micrographs of HAp/TiO_2 composites: **a** 5 wt%, **b** 10 wt%, **c** 15 wt%, **d** 20 wt% and **e** 25 wt%

Figure 56.6a shows energy-dispersive X-ray spectroscopy (EDX) analysis of HAp/TiO_2 composite. The presence Ca, P and Ti can be seen from the figure. Figure 56.6b shows the EDX analysis of HAp/TiO_2 + ZrO_2 composite. The EDX shows that Ca, P, Zr and Ti can be seen from Fig. 56.6b, whereas background grains exhibited strong Zr intensity but weak Ti intensity. The EDX results show that samples were free of metallic contaminants. However, contaminants could be found in very little percentages which are lower than the accuracy of the applied instrument.

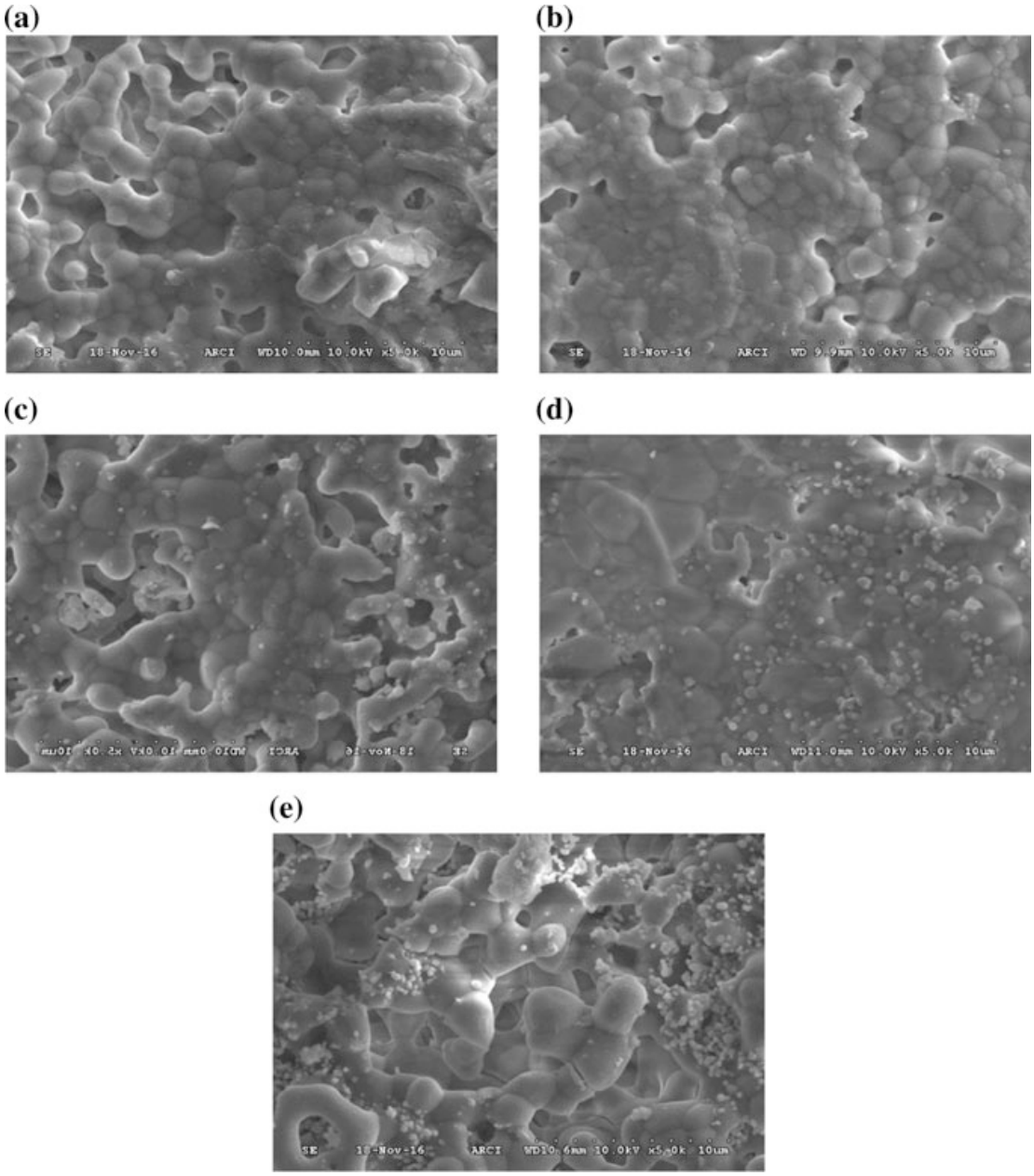


Fig. 56.5 FE-SEM micrographs of HAp/TiO_2 + ZrO_2 composites: **a** 5 wt%, **b** 10 wt%, **c** 15 wt%, **d** 20 wt% and **e** 25 wt%

Figure 56.7. shows the grain size of the HAp composites estimated from the ASTM line method [21] The grain size HAp/TiO_2 composite is observed to decrease gradually from 52 nm (5 wt% of TiO_2) to 20.1 nm with increase in TiO_2 content up to 25 wt%, whereas in HAp/ZrO_2 + TiO_2 composites, it is observed that the grain size is in the range of 42.1–19.09 nm.

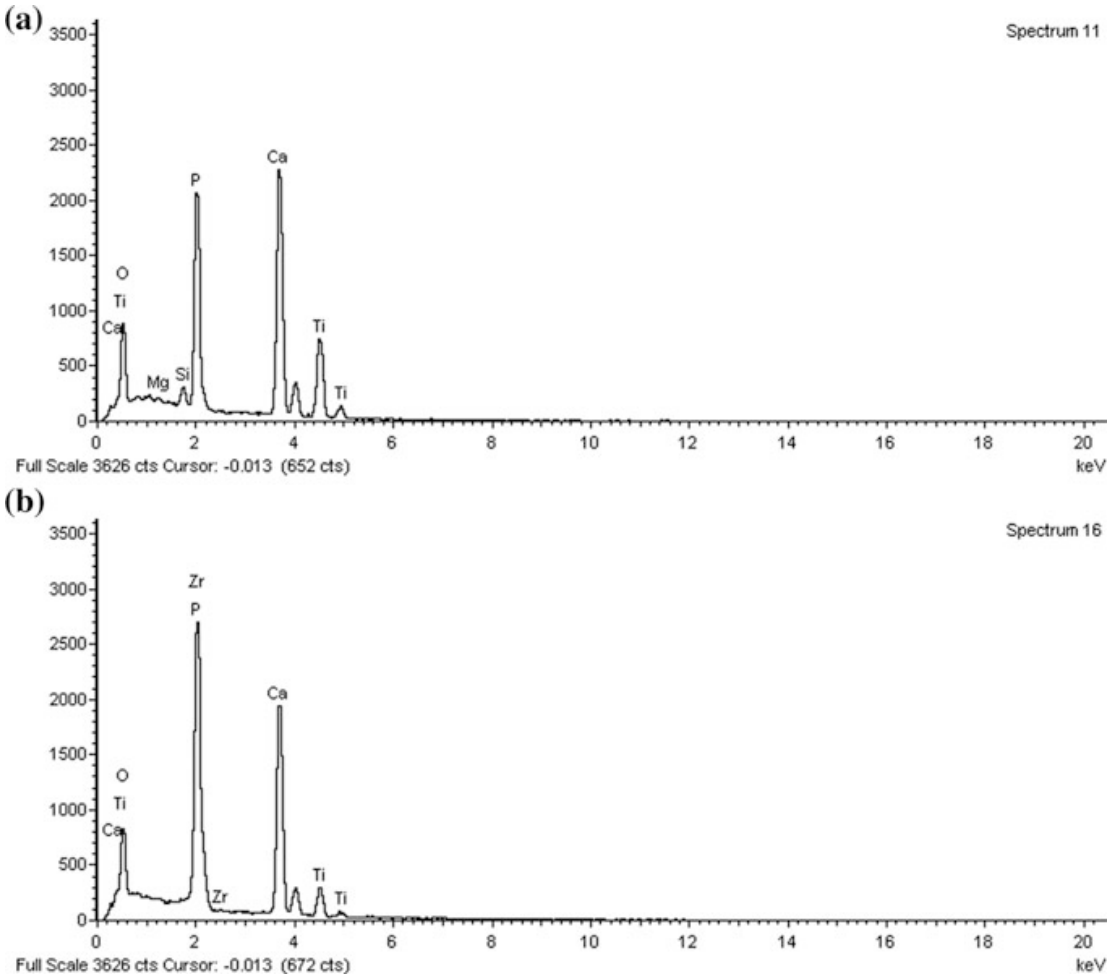


Fig. 56.6 EDAX composition of **a** HAp/25 wt% TiO_2 and **b** HAp/25 wt% ZrO_2 + TiO_2

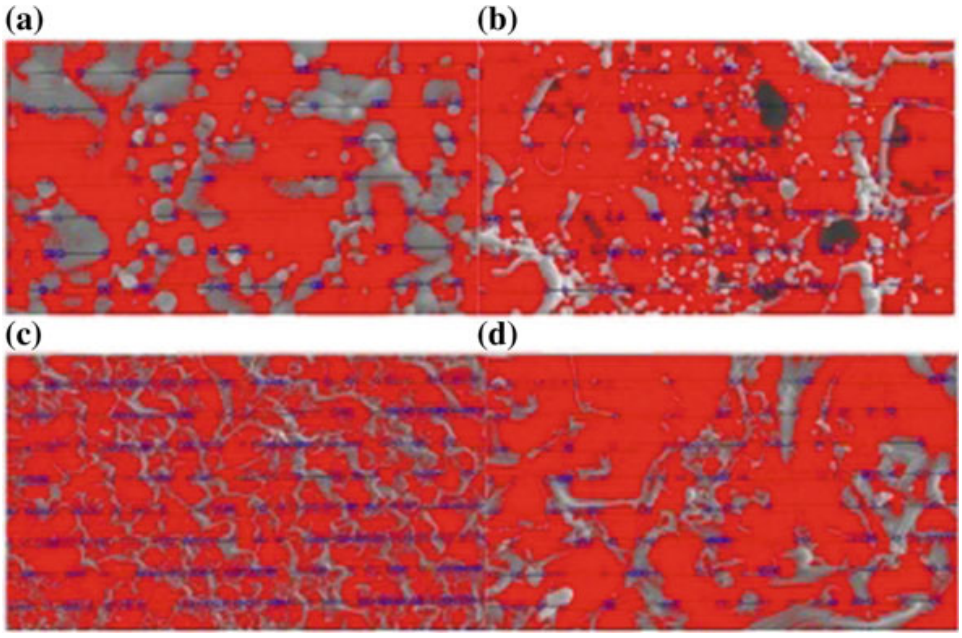


Fig. 56.7 Grain size of the composites (ASTM line method): **a** HAp + 10 wt% TiO_2, **b** HAp + 25 wt% TiO_2, **c** HAp/10 wt% ZrO_2 + TiO_2, **d** HAp/25 wt% ZrO_2 + TiO_2

56.3.2 Mechanical Properties

Based on the structural and morphological analysis, an attempt has been made to understand the formation mechanism of the composites TiO_2 and ZrO_2 + TiO_2. The measured density of HAp/TiO_2 and HAp/ZrO_2 + TiO_2 composites is shown in Fig. 56.8.

Calculated densities of HAp/TiO_2 and HAp/TiO_2 + ZrO_2 (5, 10, 15, 20 and 25 wt%) samples sintered at 1200 °C for 1 h are in close agreement with theoretical density. The density is found to be increasing gradually with TiO_2 content. With increase in TiO_2 content, it can be attributed that the grain size of HAp/TiO_2 composites decreases. The maximum density for HAp/TiO_2 + ZrO_2 composites has been obtained at 20 wt% reinforcement content. When 25 wt% of reinforcement is added to HAp, the density is found to be decreased, due to the background grains with strong Zr intensity but weak Ti intensity [20] of HAp nanocomposites, Fig. 56.5. Khalil et al. [15] has observed HAp and HAp-20 vol% 3YSZ composites at different sintered temperatures by HFIHS process.

Figure 56.9 shows the variation of hardness with reference to wt% of HAp composites. The effect of reinforcement on hardness of HAp composites can be studied as a function of density and grain size.

The increment of hardness is observed from 6.01 to 10.03 GPa. With the improvement in the TiO_2 content, the hardness of the HAp composites has been increased; at that time, grain size was observed a gradual decrease, whereas the density and the hardness increased. Among all the HAp composites (5, 10, 15, 20 and 25 wt%), maximum hardness value is 10.03 GPa, observed for the 25 wt% TiO_2 content. However, for HAp/ZrO_2 + TiO_2 composites, the hardness is increased from 5.56 to 14.01 GPa, with increasing content of ZrO_2 + TiO_2 from 5 to 20 wt%. The highest hardness (14.01 GPa) has been observed at 20 wt% ZrO_2 + TiO_2 content. Interestingly, a drop in hardness (9.98 GPa) has been observed for 25 wt% ZrO_2 + TiO_2 content. The similar work was conducted by Guo et al. [12] is observed to be in good agreement with Knoop hardness values (7.5 GPa at 20 wt% of Y-TZP) observed.

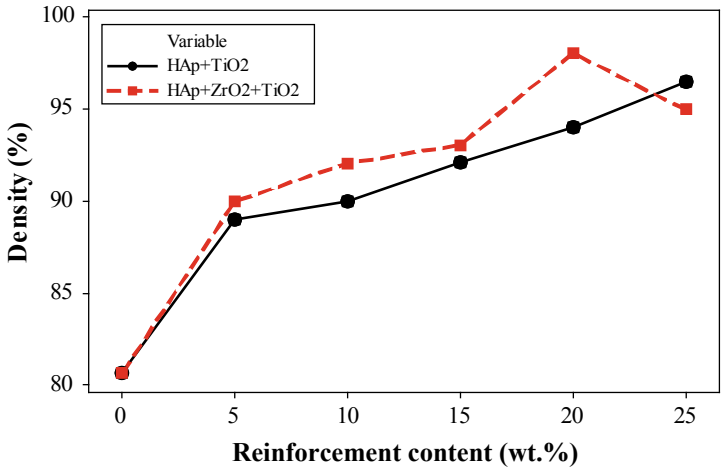


Fig. 56.8 Variation of density in HAp/TiO_2 and HAp/ZrO_2 + TiO_2 composites

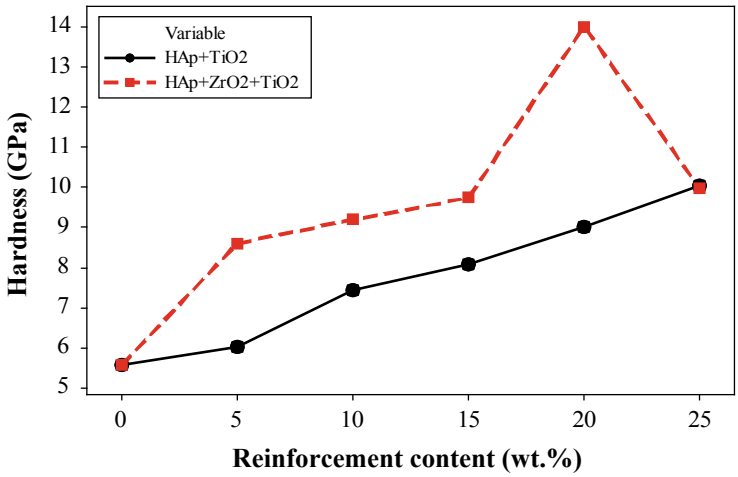


Fig. 56.9 Variation of hardness in HAp/TiO_2 and HAp/ZrO_2 + TiO_2 composites

Figure 56.10 shows Young's modulus of HAp/TiO_2 and HAp/ZrO_2 + TiO_2 sintered composite samples. With increase in wt% of TiO_2 content, Young's modulus of the HAp/TiO_2 composites is increasing from 91.2 to 128.5 GPa. The highest value of Young's modulus of 128.5 GPa has been obtained for 25 wt% of TiO_2 composite, whereas for HAP/ZrO_2 + TiO_2 composites, Young's modulus of the HAp composites was found to be increasing from 89.09 to 328.08 GPa. Young's modulus of sintered HAp composites samples increases with increasing ZrO_2 + TiO_2 content. For the HAP/ZrO_2 + TiO_2 composites, Young's modulus reaches a maximum at 20 wt%, reinforcement content followed by a slight decrease for 25 wt% composite. With increase in reinforcement content, grain size decreases and density increases; as a result, hardness is improved. The effect of the ceramic as additives is to improve the density and young's modulus values [8]. Due to the same reason, Young's modulus has also been improved with increase in reinforcement content.

Figure 56.11 shows the fracture toughness trend for HAp/TiO_2 and HAp/ZrO_2 + TiO_2 sintered composites with varying wt% of reinforcement. The fracture toughness is found to be varying from 0.742 to 3.05 MPa $m^{1/2}$. The highest value of fracture toughness of 3.05 MPa $m^{1/2}$ has been obtained for 25 wt% of TiO_2

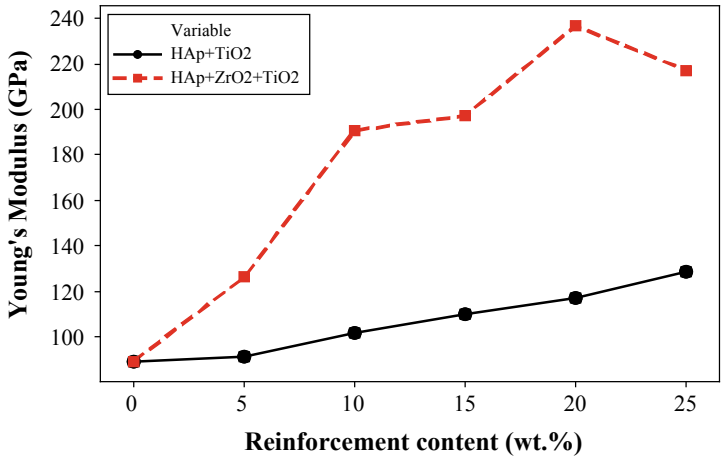


Fig. 56.10 Variation of Young's modulus in HAp/TiO_2 and HAp/ZrO_2 + TiO_2 composites

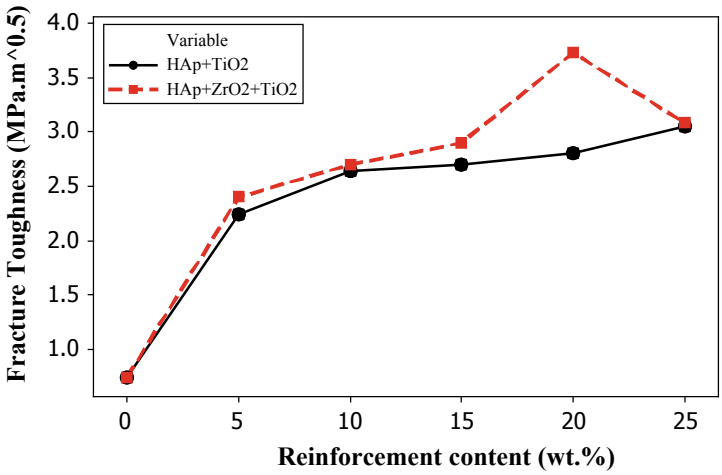


Fig. 56.11 Variation of fracture toughness in HAp/TiO_2 and HAp/ZrO_2 + TiO_2 composites

composite, whereas for HAp/ZrO_2 + TiO_2 composite, the values vary from 0.742 to 3.73 MPa $m^{1/2}$ and maximum fracture toughness 3.73 MPa $m^{1/2}$ has been obtained at 20 wt% of composite. Upon further addition of reinforcement content, i.e. 25 wt%, there is a drop in fracture toughness. It is true that hardness and Young's modulus depend on the density and grain size of composite. Further, it can also be concluded that fracture toughness in turn depends on hardness and Young's modulus.

Figure 56.12 shows the flexural strength of HAp/TiO_2 and HAp/TiO_2 + TiO_2 composites. With increasing wt% of TiO_2, the flexural strength varies from 50–206 MPa for HAp/TiO_2 composite. The highest value of fracture toughness of 206 MPa has been obtained for 25 wt% of TiO_2 composite, whereas for HAp/ZrO_2 + TiO_2 composites, the flexural strength varies from 50 to 236.8 MPa, and the maximum flexural strength is observed at 20 wt% of HAp/ZrO_2 + TiO_2 composite after which there is a dropin flexural strength for 25 wt% of reinforcement. The sintered samples with higher density improved mechanical properties, flexural strength, compression strength and hardness. Oktar [16] observed same results for bending and hardness at different temperatures, strength 86.2 MPa and hardness 143.5 HV (HAp-TiO_2 5 wt% at 1200 °C).

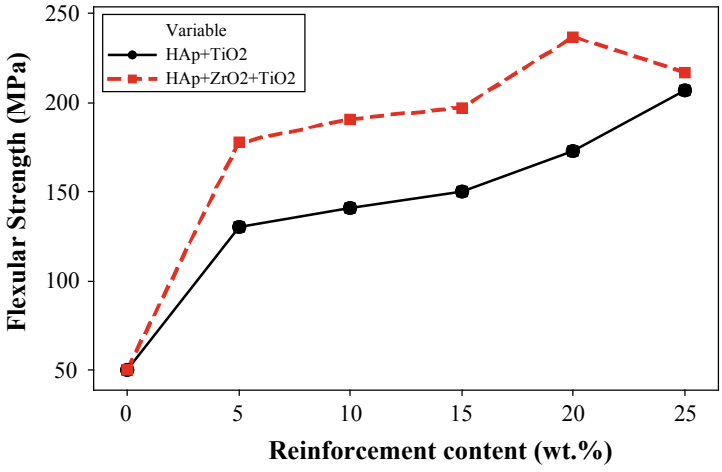


Fig. 56.12 Variation of flexural strength HAp/TiO_2 and HAp/ZrO_2 + TiO_2 composites

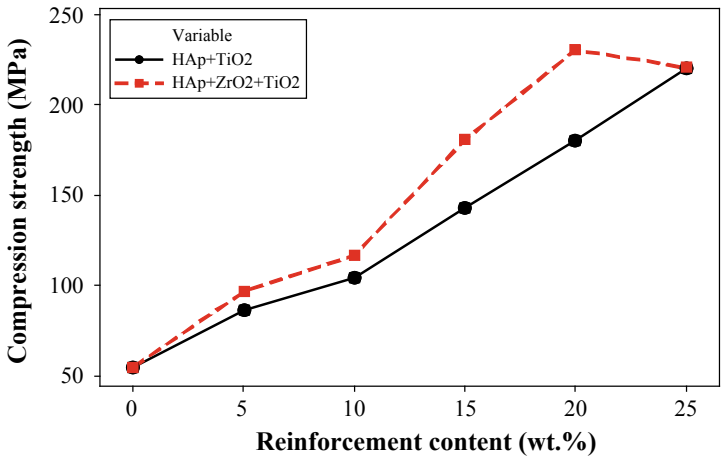


Fig. 56.13 Variation of compressive strength for HAp/TiO_2 and HAp/ZrO_2 + TiO_2 composites

Figure 56.13 can be observed that for HAp/TiO_2 composite, the compression strength varies from 50 to 220.06 MPa with increasing wt% of TiO_2. The maximum compression strength of 220.6 MPa is obtained at 25 wt% TiO_2, whereas the HAp/(ZrO_2 + TiO_2) composite has compression strength values in the range of 50–230.67 MPa. The maximum compressive strength of 230.67 MPa was obtained for 20 wt%, after which it shows a drop in compression strength. The reason can be attributed to the increase in the dispersion of reinforcement particles in the matrix phase, Figs. 56.4 and 56.5. Que et al. [8] observed the hardness and young's modulus of the pure HAp samples and the HAp/TiO_2 (10 mol%) composites as a function of the sintering temperature. The highest hardness and Young's modulus, which are 2.81 and 45.33 GPa, respectively, are obtained at the sintering temperature of 1100 °C. Salman et al. [22] presented the fabrication and characterization of composite materials of hydroxyapatite and Ti. Hydroxyapatite (HA) powder was obtained from bovine bones (BHA) and human enamel (EHA) via calcination technique. Powder-compacts were sintered at different temperatures between 1000 and 1300 °C. The best mechanical properties for BHA–Ti composites were obtained after sintering in the range of 1200–1300 °C and for EHA–Ti composites in the range of 1100–1300 °C.

56.4 Conclusions

The HAp/TiO_2 and HAp/ZrO_2 + TiO_2 composites are successfully prepared using HEBM. All the mechanical properties are investigated with HAp and its composites. It is observed that with increase in reinforced content wt% of composite the mechanical properties are found to be improved.

The following major observations can be drawn from the present investigation:

i. All the composites of HAp and TiO_2 possess hardness values of 5.56–10.3 GPa, and the hardness values of HAp/ZrO_2 + TiO_2 composites lie in the range of 5.56–14.01 GPa.

ii. Young's modulus of the sintered HAp/TiO_2 composites varied from 89.09 to 128.5 GPa. Young's modulus of the sintered HAp/ZrO_2 + TiO_2 composites varied from 89.09 to 328.08 GPa which are high compared to the pure HAp.
iii. The composites of HAp and TiO_2 fracture toughness are found to be in the range of 0.742–3.05 MPa $m^{1/2}$, and HAp/ZrO_2 + TiO_2 composites fracture toughness was found to be in the range of 0.742–3.73 MPa $m^{1/2}$. Comparing the hardness, young's modulus and toughness of pure HAp, HAp/TiO_2 and HAp/(ZrO_2 + TiO_2), composites with 20 wt% are found to yield better.
iv. The flexural strength of HAp/TiO_2 composites increased from 50 to 206 MPa, whereas the flexural strength of HAp/ZrO_2 + TiO_2 composites was found to increase from 50 to 236.8 MPa.
v. The compressive strength of HAp/TiO_2 composites for 25 wt% addition is 220.67 MPa significantly higher than pure HAp, whereas the compressive strength of the 20 wt% HAp/ZrO_2 + TiO_2 composites is 230.67 MPa which is higher than pure HAp.
vi. By adding ZrO_2 + TiO_2, all the mechanical properties have been enhanced. HAp/ZrO_2 + TiO_2 composites had better mechanical properties than that HAp/TiO_2. The investigated biocomposite materials can be used for implant applications.

References

1. Pushpakanth, S., Srinivasan, B., Sreedhar, B., Sastry, T.P.: An in situ approach to prepare nanorods of titania-hydroxyapatite(TiO2-HAp) nanocomposite by micro wave hydrothermal technique. Mater. Chem. Phys. **107**(2–3), 492–498 (2008)
2. Xihua, Z., Changxia, L., Musen, L., Yunqiang, B., Junlong, S.: Fabrication of hydroxyapatite/diopside/alumina composites by hot press sintering process. Ceram. Int. **35**(5), 1969–1973 (2008)
3. Miao, X., Hu, Y., Liu, J., Huang, X.: Hydroxyapatite coating on porous zirconia. Mater. Sci. Eng., C **27**(2), 257–261 (2007)
4. Nath, S., Biswas, K., Basu, B.: Phase stability and micro structure development in hydroxyapatite-mullite system. Scr. Mater. **58**(12), 1054–1057 (2008)
5. He, L.-H., Standard, O.C., Huang, T.T.Y., Latella, B.A., Swain, M.V.: Mechanical behavior of porous hydroxyapatite. Acta Biomater. **4**(3), 577–586 (2008)
6. Prokopiev, O., Sevostianov, I.: Dependence of mechanical properties of sintered hydroxyapatite on the sintering temperature. Mater. Sci. Eng., A **431**(1–2), 218–227 (2006)
7. Lee, B.T., Jang, D.H., Kang, I.-C., Lee, C.-W.: Relation between microstructures and material properties of novel fibrous Al_2O_3-(m-ZrO_2)/t-ZrO_2 composites. J. Am. Ceram. Soc. **88**(10), 28974–2878 (2005)
8. Que, W., Khor, K.A., Xu, J.L., Yu, L.G.: Hydroxyapatite/titaniana—nanocomposites derived by combining high-energy ball milling with spark plasma sintering processes. J. Eur. Ceram. Soc. **28**(16), 3083–3090 (2008)
9. Viswanath, B., Ravishankar, N.: Interfacial reactions in hydroxyapatite/alumina nanocomposites. Scripta Mater. **55**(10), 863–866 (2006)

10. Xihua, Z., Changxia, L., Musen, L., Yunqiang, B., Junlong, S.: Fabrication of hydroxyapatite/dropsied/alumina composites by hot-press sintering process. Ceram. Int. **35**(5), 1969–1973 (2009)
11. Yu, W., Wang, X., Zhao, J., Tang, Q., Wang, M., Ning, X.: Preparation and mechanical properties of reinforced hydroxyapatite bone cement with nano-ZrO_2. Ceram. Int. **41**(9-A), 10600–10606 (2015)
12. Guo, H., AikKhor, K., Boey, Y.C., Miao, X.: Laminated and functionally graded hydroxyapatite/yttria stabilized tetragonal zirconia composites fabricated by spark plasma sintering. Biomaterials **24**(4), 667–675 (2003)
13. Aminzare, M., Eskandari, A., Baroonian, M.H., Berenov, A., Razavi Hesabi, Z., Taheri, M., Sadmezhaad, S.K.: Hydroxyapatite nanocompatites: synthesis, sintering and mechanical properties. Ceram. Int. **39**(3), 2197–2206 (2013)
14. Hannora, A.E., Ataya, S.: Structure and compression strength of hydroxyapatite/titania nanocomposites formed by high energy ball milling. J. Alloy. Compd. **658**, 222–233 (2016)
15. Khalil, K.A., Kim, S.W., Kim, H.Y.: Consolidation and mechanical properties of nanostructured hydroxyapatite-(ZrO_2 + 3 mol% Y_2O_3) bioceramics by high-frequency induction heat sintering. J. Mater. Sci. Eng. A **456**(1–2), 368–372 (2007)
16. Oktar, F.N.: Hydroxyapatite—TiO_2 composites. Mater. Lett. **60**(17–18), 2207–2210 (2006)
17. Mobasherpour, I., Solati Hashjin, M., Razavi Toosi, S.S., Darvishi Kamachali, R.: Effect of the addition ZrO_2-Al_2O_3 on nanocrystalline hydroxyapatite bending strength and fracture toughness. Ceram. Int. **35**(4), 1569–1574 (2009)
18. Shin-Ike, M., Tsutsui, J., Tanaka, A., Murayama, S., Fujita, A.: Attempts to improve the strength of sintered lanthanum-containing hydroxyapatite. J. Osaka Odontol. Soc. **52**, 854–861 (1989)
19. Colon, G., Aviles, M.A., Navio, J.A., Sanchez-Soto, P.J.: Thermal behaviour of a TiO_2–ZrO_2 microcomposite prepared by chemical coating. J. Therm. Anal. Calorim. **67**(1), 229–238 (2002)
20. Miao, X., Sun, D., Hoo, P.W., Lu, J., Hu, Y., Chen, Y.: Effect of titania addition on yttria-stabilised tetragonal zirconia ceramics sintered at high temperatures. Ceram. Int. **30**(6), 1041–1047 (2004)
21. Standard Test Methods for Determining Average Grain size Designation: E112–96 (ASTM International)
22. Salman, S., Gunduz, O., Yilmaz, S., Ovecoglu, Robert, M.L., Snyder, S. Agathapoulos, Oktar, F.N.: Sintering effect on mechanical properties of composites of natural hydroxyapatites and titanium. Ceram. Int. **35**(7), 2965–2971 (2009)

Chapter 57
Parametric Optimization of Corrosion Resistance of Electroless Ni–Co–P Coating

Subhasish Sarkar, Rishav Kumar Baranwal, Ishita Koley, Rupam Mandal, Tapendu Mandal and Gautam Majumdar

Abstract This research paper investigates the effect of cobalt on the corrosion resistance of electroless Ni–P coating by forming a ternary alloy with Ni–P. The corrosion resistance offered by the ternary coating was measured for coating formed by varying the composition of the bath. The varying bath parameters were the concentration of cobalt source (cobalt sulphate) (A), concentration of reducing agent (sodium hypophosphite) (B) and the temperature (C) of the bath. The corrosion resistance was measured using electrochemical impedance spectroscopy (EIS). Taguchi's method has been applied with corrosion resistance as response operator to find the optimum resistance to corrosion and the optimum bath parameters, and along with it, ANOVA is done to find significant interactions between factors. The scanning electron microscopy (SEM) and energy-dispersive X-ray spectroscopy (EDAX) were performed to find the surface morphology and surface composition of the optimized coating.

Keywords Taguchi · Corrosion resistance · ANOVA · Surface morphology

57.1 Introduction

Coatings in the recent time have gathered attention of many researchers as they enhance and introduce different surface properties. Their increased application in the industry [1] is one of the main reasons for it. Coatings can be done by various techniques like electrodeposition, physical vapour deposition, chemical vapour deposition and electroless coating. We have used the electroless technique as it is advantageous. Ni–P coatings are better known for their corrosion resistance [2]. The mechanical and tribological properties can be improved with the help of various heat

S. Sarkar · R. K. Baranwal (✉) · R. Mandal · G. Majumdar
Department of Mechanical Engineering, Jadavpur University, Kolkata 700032, India
e-mail: rishavbaranwal22@gmail.com

I. Koley · T. Mandal
Department of Materials and Mechanical Engineering, IIEST, Howrah 711103, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_57

treatment techniques. However, this can also be achieved with the help of using different elements and complexes. Corrosion is such a phenomenon which degrades the life of metals and causes unpredictable failure in various applications of such metals. Hence, it is very essential to improve the corrosion resistance of the metals or superimpose a coating which has ameliorated anticorrosion properties. This coating has a sacrificial nature. So, in order to increase the corrosion resistance, the introduction of a third element/complex has been preferred [3–5]. The use of cobalt in binary coatings has provided better microhardness compared to the binary coatings [12]. Cobalt introduces a new property of electromagnetic shielding [6, 7] in Ni–P coatings while increasing the corrosion resistance of the coating at the same time.

A number of optimization techniques have been used in the past to optimize the coating [8–11]. We have applied the Taguchi L_{27} orthogonal array design to optimize the corrosion resistance. The Taguchi analysis showed that the level 3 of A, level 3 of B and level 1 of C were the optimized parameters. Analysis of variance (ANOVA) has been applied to find the significant factors and the significant interactions which affect the corrosion resistance. It has been found that all the interactions between the factors are significant for the corrosion resistances. The corrosion resistance of the coatings was measured using EIS. The optimized corrosion resistances were 24.525 and 476.17 Ω cm^2. The EDAX analysis showed the weight percentage of cobalt at 10.90% which is just enough to improve the hardness and induce electromagnetic shielding. The nodular grains present in the coatings were shown by SEM analysis.

57.2 Experimental Methods

The coating took place over copper substrates of thickness 0.1 mm and dimensions 2 × 1.5 mm^2. The pre-treatment of the substrate surface was done by acid cleaning process in 25% HCl for 90 s. The surface of the substrate was activated by using palladium chloride solution at 55 °C for 10 s. It was then rinsed with distilled water, and then, it was immersed in the electroless bath.

The electroless bath consisted of nickel sulphate hexahydrate, cobalt sulphate, sodium hypophosphate, tri-sodium citrate dehydrate and ammonium sulphate. Their respective concentrations are given in Table 57.1.

Table 57.1 Bath composition

Bath constituents	Reagents	(g/L)
Nickel source	Nickel sulphate hexahydrate	5
Cobalt source	Cobalt sulphate	10/15/20
Reducing agent	Sodium hypophosphate	20/25/30
Complexing agent	Tri-sodium citrate dehydrate	15
Buffering agent	Ammonium sulphate	10

The temperature of the bath was varied between 80, 85 and 90 °C. The bath volume was 200 ml, and the deposition time was 60 min.

Cobalt sulphate and sodium hypophosphate control the percentage of cobalt and phosphorus in the coating, respectively, while temperature controls the rate of deposition of the coating.

To study the electrochemical behaviour or corrosion behaviour, standard electrochemical potentiodynamic polarization tests of the above specimens were performed using three electrode corrosion cells in 3.5% Na Cl solution (3.5 g of NaCl in 100 ml double-distilled water). Polarization experiments were carried out as per ASTM standard methods (ASTM G5: Potentiostatic and Potentiodynamic Anodic Polarization Measurements and ASTM G59: Polarization Resistance Measurements) using OrigaLys Potentiostat (software: OrigaMaster 5) with a voltage range from −900 to +600 mV and a scan rate of 1 mV/sec. Electrochemical impedance spectroscopy (EIS) was done to know the surface morphology of the effected region after potentiocyclic voltammetry testing. The frequency range for EIS was from 100 kHz to 100 Hz with initial voltage of 10 mV. The corrosion resistances of each coating are shown in Table 57.2. The corrosion resistance was calculated using the Randles circuit, which forms a double-layer capacitance at the interaction of the coatings'

Table 57.2 Three levels of three factors and their respective resistances

S. No.	$CoSO_4$ (g/L)	$NaHPO_2$ (g/L)	Temperature (°C)	Resistance (Ω cm^2)
1	10	20	80	442.65
2	10	20	85	223.24
3	10	20	90	977.51
4	10	25	80	1309.9
5	10	25	85	49.969
6	10	25	90	1198.5
7	10	30	80	476.17
8	10	30	85	180.31
9	10	30	90	10.115
10	15	20	80	873.60
11	15	20	85	49.079
12	15	20	90	576.79
13	15	25	80	236.67
14	15	25	85	159.60
15	15	25	90	6.74
16	15	30	80	246.49
17	15	30	85	655.58
18	15	30	90	664.24
19	20	20	80	265.12
20	20	20	85	322.48
21	20	20	90	21.9
22	20	25	80	196.72
23	20	25	85	140.15

(continued)

Table 57.2 (continued)

S. No.	$CoSO_4$ (g/L)	$NaHPO_2$ (g/L)	Temperature (°C)	Resistance (Ω cm^2)
24	20	25	90	133.87
25	20	30	80	233.98
26	20	30	85	876.79
27	20	30	90	344.27

surface and the corroding medium. The two layers are formed due to charge transfer across the two different phases at the interaction point.

Design of experiments is an effective tool to determine the optimum conditions. The Taguchi method uses 'larger the better' technique in the present case to optimize the coatings' response. The varying parameters in the bath are called factors, while the different values of a parameter are called levels. We used three factors, and each factor had three levels. The Taguchi method uses the signal-to-noise ratio for finding the optimum conditions for optimum resistance. More positive value of S/N ratio means lesser noise which in turn indicates the best level in that factor. With the help of ANOVA, we can identify the significant parameters and the significant interactions. This is done by interpreting the interaction plots between AxB, BxC and AxC.

The surface morphology and its composition were determined by SEM and EDAX, respectively.

57.3 Results

Taguchi Analysis:

The larger the better uses the following formula for calculating the S/N ratio:

$$\frac{\mathrm{S}}{\mathrm{N}} = -10\log_{10}\left[1/n\left[\sum\frac{1}{y^2}\right]\right]$$

The S/N ratio is the ratio of the desired signal to the undesired noise, which is calculated using the above-mentioned formula. If for a particular level of a factor, signal is high with lower noise; it will indicate a higher S/N ratio which is desired.

Using this concept, the plots of S/N ratio for different levels of each factor are shown in Fig. 57.1.

The average S/N ratio for each factor is shown by the dotted lines in Fig. 57.1.

From Fig. 57.1, we can see the most positive values are achieved for level 3 of A (cobalt sulphate), level 3 for B (sodium hypophosphate) and level 1 for C (temperature of bath). Thus, the optimized coating is A3B3C1, which is the 24th experiment having resistance as 233.98 Ω cm^2.

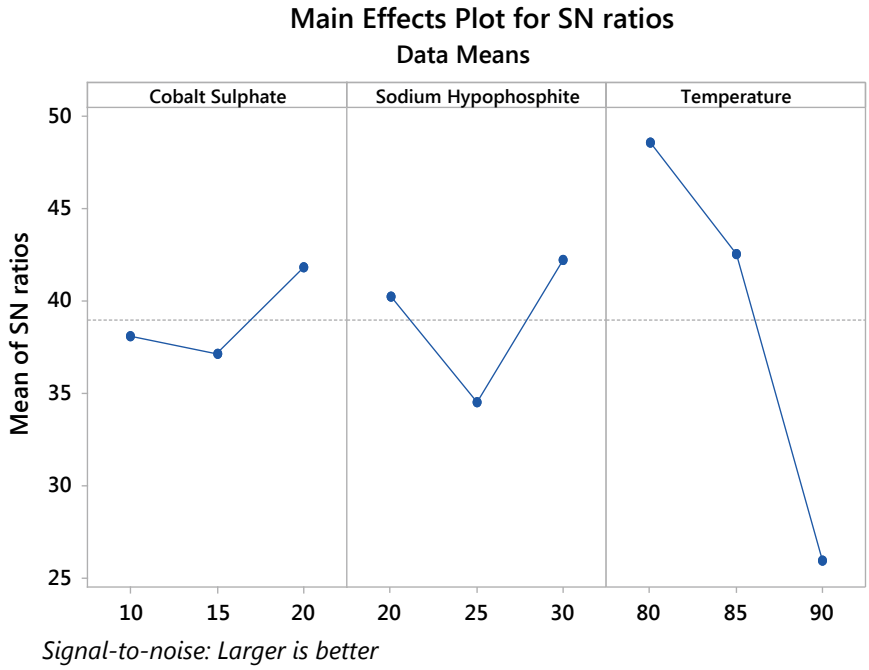


Fig. 57.1 S/N ratio of three levels of each parameter

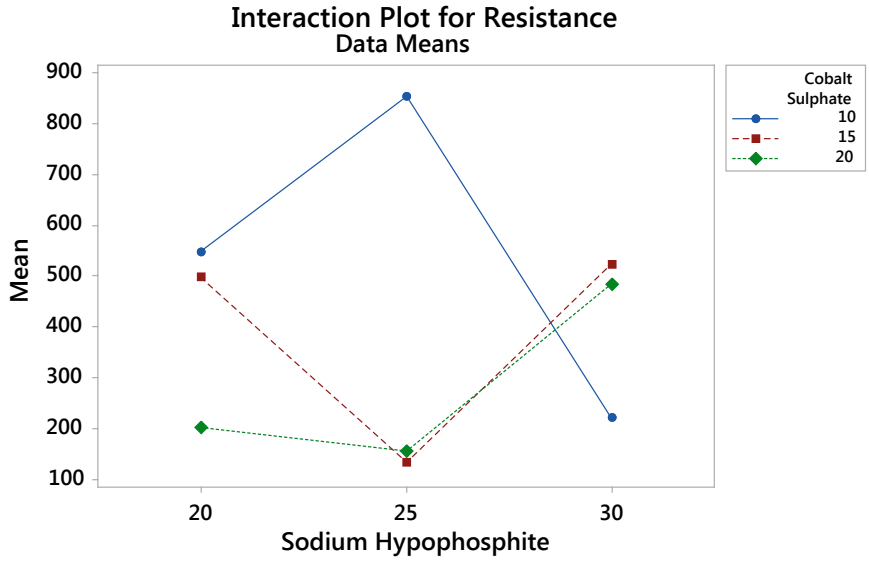


Fig. 57.2 Interactions between cobalt sulphate and sodium hypophosphite

The ANOVA results for R1 and R2 can be interpreted from Figs. 57.2, 57.3 and 57.4.

The blue, red and green lines represent the levels 1, 2 and 3, respectively, of the factors. If these lines remain horizontal or parallel to each other, then insignificant interactions between the factors have taken place. On the other hand, if the lines are inclined to each other and intersect at many points then good interactions have taken place between the factors. So for resistance offered to corrosion, we can say clearly, by observing the graphs that good interactions exist between A, B and C.

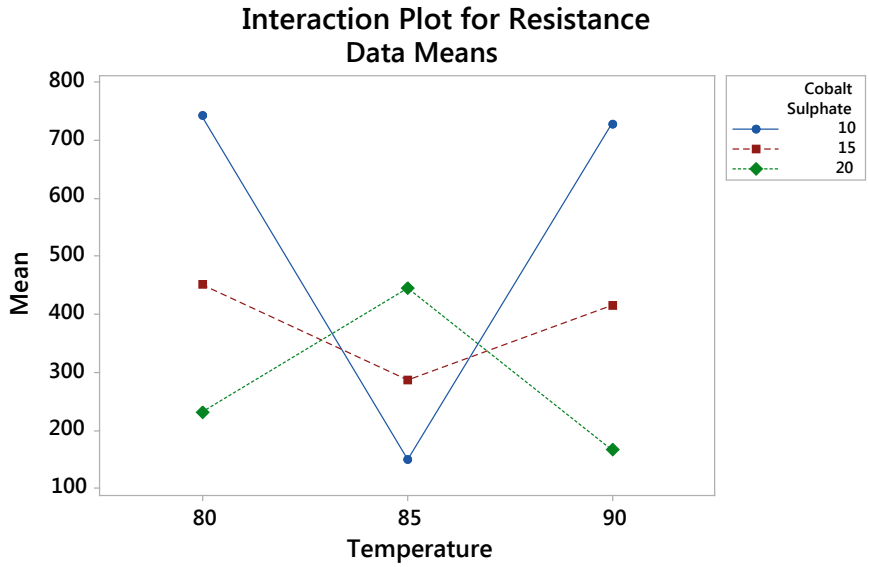


Fig. 57.3 Interactions between cobalt sulphate and temperature

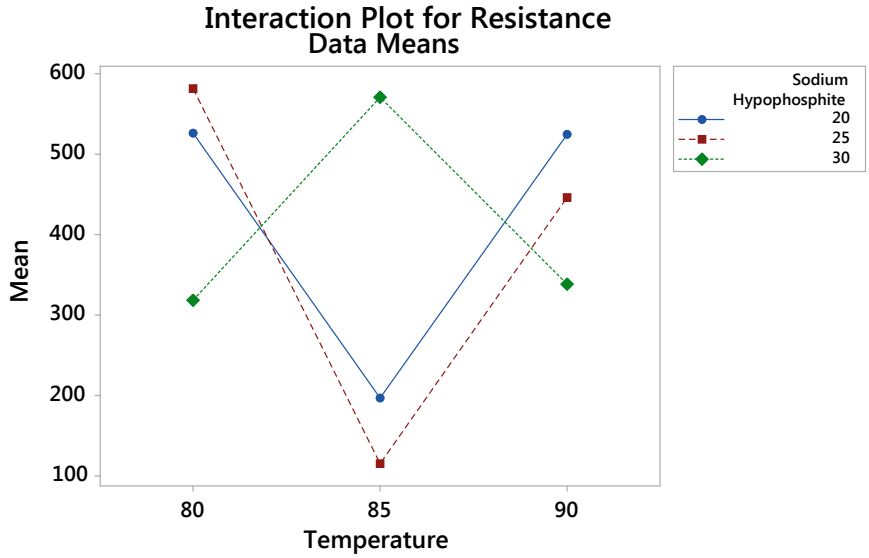


Fig. 57.4 Interactions between temperature and sodium hypophosphite

The energy-dispersive X-ray analysis shows the percentage of different elements on the surface of the coatings (Fig. 57.5). Its results are shown in Table 57.3. The cobalt % is 10.9% which is very influential in improving the various properties [12] and in this case corrosion resistance.

The composition is as shown.

Scanning electron microscopy was done to study the deposition on the surface of copper. Figure 57.6 shows the homogeneous deposition with very small agglomerated materials as globules. Since the surface is fully intact with coating. Hence, the porosity of the coating is very low. With low porosity, the corrosion agents which when gets deposited on the surface are unable to penetrate the surface of the coating which allows the coating to have higher corrosion resistance.

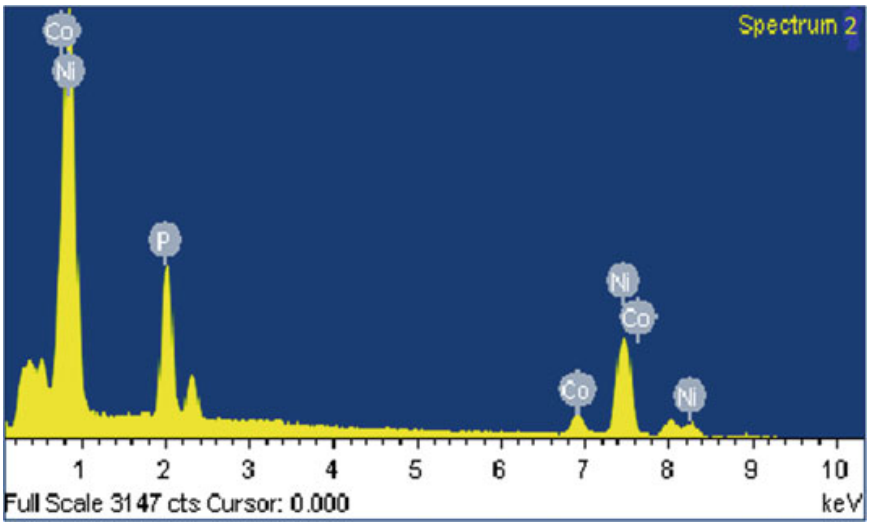


Fig. 57.5 EDAX image of the optimized sample

Table 57.3 EDAX result

Element	Weight %	Atomic %
Ni	76.89	69.34
Co	10.90	9.79
P	12.21	20.87

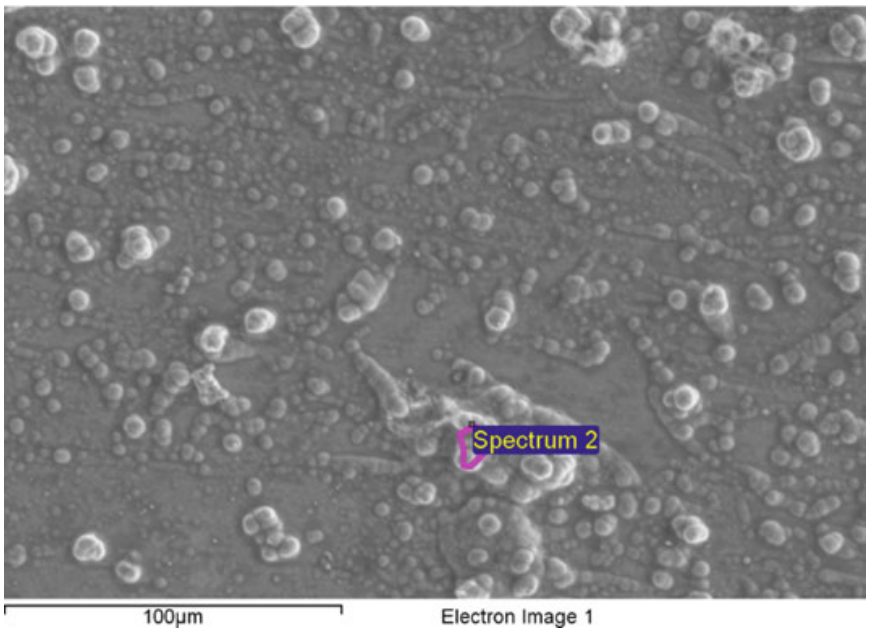


Fig. 57.6 SEM image of the coating

57.4 Conclusion

The electroless coating was optimized using Taguchi's method, and the optimized conditions are as follows: cobalt sulphate—10 g/L, sodium hypophosphite—30 g/L and temperature of bath at 80 °C. The increased resistance of the coating will have sacrificial behaviour. ANOVA results show that the interactions between each level (3 levels) of different factors (cobalt sulphate, nickel sulphate and temperature) are significant in determining the corrosion resistance of the coating. SEM results reveal intact coating taking place which allows the corrosion resistance to increase by hindering the movement of corroding materials into the substrate. The EDAX results show the weight percentage of cobalt at 10.9% which is enough to increase the corrosion resistance of the substrate material.

Acknowledgements We would like to thank Professor Deb Dulal Das of IIEST Shibpur, for the SEM and EDAX analysis.

References

1. Baudrand, D., Durkin, B.: Metal Finishing, pp. 20–24 (1998)
2. Narayanan, T.S.N.S., Baskaran, I., Krishnaveni, K., Parthiban, S.: Surf. Coat. Technol. **200**, 3438–3445 (2006)
3. Song, Y.-W., Shan, D.-Y., Chen, R.-S., Han, E.-H.: Surf. Eng. **23**(5), 334–338 (2007)
4. Rezagholizadeh, M., Ghaderi, M., Heidary, A., Vaghefi, S.M.M.: Prot. Metals Phys. Chem. Surf. **51**(2), 234–239 (2015)
5. Jiaqiang, G., Lei, L., Yating, W., Bin, S., Wenbin, H.: Surf. Coat. Technol. **200**, 5836–5842 (2006)
6. Narayanan, T.S.N.S., Selvakumar, S., Stephen, A.: Surf. Coat. Technol. **172**, 298–307 (2003)
7. Gao, Y., Huang, L., Zheng, Z.J., Li, H., Zhu, M.: Appl. Surf. Sci. **253**, 9470–9475 (2007)
8. Oraon, B., Majumdar, G., Ghosh, B.: Mater. Des. **29**, 1412–1418 (2008)
9. Oraon, B., Majumdar, G., Ghosh, B.: Mater. Des. **28**(7), 2139–2147 (2007)
10. Sahoo, P.: Mater. Des. **30**, 1341–1349 (2009)
11. Basavaraja, D.K., Skanda, M.G., Soumya, C., Ramesh, V.: Int. J. Mech. Aerosp. Ind. Mechatron. Manuf. Eng. **11**(4), 698–704 (2017)
12. Sarkar, S, Baranwal, R.K., Lamichaney, S., De, J., Majumdar, G.: Optimization of electroless Ni–Co–P coating with hardness as response parameter: a computational approach. Jurnal Tribologi **18**, 81–96 (1 Sept 2018)

Chapter 58
Investigation on Dimensional Accuracy for CFRP Antenna Reflectors Using Autoclave and VARTM Processes

Tushar Gajjar, Dhaval Shah, Shashikant Joshi and Kaushik Patel

Abstract The carbon fibre reinforcement polymer (CFRP) is widely used in space applications, automobile industries, aerospace and sports equipment because of good specific properties over conventional metals and alloys. Autoclave processing is one of the best manufacturing techniques to achieve high-dimensional accuracy for complex-shaped component. But, due to high initial and processing cost, vacuum-assisted resin transfer moulding (VARTM) is the alternative manufacturing method. In this research paper, CFRP antenna reflectors have been manufactured using autoclave process and VARTM manufacturing process to evaluate dimensional accuracy. The part thickness of laminate and layup orientation has been considered as a processing parameter in reflector manufacturing. The dimensions of the inner surface of the manufactured reflectors have been measured using 3D scanning techniques at the different locations. The deviation between theoretical CAD model and reflector dimensions at each location provides dimensional inaccuracy for the fabricated reflectors. The investigation on the dimensional inaccuracy of cured reflectors manufactured using the autoclave and VARTM manufacturing techniques has been carried out and the percentage difference has been found within the acceptable limit.

Keywords Dimensional accuracy · Autoclave · VARTM · Antenna reflector · Spring-back deformation

T. Gajjar
Mechanical Engineering Department, Sardar Vallabhbhai National Institute of Technology, Surat 395007, India
e-mail: tushar.gajjar79@gmail.com

D. Shah (✉) · S. Joshi · K. Patel
Mechanical Engineering Department, Institute of Technology, Nirma University, Ahmedabad 382481, India
e-mail: dbshah@nirmauni.ac.i

S. Joshi
e-mail: s.j.joshi@nirmauni.ac.in

K. Patel
e-mail: kaushik.patel@nirmauni.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_58

58.1 Introduction

The composite material is a combination of two or more materials which are widely used in aerospace industries, space applications, automobiles and sports industries due to its high strength-to-weight ratio [1]. A prepreg is a pre-impregnated fibre and available as unidirectional, bi-directional and twill fabric tape [2]. The autoclave curing process is widely used for making complex-shaped composite components because of high-dimensional accuracy. But, due to high capital as well as processing cost, VARTM process is the alternative manufacturing method. The variation between cured laminated composite part dimensions and theoretical dimension is usually referred as spring-back deformation or warpage depending on the shape of components.

Many researchers worked on the prediction of the spring-back deformation and warpage for the composite products. Darrow and Smith [3] occupied three process parameters—thicknesses cure shrinkage, fibre volume fraction gradients and mould expansion that effects on spring-back deformation. Fernlund and Poursartip [4] exhibited the effect of the cure cycle and tool surface condition on spring-back deformation. Kaushik et al. [5] calculated the effectiveness of various parameters like pressure, degree of cure and ramp rate on the coefficient of friction in the autoclave process. Kappel et al. [6] measured the effect of crucial parameters such as part thickness, layup, part radius and scattering of occurring distortion during the autoclave manufacturing process. Correia et al. [7] developed an analytical formulation of governing equation for incompressible flow and this model successively used to measure the effect of process parameters such as inlet and outlet pressure, fibre architecture and layup orientation. Govignon et al. [8] performed experimentation on thickness variation during compaction to replicate reinforcement behaviour while pre-filling, filling and post-filling stage. Hammami et al. [9] studied the effects of the processing variables in vacuum infusion moulding process by developing the one-dimesional model. The compaction test was conducted with the maximum compaction pressure not beyond 0.1 MPa. It was established that for low compaction pressures (≤ 0.1 MPa), the number of layers did not have a significant effect on the compaction behaviour of the preform. Arulappan et al. [10] performed experiments on CFRP material with flat plate, *L*-shaped plate and the flat plate with a central circular hole. The pressure was varried for different shape. The flow of resin depends on flow medium. For large structure, high permeability medium (HPM) was selected for getting high flow rate and reduced filing time. Song et al. [11] developed an experimental model for incorporating resin flow through the preform as well as relaxation of the preform because resin pressure and part thickness variation were parameters in VARTM which play a vital role in manufacturing component.

In this research paper, autoclave and VARTM manufacturing processes have been used to predict dimensional accuracy of CFRP antenna reflectors. The AISI 430 grade stainless steel material has been chosen for manufacturing of parabolic mould. The thickness of the laminate and layup orientation has been considered as

process parameters. The Hinpreg A45 CFRP prepreg is used in autoclave process, whereas carbon fibre dry fabric along with epoxy resin 520 and D-type hardener is used in the VARTM process. The parabolic-shaped reflectors have been cured using autoclave and VARTM processes. The dimensions at the inner surface of manufactured CFRP reflectors have been measured using 3D scanning techniques at different 200 locations. The deviation between the theoretical CAD model and reflector dimensions at each location gives dimensional inaccuracy or spring-back deformation for the manufactured CFRP reflectors. The comparison of spring-back deformation of cured reflectors manufactured using autoclave and VARTM manufacturing techniques has been carried out. The reflectors with different thickness and layup sequences have been fabricated using both manufacturing techniques, and dimensional accuracy has been matched.

58.2 Experimental Preparation

The stainless steel AISI 430 grade has been identified as mould material with 300 × 300 × 70 mm size. The 3D model of the parabolic mould has been prepared using solid modelling software. The dimension of the parabolic profile at one side is 240 mm circumferential diameter with 50 mm depth. The stainless steel mould block has been machined using NC code in 5–axis vertical milling centre with very high surface finish as shown in Fig. 58.1.

58.2.1 Reflectors Manufacturing Using Autoclave Process

The Hinpreg A45 unidirectional prepreg tape has been selected for the autoclave experimental process. Layers of prepreg are cut into the round shape of 300 mm diameter depending upon the required number of laminates and thickness as shown in Fig. 58.2. The consolidation materials like release agent, peel ply, breather cloth, vacuum bagging, etc. have been used in the autoclave process. In the autoclave process, the entire release agent has been applied on the parabolic mould surface.

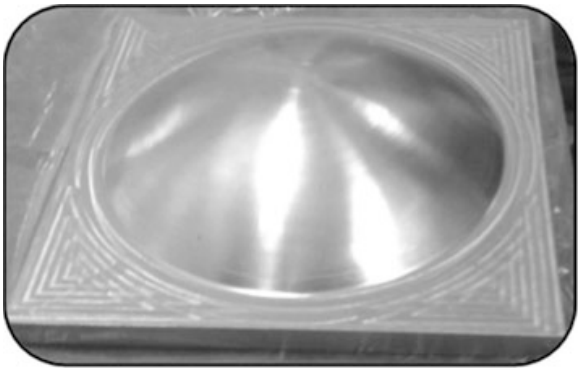


Fig. 58.1 Manufactured stainless steel mould

Fig. 58.2 Unidirectional cut prepreg

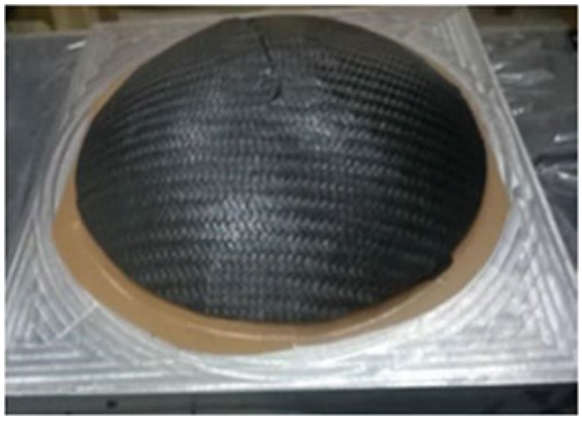


Fig. 58.3 Layers of CFRP prepreg on steel mould

The arrangement of prepreg layers with proper orientation is shown in Fig. 58.3. Peel ply permits free passage of volatiles and excess matrix while curing. To apply the vacuum and guide the removal of air and volatiles from the assembly, the breather fabric has been provided. Thicker breathers are required during great autoclave pressures. Sealant tape and vacuum bag are used for sealing the entire assembly. The whole set-up has been vacuum bagged with the help of vacuum film, vacuum valve and hose pipe as shown in Fig. 58.4.

Fig. 58.4 Vacuum bagging and air suction process

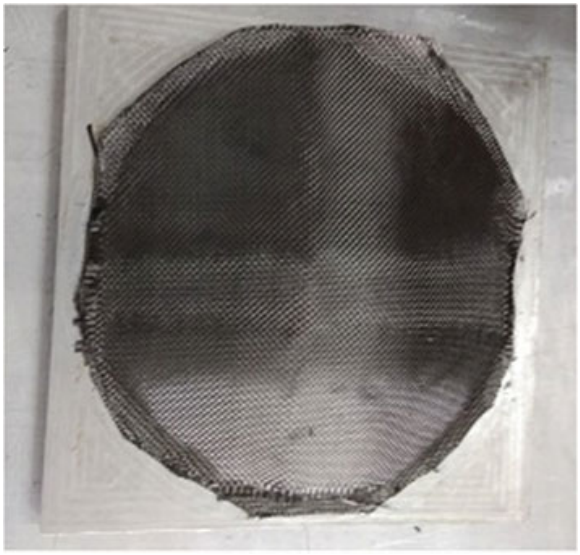


Fig. 58.5 Four layers of carbon fibre

58.2.2 Reflectors Manufacturing Using VARTM Process

The parabolic shape reflectors are also manufactured using the VARTM process. The dry carbon fabric has been selected as a reinforcement and epoxy resin 520 and D-type hardener as a matrix material in the VARTM experimental process. The epoxy resin 520 having mix viscosity at 27 °C of 650 ± 100 MPas, the specific gravity of 1.15–1.18 g/cm^3 and pot life at 27 °C of 1.5–2 h has been considered. Similarly, D-type hardener consists of tensile strength of 70 MPa, the flexural strength of 120 MPa and impact strength of 80 N/m is mixed with one-fourth of resin ratio. The vacuum pump, infusion mesh, purported plastic, catch pot, etc. have been used for VARTM process. The layer of the release agent is applied on the paraboloid surface of the mould for ease of demoulding. The four layers of carbon fibre have been arranged on the mould surface in the proper sequence as shown in Fig. 58.5. Similarly, peel ply, purported plastic and infusion mesh are placed on fibre layers simultaneously and the whole assembly is sealed using sealant tape and vacuum bag as shown in Fig. 58.6. The PVC hose pipe is attached at resin inlet and another side for obtaining vacuum pressure. The resin is sucked into the system because of vacuum pressure as shown in Fig. 58.7.

58.2.3 Curing of CFRP Reflectors

In the autoclave process, the entire set-up—mould, composite prepreg along with consolidation materials, is kept in autoclave machine for curing purpose as shown in Fig. 58.8. The mould along with consolidate material has been cured for twenty-four hours at an atmospheric temperature in the VARTM process as shown in Fig. 58.9. The atmospheric pressure of 1 bar in VARTM, whereas 5 bar pressure in autoclave is retained machine throughout the curing cycle.

Fig. 58.6 Vacuum bag along with sealant tape

Fig. 58.7 Resin inlet hose and vacuum hose

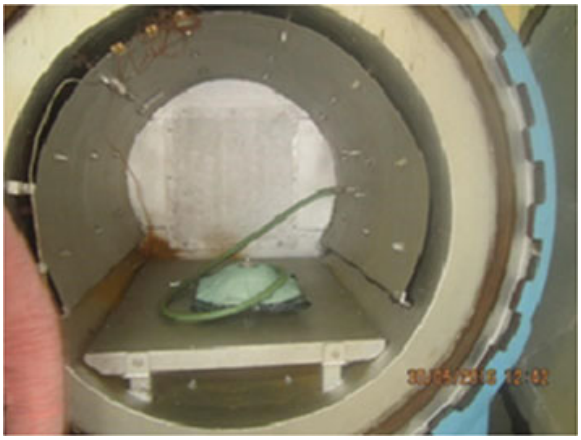


Fig. 58.8 Curing in autoclave machine

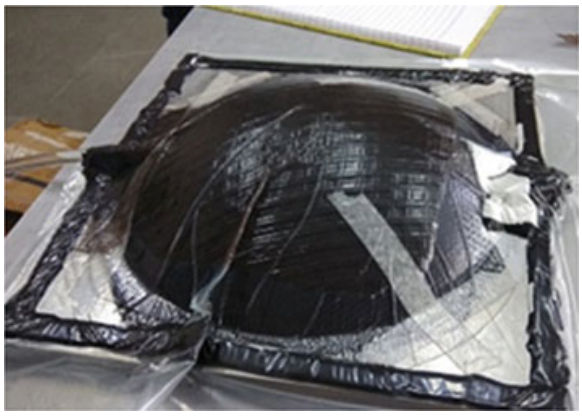


Fig. 58.9 Curing in VARTM process

Fig. 58.10 Autoclave demoulding process and cured reflector

Fig. 58.11 VARTM demoulding process and cured reflector

The entire assembly is taken out from the autoclave machine after the curing process and consolidation materials are detached from the assembly at room temperature as shown in Fig. 58.10. The sequence of demoulding process for VARTM process and final cured reflector is shown in Fig. 58.11.

58.3 Dimensional Measurement Using 3D Scanning

The inner surface of reflectors is selected for dimension measurement and 200 number of coordinate points are identified on this surface. These points are obtained projecting circles of diameter 50, 100, 150 and 240 mm on parabolic surface of the reflector. Measurement of specified points is performed to ensure the accuracy of the parabolic reflector using the 3D scanner. This scanned model of the composite reflector is superimposed with the theoretical CAD model for the same to achieve deviation on each prescribed coordinate points. The deviation at each coordinate point is given using contour plot as shown in Fig. 58.12. The root means square of all 200 points deviations is referred as spring-back deformation for that reflector. The theoretical X, Y and Z coordinates and scanned X, Y and Z coordinates are given in Table 58.1 for autoclave curing, whereas Table 58.2 for VARTM curing process.

The spring-back deformation for the reflector manufactured using autoclave manufacturing process is 0.37 mm, whereas for the reflector manufactured using VARTM process is 0.35 mm. This value has been obtained for the same configuration reflector in each case. The difference between the autoclave manufacturing process and VARTM manufacturing process is around 5%. The different antenna reflectors have been manufactured with different laminate thickness and layup orientations. The 4, 8 and 12 number of layers and unidirectional, cross- and quasi-isotropic layup orientations have been considered. The deviation, i.e. dimensional inaccuracy, is observed and root mean square value is compared from VARTM and autoclave curing process for each case.

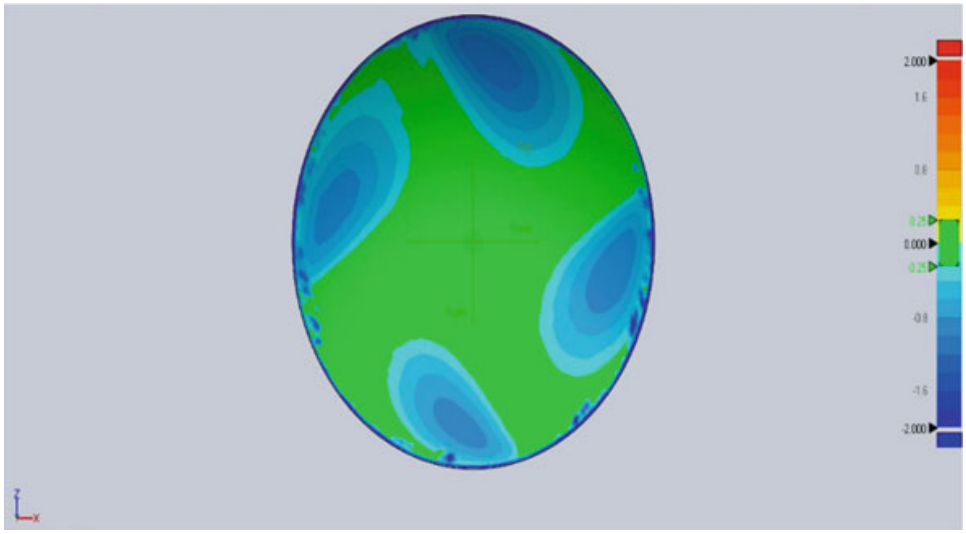


Fig. 58.12 Deviation contour of reflector at prescribed coordinate points

Table 58.1 Deviation of each coordinate point in autoclave curing process

Sr. No.	Predefined coordinate (Nominal) points (mm)			Scanned coordinate points (mm)			Deviation (mm)
1	0	−2.003	25	0	−2.051	24.992	0.04866
2	−3.911	−2.003	24.692	−3.911	−2.007	24.691	0.00412
3	−7.725	−2.003	23.776	−7.726	−1.978	23.78	0.02534
4	−11.35	−2.003	22.275	−11.353	−1.957	22.281	0.04649
5	−14.694	−2.003	20.225	−14.7	−1.947	20.232	0.05675
6	−17.677	−2.003	17.677	−17.683	−1.948	17.683	0.05565
.							
.							
.							
200	18.772	−46.083	118.521	18.728	−46.436	118.253	0.44524
Min	−120	−46.084	−120	−120.303	−47.054	−119.51	0.001
Max	120	0	120	120.454	0.0057	119.784	1.225
RMS							**0.37**

Table 58.2 Deviation of each coordinate point in VARTM curing process

Sr. No.	Predefined coordinate (Nominal) points (mm)			Scanned coordinate points (mm)			Deviation (mm)
1	25	−2.003	0	25	−1.998	0	−0.005
2	24.692	−2.003	3.911	24.693	−1.998	3.911	−0.005
3	23.776	−2.003	7.725	23.777	−1.998	7.726	−0.005
4	22.275	−2.003	11.35	22.277	−1.988	11.351	−0.015
5	20.225	−2.003	14.694	20.231	−1.957	14.699	−0.046
6	17.677	−2.003	17.677	17.683	−1.955	17.683	−0.049
.							
.							
.							
200	18.772	−46.083	118.521	18.831	−45.612	118.877	−0.594
Min	−119.99	−46.084	−119.99	−120.217	−46.265	−120.3	−1.168
Max	119.999	0	119.997	120.258	0.023	120.441	0.229
RMS							**0.35**

58.4 Conclusions

The CFRP antenna reflectors have been manufactured using autoclave and VARTM manufacturing processes with stainless steel AISI 430 mould. Hinpreg A45 uni-directional prepreg tape is used in the autoclave and bi-axial dry fabric with epoxy 520 resin along with hardener D has been used in the VARTM process. The dimension at the inner surface of CFRP manufactured reflectors has been measured

using 3D scanning techniques at 200 different locations. The comparison of spring-back deformation of cured reflectors manufactured using autoclave manufacturing techniques and VARTM process has been carried out. The percentage deviations in spring-back deformation have been found within 5% range so the VARTM process gives more dimensional accuracy compared to autoclave process and VARTM is the cheapest process compare to autoclave process. The effect of laminate thickness and layup orientation has been considered as a process parameter. As laminate thickness increases, the spring-back deformation decreases. The quasi-isotropic layup orientation has minimum spring-back deformation.

References

1. Mazumdar, S.K.: Composites Manufacturing Materials, Product, and Process Engineering. CRC Press
2. White, S.R., Hahn, H.T.: Process modeling of composite materials: residual stress development during cure. Part I. Model formulation. J. Compos. Mater. (1992)
3. Darrow, D.A., Smith, L.V.: Isolating components of processing induced warpage in laminated composites. J. Compos. Mater. (2002)
4. Fernlund, G., Poursartip, A.: The effect of tooling material, cure cycle, and tool surface finish on spring-in of autoclave processed curved composite parts. In: Proceeding of ICCM (2003)
5. Kaushik, V., Raghavan, J., Zeng, X.: Experimental study of tool—part interaction during autoclave processing of thermoset polymer composite structures. Compos.: Part A **41**, 1210–1218 (2010)
6. Kappel, E., Stefaniak, D., Hühne, C.: Process distortions in prepreg manufacturing—an experimental study on CFRP L-profiles. Compos. Struct. **106**, 615–625 (2013)
7. Correia, N.C., Robitaille, F., Long, A.C., Rudd, C.D., Simacek, P., Advani, S.G.: Analysis of the vacuum infusion moulding process: I. Analytical formulation. Compos. Part A: Appl. Sci. Manuf. **36**(12), 1645–1656. (2005)
8. Govignon, Q., Simon, B., Kelly, P.A.: Simulation of the reinforcement compaction and resin flow during the complete resin infusion process. Compos. A Appl. Sci. Manuf. **41**(1), 45–57 (2010)
9. Hammami, A., Gebart, B.R.: Analysis of the vacuum infusion molding process. Polym. Compos. **21**(1), 28–40 (2000)
10. Arulappan, C., Duraisamy, A., Adhikari, D., Gururaja, S.: Investigations on pressure and thickness profiles in carbon fiber-reinforced polymers during vacuum assisted resin transfer moulding. J. Reinf. Plast. Compos. **34**(1), 3–18 (2015)
11. Song, X., Loos, A.C., Grimsley, B., Cano, R., Hubert, P.: Simulation of the vacuum assisted resin transfer moulding process. In: International Conference on Flo Process in Composite Materials, Newark, Delaware (2004)

Chapter 59
On Finite Element Analysis of Melting of Metallic Powders Using Microwave Energy at 2.45 GHz

Mohit Kumar, Gaurav Kumar and Vikas Kukshal

Abstract Use of microwave energy for processing of materials is gaining vast popularity due to its inherent superior characteristics like reduced processing time, better control on microstructure and enhanced mechanical properties. In this paper, the interaction of the microwave with metallic powder is analyzed using COMSOL Multiphysics software. Further, the effect of particle size on microwave absorption is also been analyzed. Microwave hybrid heating (MHH) technique is used to heat the metallic powder. Initially, metallic powder gets heated conventionally, i.e., through the heat transfer with separator plate and susceptor material. Once the metallic powder reaches the critical temperature, metallic powder starts to absorb the microwave and volumetric heating takes place. Also, the effects of input power on electric field strength and melting time of charge were analyzed. Further, it was found that as the size of the metallic powder particle is reduced, the time required for melting and casting of metallic powder reduces.

Keywords Microwave (MW) · Metallic powder · Temperature

Nomenclature

T	Temperature, K
t	Time of exposure, s
E	Time-harmonic electric field, V m^{-1}
H	Time-harmonic magnetic field, A m^{-1}

M. Kumar · G. Kumar (✉) · V. Kukshal
National Institute of Technology, Uttarakhand, Srinagar Garhwal 246174, Uttarakhand, India
e-mail: grv.kmr2015@gmail.com

M. Kumar
e-mail: mohitchauhan657@gmail.com

V. Kukshal
e-mail: vikaskukshal@nituk.ac.in

G. Kumar
Microwave Material Processing Laboratory, IIT Roorkee, Roorkee 247667, Uttarakhand, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_59

C_p Specific heat of material, $J\,kg^{-1}\,K^{-1}$
h Convection heat transfer coefficient, $W\ m^{-2}\,k^{-1}$
n Unit vector perpendicular to the surface of metal boundary
k Thermal conductivity, $W\ m^{-2}\,K^{-1}$
ε Permittivity, $F\,m^{-1}$
ε_r Relative permittivity
ε_0 Permittivity in free space, $F\,m^{-1}$
ε^* Complex electric permittivity, $F\,m^{-1}$
ε' Dielectric constant, $F\,m^{-1}$
ε'' Dielectric loss factor, $F\,m^{-1}$
μ Magnetic permeability, $H\,m^{-1}$
μ_0 Magnetic permeability in free space, $H\,m^{-1}$
μ_r Relative permeability, $H\,m^{-1}$
μ'' Magnetic loss factor, $H\,m^{-1}$
ω Angular frequency, s^{-1}
σ Electrical conductivity, $S\,m^{-1}$
σ_m Static electrical conductivity of metal, $S\,m^{-1}$
c_0 Speed of light in free space, $m\,s^{-1}$
γ Collision frequency of electrons, s^{-1}
β Coupling parameter
ρ Density of material $kg\,m^{-3}$
$\partial\Omega$ Small volume, m^{-3}
ε_{rad} Emissivity
k_0 Wave number in free space

59.1 Introduction

The casting process is used to produce products with a wide range of variety in terms of complex shape, size and order quantity [1]. It is the most efficient manufacturing process to overcome complexity. Nowadays, cost, power and time saving are critical issues in the casting process. Microwave heating is a novel process which has a wide range of application in manufacturing processes like sintering, surface modification, casting, joining, drilling and cladding [2–7]. The past studies focused on the heating of metal powder [8, 9], melting [9] and composite casting with graphite mold [10] using microwave energy at 2.45 GHz. It is not possible to heat bulk metal directly using the microwave as bulk metal acts as a reflector to microwave at low temperature. But a recent study has shown that microwave hybrid heating (MHH) can be used to heat bulk metal and consequently, bulk metal can be melted and further cast in graphite mold [4].

Roy et al. [2] reported the microwave sintering process of metal powders (Fe–Ni and Fe–Cu) at a 2.45 GHz frequency and compared MW processed results with

conventionally processed results. Properties like modulus of rupture (MOR), Rockwell hardness and sintered density of MW processed samples are substantially greater than conventional samples. It was reported that similar heating occurred with or without using susceptor, but the presence of susceptor provided rapid heating. Ghosh et al. [3] examined the characterization of microwave-heated aluminum powder (average particle size 25–30 μm with 99.7% purity) by XRD, SEM, TEM, EDX analysis and Fourier transform infrared (FTIR) spectroscopy. EDX analysis and FTIR studied revealed that Al powder transformed into Al–Al_2O_3 core–shell composite powder after the microwave heating process. SEM indicated that the morphology of microwave-processed Al powder was largely different from unprocessed Al powder [3]. Gupta and Sharma [7] developed nickel-based powder (EWAC, the average particle size of 40 μm) cladding on austenitic stainless steel (SS-316) substrate by using a domestic microwave system (1 kW and 2.45 GHz). The cladding was done successfully with an exposure time of 360 s. It was found that average Viker's micro-indentation hardness (304 ± 48 HV) of microwave clad is twice that of the SS-316 substrate (175 HV).

Chandrasekaran et al. [9] investigated the experimental and theoretical results of microwave melting of metals (Pb, Sn, Al and Cu) at 2.45 GHz and different power levels (520, 910 and 1300 W). Also, they compared the results with conventional heating and found that the energy required in microwave heating was half of the conventional heating (muffle furnace) and also processing time was very less as compared to the conventional heating. Singh et al. [10] reported the microwave melting and casting of metallic powders (EWAC or Ni-based powder + SiC) at 900 W and 2.45 GHz through MHH. It was observed that sample of EWAC + 10% SiC metal–ceramic composite casting is approximately 2.64 times harder than the sample of pure EWAC casting. Mishra and Sharma [4] reported the microwave casting of bulk aluminum alloy (AA-7039) by MHH at 1400 W and 2.45 GHz. The whole process was carried out inside the microwave oven cavity, and a new mechanism of in situ microwave casting was proposed. Four different temperature-dependent phases were also found during the study of time–temperature characteristics. Further, it was reported that the mode of heating in each phase was different. The time–temperature plot was validated using simulated data obtained from COMSOL Multiphysics software [11]. It was also reported that AA-7039 cast developed using in situ microwave casting through MHH showed superior property in the cast with less than 2% porosity and ultimate tensile strength of 148.46 ± 10 MPa [12].

The basic objective of the present work-study is to investigate the simulation result of the melting and casting of AA-7039 powder using MHH. The electromagnetic wave, frequency domain and heat transfer in solids were studied in COMSOL Multiphysics 5.2a software tool. The electric field distribution, temperature distribution and time–temperature profile of metallic powder were determined. The effect of input power on the electric field norm inside the microwave oven and melting time of charged or metallic powder were also discussed. In addition, the results showing the effect of particle size on melting time are also reported.

59.2 Modeling and Simulation

59.2.1 Model Development

A three-dimensional model of microwave (MW) casting of aluminum alloy 7039 powder was developed using COMSOL Multiphysics 5.2a software tool. The dimensions of the developed 3-D model of the microwave oven were taken from commercially available LG domestic microwave, and it operates at a power of 900 W with 2.45 GHz frequency source of a rectangular waveguide working in TE_{10} mode. The microwave oven cavity with dimensions of 335 × 335 × 215 mm^3 (W × D × H) and rectangular waveguide with dimensions of 50 × 43 86 mm^3 (W × D × H) were modeled to simulate the casting process. The powder material (particle radius 0.44 mm) was cast into a graphite mold cavity with dimensions of 24 × 45 × 10 mm^3 (W × D × H), and ceramic brick was used as a base of the setup. In the simulation, the MHH technique was used by taking SiC plate as a separator and SiC powder as a susceptor with a 4-mm thickness of both. The chemical composition of aluminum alloy (AA) 7039 used is shown in Table 59.1, and material properties used in the simulation analysis are shown in Table 59.2.

A simplified 3-D model of MW casting of metallic powder is shown in Fig. 59.1. The inset shows the zoomed view of metallic powder. A single sphere particle was generated with a preferred dimension, and an array operation was performed on the sphere particle. A number of particles were added in three dimensions (*X*-, *Y*- and *Z*-axes) by giving offset distance according to the mold dimensions using array operation.

59.2.2 Assumptions

In order to overcome the complexity and to simplify the model, the following assumptions were made:

(a) The ambient temperature and pressure of the system were assumed to be 27 °C and 1 atm, respectively.
(b) The oxidation of powder material and other chemical changes were not considered.
(c) Susceptor powder was taken as bulk metal (plate) to simplify the study.

Table 59.1 Chemical composition of AA 7039 [13]

Element	Zn	Mg	Mn	Fe	Cr	Si	Cu	Al
% weight	4.12	2.88	0.28	0.22	0.22	0.14	0.01	Balance

Table 59.2 Material properties used in the simulation study [11]

Materials	Density (kg m^{-3})	Specific heat at constant pressure (J kg^{-1} K^{-1})	Thermal conductivity (W m^{-1} K^{-1})	Electric conductivity (S m^{-1})	Permittivity	Magnetic permeability	Emissivity
AA-7039 (powder)	2700	$741.56 + 0.5071 \times T$	$238/(0.000011 \times T^{2.30} + (0.124/T))$	$3.774 \times 10^7/(1 + 0.0043 \times (300.15))$	1	1	0.19
Graphite	2100	830	470	1700	15	1	–
SiC	3210	1000	250	0.02	9.6	1	0.3
Ceramic brick	2000	900	0.5	0.001	1	1	–

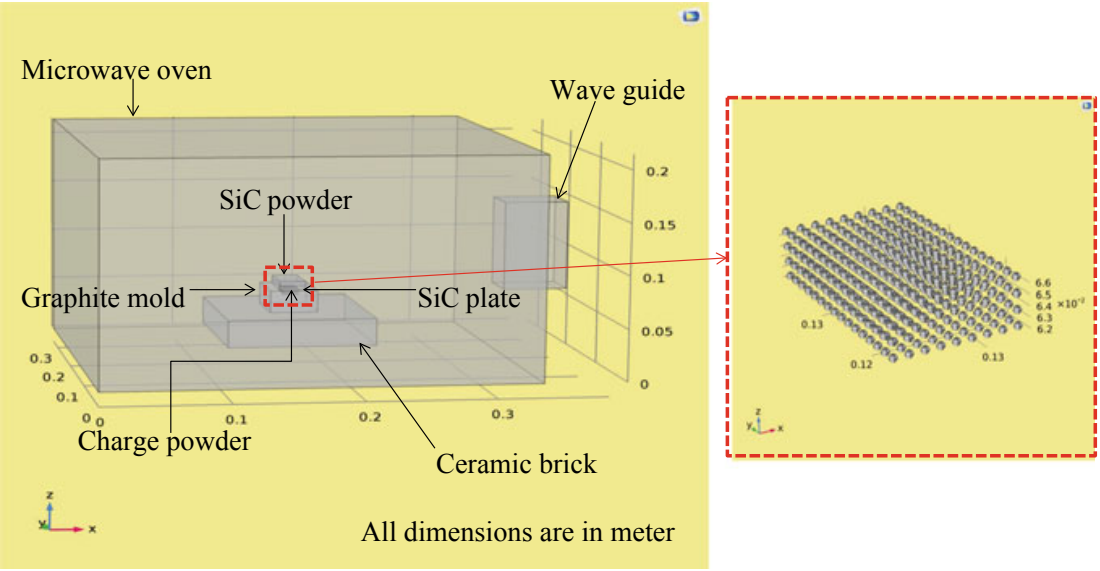


Fig. 59.1 A 3-D model of microwave casting of metallic powder (*inset* zoomed view of powder material)

(d) The wall of the microwave oven and waveguide were considered as copper metal.
(e) The ceramic bricks and air domain were not considered in the thermal analysis.
(f) The rectangular port was excited by transverse electric (TE_{10}) field in *Z*-direction only.

59.2.3 Governing Equations

A Hewlett-Packard (HP) work system (16 GB RAM) was used for simulation. Finite element-based COMSOL Multiphysics software was used to solve the current problem. The software couples the physics of electromagnetic wave in the frequency domain and heat transfer in the transient domain. Maxwell's equations were used to solve the electromagnetic wave of propagation. The electric field (E) inside the microwave oven is given by Eq. (59.1) [11]:

$$\nabla \times \mu_r^{-1}(\nabla \times E) - k_0^2\left(\varepsilon_r - \frac{j\sigma}{\omega\varepsilon_0}\right)E = 0 \tag{59.1}$$

where μ_r is the relative permeability, E is the time-harmonic electric field ($V\,m^{-1}$), ε_r is the relative permittivity, $j = \sqrt{(-1)}$, σ is the electrical conductivity, ω is angular frequency, ε_0 is the permittivity of vacuum, i.e., 8.85×10^{-12} ($F\,m^{-1}$), and k_0 is the wave number in free space which can be formulated in Eq. (59.2) [11]:

$$k_0 = \omega\sqrt{\varepsilon_0\mu_0} = \frac{\omega}{c_0} \tag{59.2}$$

where c_0 is the speed of light in vacuum. The expression of complex electric permittivity (ε^*) of dielectric material is given by Eq. (59.3) [11]:

$$\varepsilon^* = \varepsilon' - j\varepsilon'' \tag{59.3}$$

where ε' is the dielectric constant $(\mathrm{F\,m^{-1}})$ and ε'' is the dielectric loss factor $(\mathrm{F\,m^{-1}})$. The complex electric permittivity $\left(\varepsilon^*_{\mathrm{m}}\right)$ for metallic material can be expressed as in Eq. (59.4) [14]:

$$\varepsilon^*_{\mathrm{m}} = 1 - \frac{4\pi\sigma_{\mathrm{m}}}{\gamma\left[\left(\frac{\omega^2}{\gamma^2}\right)+1\right]} + j\frac{4\pi\sigma_{\mathrm{m}}}{\omega\left[\left(\frac{\omega^2}{\gamma^2}\right)+1\right]} \tag{59.4}$$

where σ_{m} is the static electrical conductivity of metals, its range for well-conducting metals is $10^{13} - 10^{14}\,\mathrm{s^{-1}}$, and γ is the collision frequency of electrons, $\gamma = 8.5 \times 10^{14}\,\mathrm{s^{-1}}$ (for solid aluminum at room temperature) [15]. In Eq. (59.4), $\left(\frac{\omega^2}{\gamma^2}\right) \ll 1$; therefore, the second term approaches zero for the real part of $\varepsilon^*_{\mathrm{m}} \to 1$. For aluminum, Eq. (59.4) can be expressed as:

$$\varepsilon^*_{\mathrm{m}} = 1 + j\frac{4\pi\sigma_{\mathrm{m}}}{\omega} \tag{59.5}$$

As from the above equation, it is clear that heat dissipated (dielectric loss) in the metallic material depends on σ_{m}. Thus, the dependence of static electrical conductivity of metal (σ_{m}) on temperature (T) is as follows:

$$\sigma_{\mathrm{m}}(T) = \frac{\sigma_0}{1 + \beta(T - T_0)} \tag{59.6}$$

where β is the temperature coefficient of metal and σ_0 is the electrical conductivity of metal at a temperature T_0.

Fourier's energy balance Eq. (59.7) can be used to estimate the microwave (MW) energy dissipation as heat inside the powder material at the time of microwave exposure as [16, 17]:

$$\rho C_P \frac{\partial T}{\partial t} + Q_{\mathrm{Loss}} = Q(r,T) + \nabla.(k\nabla T) \tag{59.7}$$

where C_P is the specific heat capacity $\left(\mathrm{J\,kg^{-1}\,K^{-1}}\right)$, ρ is the density $(\mathrm{kg\,m^{-3}})$, k is the thermal conductivity $\left(\mathrm{W\ m^{-2}\,K^{-1}}\right)$, Q_{Loss} is the heat losses from material inside the microwave, and $Q(r,T)$ is the heat source or microwave (MW) energy

absorbed by the powder material. For metal, it can be expressed as in Eq. (59.8) [16, 17]:

$$Q(r,T)=\frac{1}{2}\left[\omega\varepsilon''_{\mathrm{m}}(r,T)|E(r,T)|^2+\omega\mu''_{\mathrm{m}}(r,T)|H(r,T)|^2+\sigma_{\mathrm{m}}|E(r,T)|^2+\sigma_{\mathrm{i}}|E(r,T)|^2\right] \tag{59.8}$$

where $H(r,T)$ and $E(r,T)$ are the local electromagnetic field vectors at point r and temperature T in (K), μ''_{m} is magnetic loss factor $(\mathrm{H\ m^{-1}})$, $\varepsilon''_{\mathrm{m}}$ is dielectric loss factor $(\mathrm{F\ m^{-1}})$ of metallic materials, and σ_{i} is ionic conductivity of metallic ions in the metal lattice.

59.2.4 Boundary Conditions

Boundary conditions used to approach the real experimental conditions are as follows:

- The impedance boundary condition was used to define the minimum penetration depth of microwave on the wall of the microwave oven and waveguide. This boundary condition is shown in Fig. 59.2a and can be expressed by Eq. (59.9) [11]:

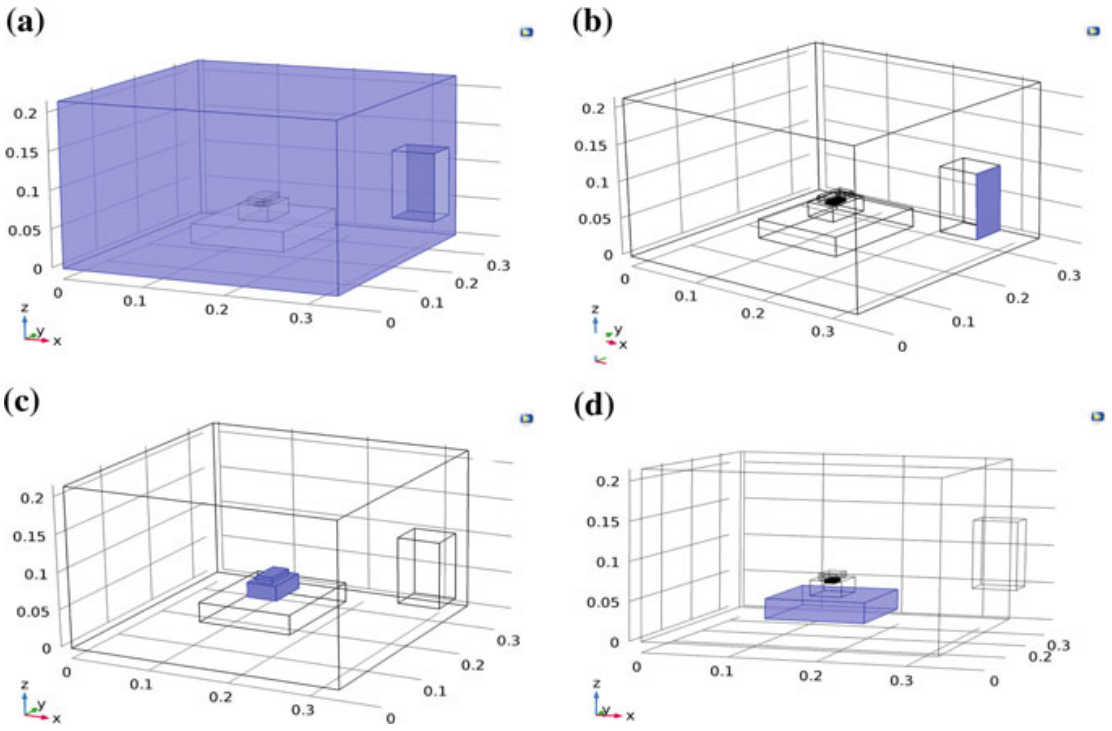


Fig. 59.2 Boundary conditions used in simulation: **a** impedance boundary condition, **b** port boundary condition, **c** heat transfer in solid and **d** thermal insulation

$$\sqrt{\frac{\mu_0\mu_r}{\varepsilon_0\varepsilon_r - j\frac{\sigma}{\omega}}}n \times H + E - (n.E)n = (n.E_s)n - E_S \tag{59.9}$$

where μ_0 is the permeability of vacuum, n is a unit vector which is perpendicular to the surface of the metal boundary, and E_S is the electric field component along the surface.

- A rectangular port boundary was used for wave excitation as the transverse electric (TE_{10}) field. The port boundary condition is shown in Fig. 59.2b. The port boundary condition at a small volume $\partial\Omega$ of edge element (S) is expressed by Eq. (59.10) [11]:

$$S = \frac{\int (E - E_1).E_1\partial\Omega}{\int E_1.E_1\partial\Omega} \tag{59.10}$$

- The heat transfer in solid condition is applied to the powder charge, mold, separator and susceptor as shown in Fig. 59.2c. The thermal insulation boundary condition (i.e., heat flux is zero) is shown in Fig. 59.2d and governed by Eq. (59.11) [11]:

$$-n \times (-k\nabla T) = 0 \tag{59.11}$$

- The heat flux and diffuse surface boundary conditions are applied on the surface of mold, separator and susceptor, which can be expressed by Eqs. (59.12, 59.13) [11]:

$$-n \times (-k\nabla T) = h(T_0 - T) \tag{59.12}$$

$$-n \times (-k\nabla T) = \varepsilon_{rad}\sigma_m\left(T_0^4 - T^4\right) \tag{59.13}$$

where h is the convective heat transfer coefficient and ε_{rad} is emissivity of the radiative surfaces.

59.2.5 Meshed Model

For simulating the above 3-D model, user-controlled mesh was used for meshing. Different types of free tetrahedral mesh like a finer, extra fine and extremely fine were used according to the suitable element size. Figure 59.3 shows mesh quality statistics (the average element quality and average growth rate) of 'finer' (Fig. 59.3a), 'extra fine' (Fig. 59.3b) and 'extremely fine' (Fig. 59.3c) mesh element sizes. The average element quality was good in case of extremely fined mesh with a low growth rate. The meshed model of microwave casting is shown in Fig. 59.4.

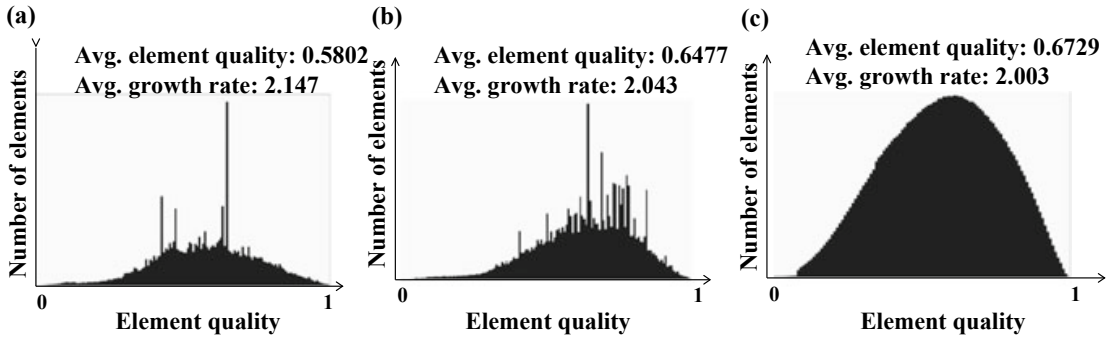
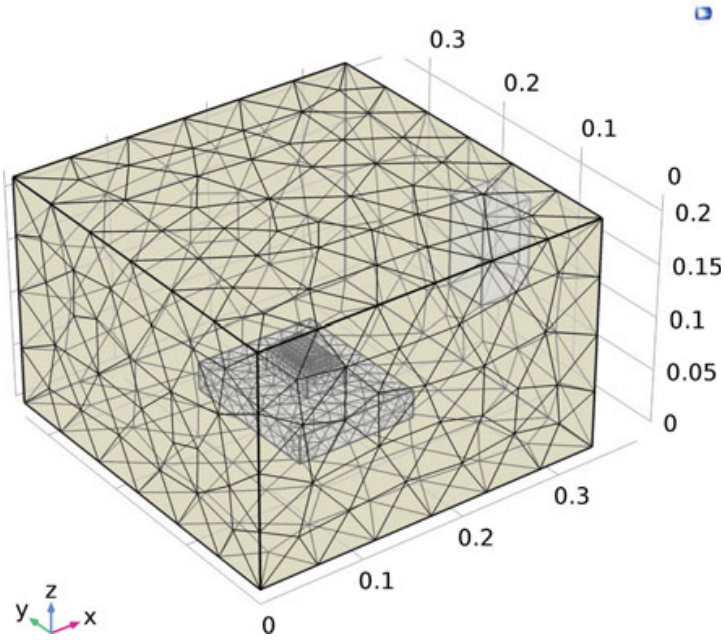


Fig. 59.3 Element quality comparison of **a** finer, **b** extra fine and **c** extremely fine element size

Fig. 59.4 Meshing of the experimental model



59.3 Results and Discussion

59.3.1 Electric Field Distribution and Temperature Distribution

The electric field distribution inside the microwave oven is shown in Fig. 59.5a. The energy is transmitted to the cavity at a power of 900 W and a frequency of 2.45 GHz. The distribution of the electric field is non-uniform inside the cavity because of the node and antinode formations due to the fluctuation of the electric field. So, the experimental setup should be kept precisely at a dense area of the electric field as it may affect the microwave absorption and hence, the energy required to melt the charged particle may be different.

As the electric field travels through dielectric materials like SiC plate and SiC powder, dielectric material starts to absorb it. This results in heating of dielectric

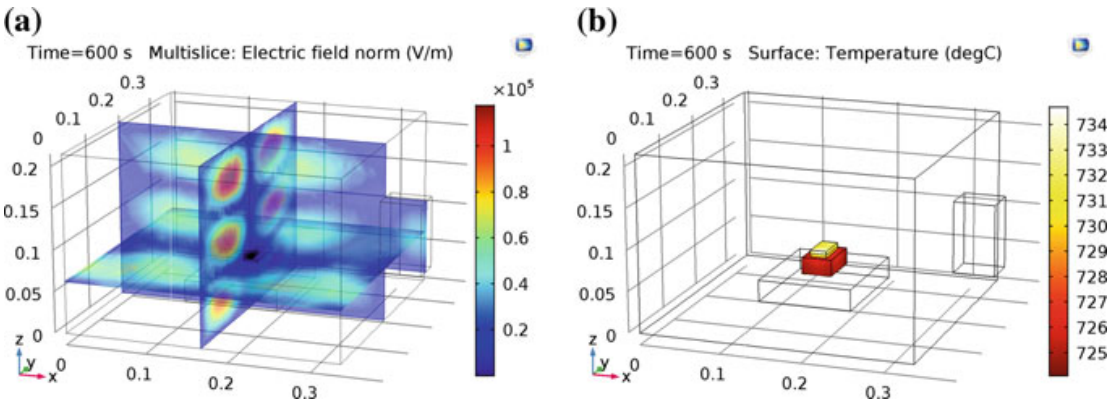


Fig. 59.5 **a** Electric field distribution inside the microwave oven and **b** temperature distribution

material. Initially, SiC powder couples with microwave radiation, starts heating up and transfers heat to SiC plate. SiC plate and graphite mold also start absorbing microwave and get heated rapidly, but charged powder absorbs very less microwave. Initially, the charged powder gets heated by conventional mode of heat transfer through SiC plate, susceptor and graphite mold. Once, the temperature of charged powder reaches up to a critical point, powder particles start absorbing microwave radiation and get heated by both conventional heating and microwave heating. The temperature increases rapidly with respect to the time, and melting of charged particles takes place. This heating phenomenon is known as microwave hybrid heating (MHH). This melted material is solidified and cast in a graphite mold.

The surface temperature distribution of heat transfer in the solid domain is shown in Fig. 59.5b. It is evident from the temperature distribution that temperature rise of SiC powder (susceptor) and SiC plate (separator) is more as compared to the metallic powder and graphite mold. This is due to the high dielectric constant, i.e., relative permittivity of SiC as compared to the metal as shown in Table 59.2. The time–temperature graph of charged powder is shown in Fig. 59.6. Initially, the heating rate of the charged powder is high but after reaching the melting point heating rate decreases. The melting of charged powder starts within 400 s, and after that temperature becomes constant. It can also be seen from Fig. 59.6 that slope of curve, i.e., rate of heating, increases after a certain time duration. This is the time duration after which powder particles start to absorb microwave radiation.

59.3.2 Effect of Input Power

The variation in input power affects the electric field norm inside the microwave oven and consequently affects the melting time of charged powder as shown in Fig. 59.7. The input power was varied 700–1300 W with step size 200 W with an exposure time of 600 s. It is observed from the figure that the maximum electric

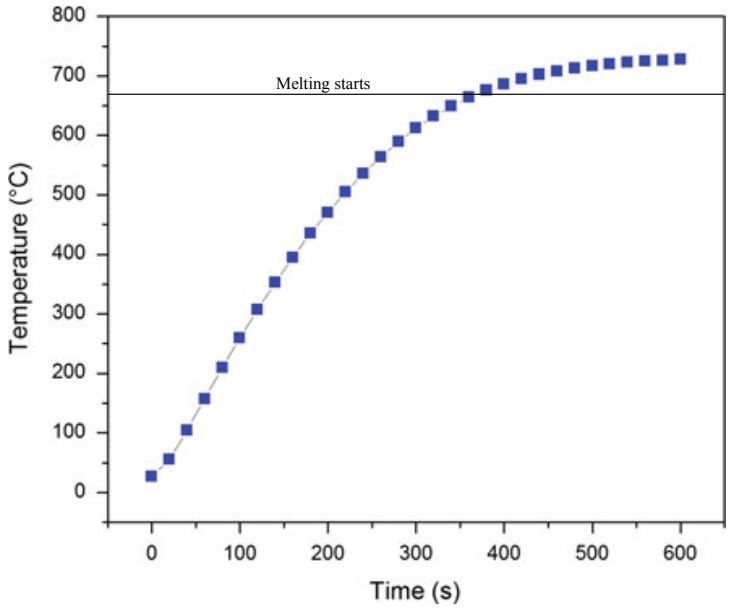


Fig. 59.6 Time–temperature graph at 900 W

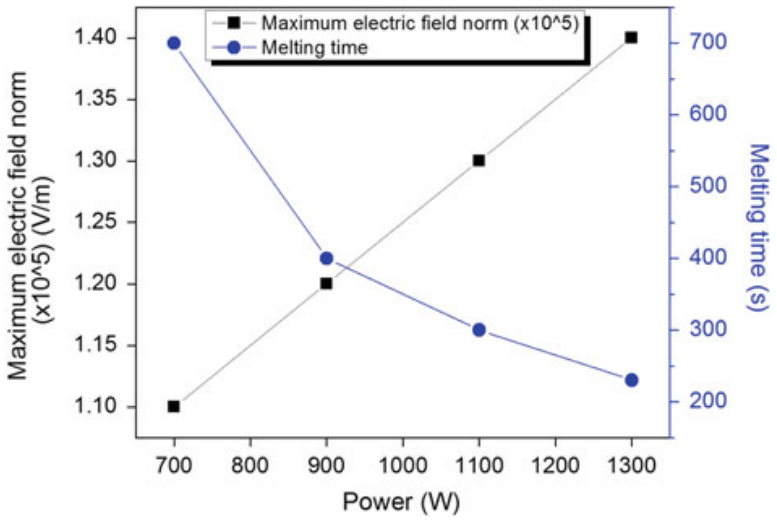


Fig. 59.7 Effect of power on electric field norm inside and the maximum temperature inside the charge

field norm inside the microwave increases linearly with input power. This happens because power absorbed inside the material is directly proportional to the electric field strength (Eq. 59.8). Also, the melting time of charged powder decreases with an increase in input power. The maximum electric field norm value varies from 1.1×10^5 to 1.4×10^5 V/m, and melting time decreased from 700 to 230 s by rising the input power from 700 to 1300 W, respectively.

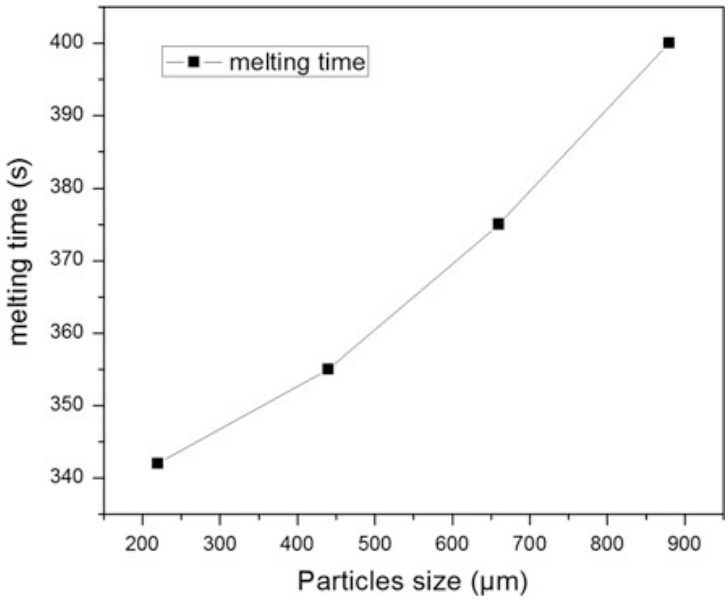


Fig. 59.8 Effect of powder particles in terms of melting time

59.3.3 Effect of Size of Powder Particles

Figure 59.8 shows the variation in melting time of charged powder with variation in particle size at 900 W. As the AA 7039 powder size reduces, the melting time of the charged powder reduces. Thus, it indicates that smaller-sized particles have higher microwave absorption capacity [18]. Microwave gets absorbed inside the material up to a particular depth called skin depth. As the particle size is reduced, the effective surface area which interacts with microwave increases and consequently, factors like Joule loss increase which increase the microwave absorption capacity [12, 19].

59.4 Conclusion

A 3-D microwave oven setup was modeled to find the precise location of experimental setup inside the cavity. COMSOL Multiphysics software tool was used successfully to simulate the microwave hybrid heating (MHH) of AA 7039 powder at a rated power of 900 W and 2.45 GHz frequency. Following are the conclusions based on the above study:

- It was observed that the experimental setup should be placed at the high magnitude or dense area of the electric field.
- It was found that the melting of the metallic powder achieved within 400 s and temperature rises to −730 °C in 600 s at the power of 900 W.
- It was observed that higher is the input power, more is the electric field norm, consequently reducing the melting time of charged powder.

- From the simulation results obtained, it can be concluded that decreasing the powder particle size results in more absorption of microwave energy and hence the time required for melting and casting of metallic powder reduces.

Acknowledgements The authors would like to express sincere gratitude to Prof. Apurbba Kumar Sharma for allowing and doing the simulation work using COMSOL Multiphysics in Microwave Material Processing Laboratory, IIT Roorkee.

References

1. Ravi, B.: Metal casting: computer-aided design and analysis. PHI Learning Pvt. Ltd. (2005)
2. Roy, R., Agrawal, D., Cheng, J., Gedevanishvili, S.: Full sintering of powdered-metal bodies in a microwave field. Nature **399**(6737), 668 (1999)
3. Ghosh, S., Pal, K.S., Dandapat, N., Mukhopadhyay, A.K., Datta, S., Basu, D.: Characterization of microwave processed aluminium powder. Ceram. Int. **37**(3), 1115–1119 (2011)
4. Mishra, R.R., Sharma, A.K.: On mechanism of in-situ microwave casting of aluminium alloy 7039 and cast microstructure. Mater. Des. **112**, 97–106 (2016)
5. Bansal, A., Sharma, A.K., Das, S.: Metallurgical and mechanical characterization of mild steel-mild steel joint formed by microwave hybrid heating process. Sadhana **38**(4), 679–686 (2013)
6. Kumar, G., Sharma, A.K.: Role of dielectric fluid and concentrator material in microwave drilling of borosilicate glass. J. Manufact. Processes **33**, 184–193 (2018)
7. Gupta, D., Sharma, A.K.: Development and microstructural characterization of microwave cladding on austenitic stainless steel. Surf. Coat. Technol. **205**(21–22), 5147–5155 (2011)
8. Yoshikawa, N.: Fundamentals and applications of microwave heating of metals. J. Microw. Power Electromagn. Energy **44**(1), 4–13 (2010)
9. Chandrasekaran, S., Basak, T., Ramanathan, S.: Experimental and theoretical investigation on microwave melting of metals. J. Mater. Process. Technol. **211**(3), 482–487 (2011)
10. Singh, S., Gupta, D., Jain, V.: Novel microwave composite casting process: theory, feasibility and characterization. Mater. Des. **111**, 51–59 (2016)
11. Mishra, R.R., Sharma, A.K.: On melting characteristics of bulk Al-7039 alloy during in-situ microwave casting. Appl. Therm. Eng. **111**, 660–675 (2017)
12. Mishra, R.R., Sharma, A.K.: Structure-property correlation in Al–Zn–Mg alloy cast developed through in-situ microwave casting. Mater. Sci. Eng. A **688**, 532–544 (2017)
13. Mishra, P., Upadhyaya, A., Sethi, G.: Modeling of microwave heating of particulate metals. Metall. Mater. Trans. B **37**(5), 839–845 (2006)
14. Buchelnikov, V.D., Louzguine-Luzgin, D.V., Anzulevich, A.P., Bychkov, I.V., Yoshikawa, N., Sato, M., Inoue, A.: Modeling of microwave heating of metallic powders. Physica B **403** (21–22), 4053–4058 (2008)
15. Eidmann, K., Meyer-ter-Vehn, J., Schlegel, T., Hüller, S.: Hydrodynamic simulation of subpicosecond laser interaction with solid-density matter. Phys. Rev. E **62**(1), 1202 (2000)
16. Mishra, R.R., Sharma, A.K.: A review of research trends in microwave processing of metal-based materials and opportunities in microwave metal casting. Crit. Rev. Solid State Mater. Sci. **41**(3), 217–255 (2016)
17. Maxim, I., Tanaka, M.: Numerical analysis of the microwave heating of compacted copper powders in single-mode cavity. Jpn. J. Appl. Phys. **50**(9R), 097302 (2011)

18. Mondal, A., Shukla, A., Upadhyaya, A., Agrawal, D.: Effect of porosity and particle size on microwave heating of copper. Sci. Sinter. **42**(2), 169–182 (2010)
19. Crane, C.A., Pantoya, M.L., Weeks, B.L., Saed, M.: The effects of particle size on microwave heating of metal and metal oxide powders. Powder Technol. **256**, 113–117 (2014)

Chapter 60
Tribological Behaviour of Nylon 6 Composites and Optimization of Process Parameters Using Hybrid Taguchi-GRA and CSA

Sudhir Kumar, Barnik Saha Roy and K. Panneerselvam

Abstract Friction and wear are the main causes observed by manufacturing industries which lead to the replacement of components. The objective of the present work is to investigate the tribological behaviours of Nylon 6 and glass-filled (GF) Nylon 6 composites using a pin-on-disc configuration. Pin materials were fabricated by injection moulding machine and tested for friction and wear against 320 grit size silicon carbide (SiC) sandpaper. The effect of the different load (N), glass-filled content (wt%) and sliding distance (m) were studied for the coefficient of friction (COF) and specific wear rate (SWR). Analysis of variance (ANOVA) procedure was used to find out the percentage contribution of process parameters (load, sliding distance and varying glass-filled content) with the COF and SWR. Process parameters were optimized by hybrid Taguchi-grey relational analysis (GRA) and cuckoo search algorithm (CSA). There is a close relationship between COF and SWR with wear mechanisms.

Keywords Polymer composites · COF · SWR · Taguchi · GRA · CSA

S. Kumar (✉) · B. S. Roy
Department of Mechanical Engineering, National Institute of Technology, Agartala 799046, India
e-mail: sudhir007gec@gmail.com

B. S. Roy
e-mail: barnik.me@nita.ac.in

K. Panneerselvam
Department of Production Engineering, National Institute of Technology, Tiruchirappalli 620015, India
e-mail: kps@nitt.edu

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_60

60.1 Introduction

Polymer matrix composites (PMCs) are replacing traditional metals and unfilled polymers due to their superior properties such as intrinsic lubrication property and wear resistance with low processing cost [1]. The application of polymer composites is mostly in sliding components in the food and chemical industry like gears and cams. This may be mainly attributed to the Nylon having inherent lubrication behaviour which avoided the use of oil lubrication [2]. Ramesh and Suresh carried out friction and wear test of polymer composites and observed grit size dominant factors followed by applied load and sliding distance [3]. Grey relational analysis (GRA) may be used to convert multiple objective optimization problems to a single objective problem [4–6]. Yang and Deb [7] investigated a new evolutionary and population-based algorithm cuckoo search for optimizing process parameters. Cuckoo search is based on the breeding behaviour of cuckoo birds. The study on hybrid Taguchi-GRA and CSA for multi-criteria optimization in tribological process parameters is very rare. The aim of present work is to investigate the friction and wear characteristic of Nylon 6 and glass-filled Nylon 6 composites against the abrasive paper. There is a close relationship between COF and SWR with wear mechanisms.

60.2 Materials and Methods

The material selected in this investigation for pin specimens was glass-filled Nylon 6 (0, 10, 20 and 30 wt%) in the form of granules as shown in Fig. 60.1. Injection moulding machine was used to fabricate the Nylon 6 composite specimens for mechanical (UTS, impact strength and hardness) and tribological test. The mould temperature was maintained at 240 °C. The size of stationary GF Nylon 6 pin specimen was 6 mm-diameter and 31 mm-length according to ASTM G99-05. Schematic illustration of the pin-on-disc set-up is shown in Fig. 60.2. Track diameter of the disc (320 grit size sandpaper) was 40 mm. The Taguchi technique is a powerful tool to optimize process parameters. The process parameters used for the experiment are presented in Table 60.1. Grey relational analysis is used to convert multi-objective (COF and SWR) into a single response optimization problem.

Cuckoo search (CS) is an evolutionary and population-based approach that mimics the breeding behaviour of cuckoo birds. The preliminary experiments show the promising results of this approach in the field of search and optimization. The idealized steps of CS can be stated as follows:

(1) Every cuckoo bird lays an egg on the randomly selected nest.
(2) The number of nests (agents) and eggs (individuals) in each nest is fixed.
(3) Nest with the high quality of eggs (elite members) is passed on to the next generation.

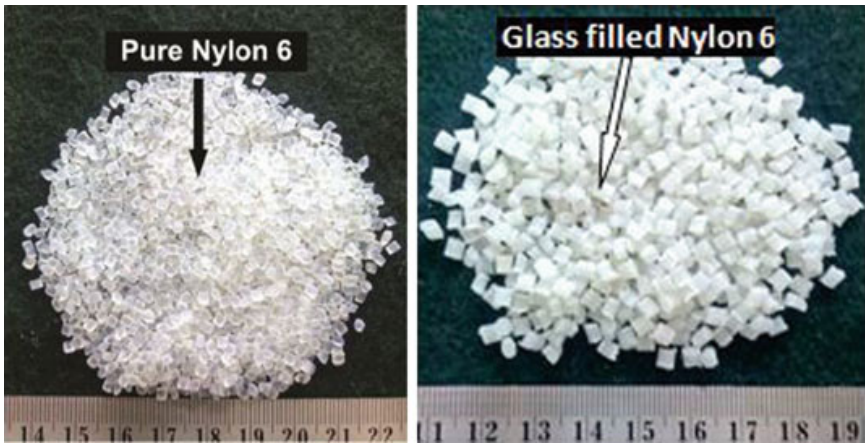


Fig. 60.1 Pure Nylon 6 and GF Nylon 6 (granule form) [1]

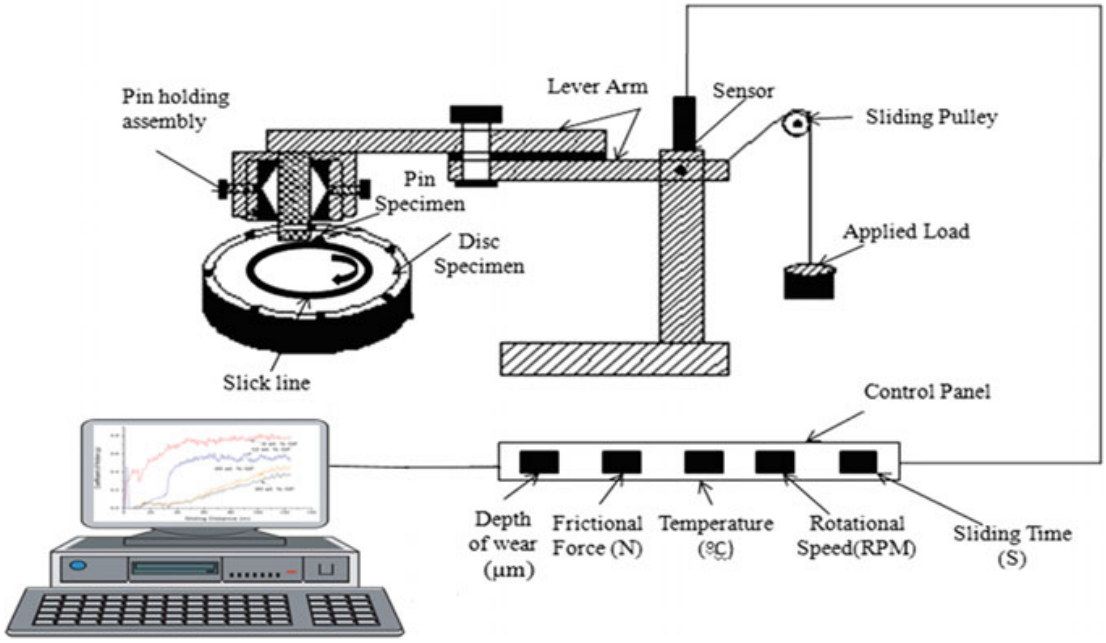


Fig. 60.2 Experimental set-up of pin-on-disc [8]

Table 60.1 Process parameters and their levels

Factor symbol	Control factors	Levels			
		1	2	3	4
A	Glass filled (wt%)	0	10	20	30
B	Load (N)	5	10	15	20
C	Sliding distance (m)	125	250	375	500

(4) If the host bird discovers an alien egg based on the external characteristics of an egg such as colour and spots, then host bird either throws the alien egg away or builds an entirely new nest at the new location.

For feature selection, the host nest n is defined as an agent containing multiple eggs (or individuals) called population. Each egg is represented as a binary string, with a 1 for inclusion of feature (or attribute) and a 0 for the exclusion of feature. Step 4 is implemented by maintaining the index of the worst nest and replacing it with better individuals. Levy flight is a random walk and named after French Mathematician Paul Levy, in which steps are defined in terms of step length having a probability distribution that is heavy-tailed. The direction of the steps in Levy flights can be isotropic or random. Various studies demonstrated the flight behaviour of many insects and animals based on Levy flights. In the past decade, these behaviours have been applied in optimization and optimal search and found to yield promising results [5–7].

60.3 Results and Discussion

The mechanical characterization includes evaluation of impact strength, UTS and hardness of the composites that are presented in Table 60.2. The UTS increased as to increase the GF wt% presented in Table 60.2. This may be attributed to the better distribution of GF and superior adhesion among the Nylon and GF. The hardness (shore D) increased as to increase the GF wt% as shown in Table 60.2. This may be attributed to brittleness character of GF. Nylon 6 material has high impact strength due to ductile nature and flexibility of the interface molecular chain resulting in absorbing more energy presented in Table 60.2. The impact strength increased as to increase the GF wt%. The improvement in impact strength of composites is due to good bonding strength between matrix and reinforcement. The experiments were conducted using L_{16} orthogonal array details which are shown in Table 60.3.

Table 60.2 Mechanical properties of various glass-filled contents (wt%) of Nylon 6 [1]

Glass-filled content (wt%)	Hardness (shore D) ASTM D2240-05	Notched impact strength (J/mm) ASTM D256-10	UTS (MPa) ASTM D638-14
0	64	0.137	51
10	65	0.075	54
20	73	0.125	71
30	74	0.125	86

Table 60.3 GRA for tribological properties of Nylon 6 composites

Exp. level	Weight loss (g)	COF (μ)	SWR (mm^3/Nm)	Normalized response		GRC		GRG	Rank
				COF	SWR	COF	SWR		
1($A_1B_1C_1$)	0.084	0.67	0.11	0	0	0.33	0.33	0.33	16
2($A_1B_2C_2$)	0.178	0.65	0.06	0.07	0.54	0.35	0.52	0.43	15
3($A_1B_3C_3$)	0.241	0.60	0.03	0.26	0.78	0.40	0.70	0.55	12
4($A_1B_4C_4$)	0.293	0.49	0.02	0.69	0.90	0.61	0.84	0.73	8
5($A_2B_1C_2$)	0.088	0.62	0.05	0.19	0.58	0.38	0.54	0.46	14
6($A_2B_2C_1$)	0.063	0.59	0.04	0.30	0.74	0.41	0.66	0.54	13
7($A_2B_3C_4$)	0.226	0.52	0.02	0.57	0.91	0.54	0.85	0.69	9
8($A_2B_4C_3$)	0.202	0.48	0.02	0.73	0.93	0.65	0.89	0.77	7
9($A_3B_1C_3$)	0.094	0.47	0.03	0.76	0.77	0.68	0.68	0.68	10
10($A_3B_2C_4$)	0.126	0.46	0.01	0.80	0.96	0.72	0.93	0.82	4
11($A_3B_3C_1$)	0.080	0.51	0.03	0.61	0.83	0.56	0.74	0.65	11
12($A_3B_4C_2$)	0.130	0.43	0.02	0.92	0.87	0.86	0.79	0.83	3
13($A_4B_1C_4$)	0.069	0.48	0.02	0.73	0.95	0.65	0.92	0.78	5
14($A_4B_2C_3$)	0.082	0.45	0.016	0.84	1	0.76	1	0.88	2
15($A_4B_3C_2$)	0.089	0.50	0.01	0.65	0.98	0.59	0.97	0.78	6
16($A_4B_4C_1$)	0.069	**0.41**	**0.02**	**1**	**0.96**	**1**	**0.93**	**0.96**	**1**

60.3.1 Effect of GF Content, Load and Sliding Distance on COF

The variation of COF w.r.t different sliding distances is depicted in Fig. 60.3. The load shared by both glass filled and Nylon 6 during sliding as depicted in Fig. 60.4, the major part of the load supported by GF resulting COF decreased. In the case of pure Nylon 6, the COF was observed more; however, further increase in glass-filled COF decreased due to rubbing of glass filled on the abrasive paper as shown in Fig. 60.5.

The COF of the Nylon 6 composites decreases as to increase load. This may be mainly due to the increase in heat generation of the surface of the specimen and contact zone of pin specimen. At 5 N load, the thermal motion of molecules of polymer observed weak, this may be because the heat generation at contact zone was low, then plastic deformation of polymer molecule cannot respond to load, and thus COF is relatively high. At 9.8 N loads, as the heat generation increases molecular segment of polymer movement gradually increases, then plastic deformation of polymer molecule can respond to load, and thus the corresponding COF gradually decreases. At 15 N loads, the heat generation increases molecular segments likely to move freely, the plastic deformed polymeric material at contact zone of pin and disc forms a low shear strength interfacial film which acts like as lubricant, and thus COF again decreases as shown in Fig. 60.5. The COF decreases

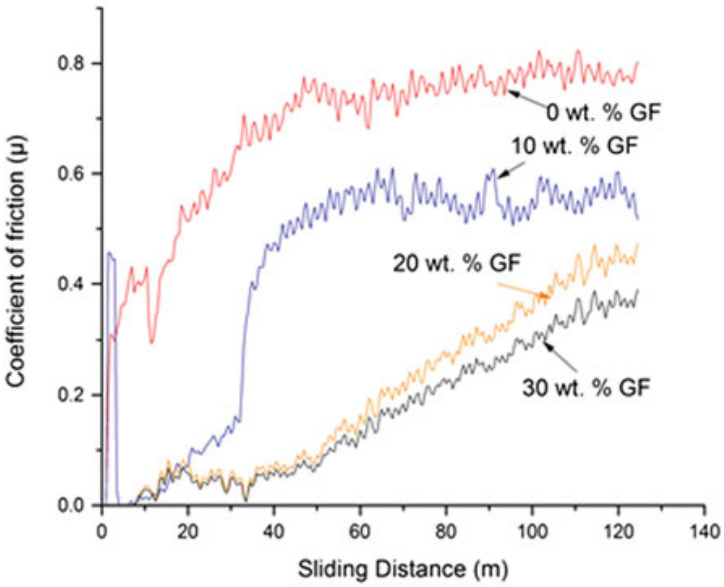


Fig. 60.3 Relation between sliding distance and coefficient of friction irrespective of load

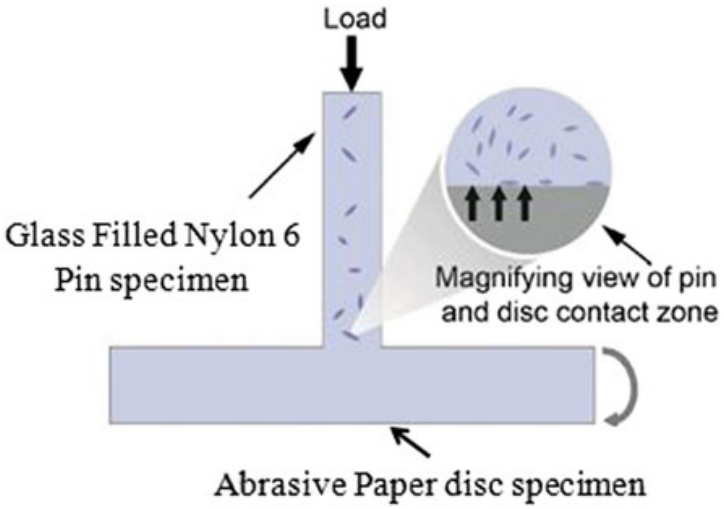


Fig. 60.4 Schematic representation of load sharing between GF and Nylon 6 [8]

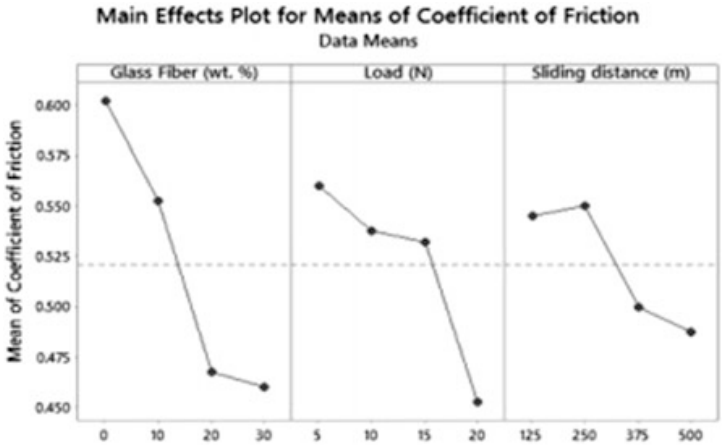


Fig. 60.5 Relation between process parameters and COF

with an increase in sliding distance for Nylon 6 composites. This may be mainly attributed to pin specimen slides over the same wear track, thus resulting in less effect of abrasiveness. Similar trends are found by other researchers [8, 9].

60.3.2 *Effect of GF Content, Load and Sliding Distance on SWR*

The values of SWR are observed from 0.0161 mm^3/Nm to 0.1189 mm^3/Nm. In the case of Nylon 6, SWR was very high and when glass filled increased then SWR decreased. When glass filled increases hardness and strength increases, so wear resistance improved. During sliding, load shared by glass filled and Nylon 6 is depicted in Fig. 60.4. This may be mainly attributed to improved wear resistance. The SWR decreases as an increase in load for Nylon 6 composites as shown in Fig. 60.6. Similar trends were found by other researchers [8, 9].

The variation of SWR w.r.t different sliding distances at a different GF is depicted in Fig. 60.8. The SWR of Nylon 6 composites decreased as an increase in the sliding distance. This may be mainly attributed to the repetition of the pin on same wear track gap on emery paper filled by debris.

60.3.3 *Grey Relational Analysis and Implementation of Cuckoo Search Algorithm*

The grey relational analysis results are shown in Table 60.3. The optimum parameter combination for minimum COF and SWR is shown in Fig. 60.7. ANOVA was done for grey relational grades, and the results are presented in Table 60.4. It was observed that glass filled has the dominant factor followed by load and sliding distance.

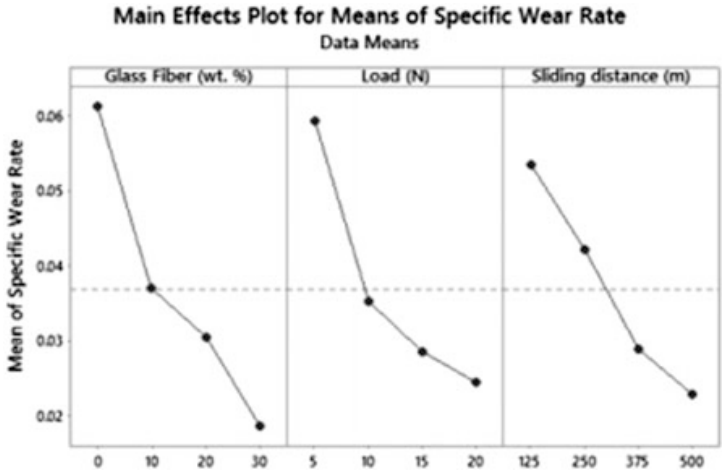


Fig. 60.6 Relation between process parameters and SWR

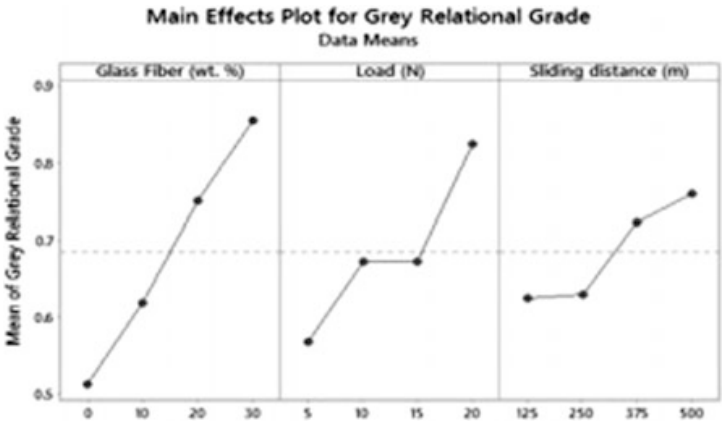


Fig. 60.7 Effect of control factors with their levels for GRG

Table 60.4 Analysis of variance of grey relational analysis

Source	DOF	Adj. SS	Adj. MS	F-cal.	P-value	% contribution
Glass filled (wt%)	3	0.26	0.08	129.85	0.000	57.92
Load (N)	3	0.13	0.04	65.07	0.000	29.02
Sliding distance (m)	3	0.05	0.01	27.25	0.001	11.74
Residual error	6	0.00	0.00	–	–	1.32
Total	15	0.46	–	–	–	–

Lower and upper limits of the parametric constraints are $0 \leq$ glass filled (wt%) ≤ 30, $125 \leq$ sliding distance (m) ≤ 500, $5 \leq$ load (N) ≤ 20. The cuckoo search parameter values are the number of nests $n = 25$, discover rate = 0.25, number of iteration = 100. Figure 60.8 shows the objective obtained during minimization for all iterations. It was revealed that firstly, the value of the objective function is at a higher value and as the iteration progresses, the objective function

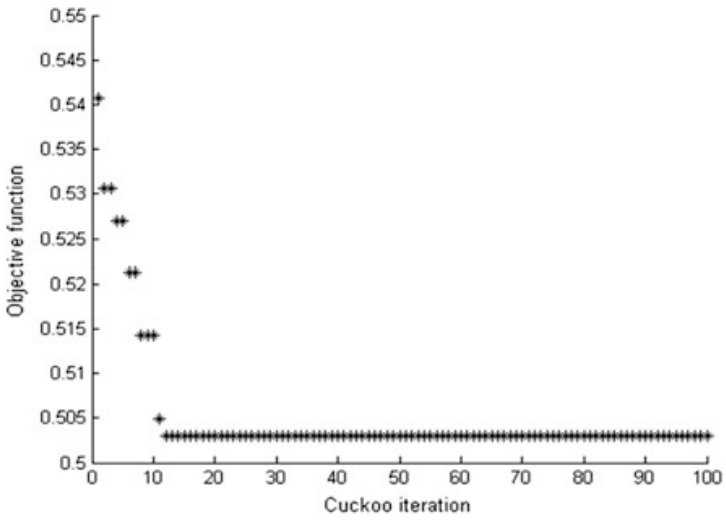


Fig. 60.8 Variation of objective function with iterations

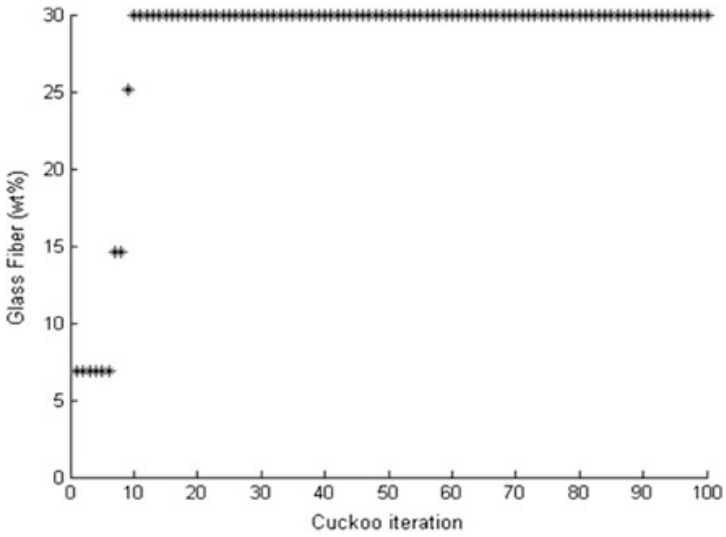


Fig. 60.9 Variation of glass filled with iterations

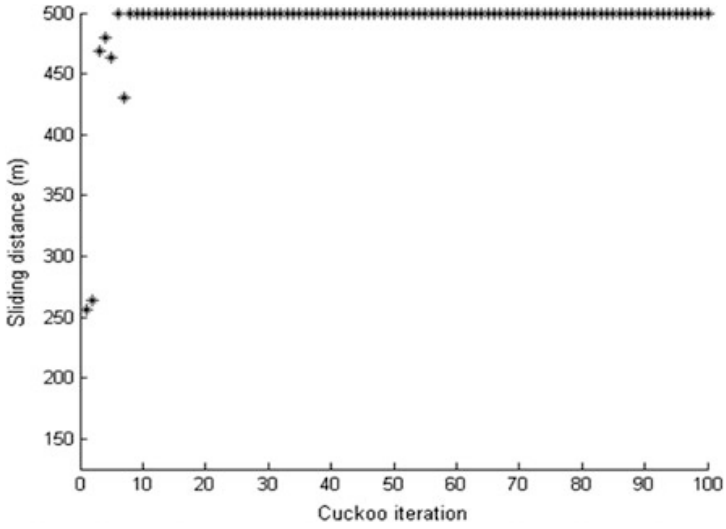


Fig. 60.10 Variation of sliding distance with iterations

values converge quickly. The best solution obtained is the objective function value of 0.5030 for glass-filled content of 30 wt%, sliding distance 500 m and load of 20 N for lower COF and SWR as shown in Figs. 60.9, 60.10 and 60.11.

60.3.4 Confirmation Tests for CSA

With the identified optimal process parameters of glass-filled content of 30 wt%, sliding distance of 500 m and load of 20 N, a confirmation experiment was conducted as shown in Table 60.5.

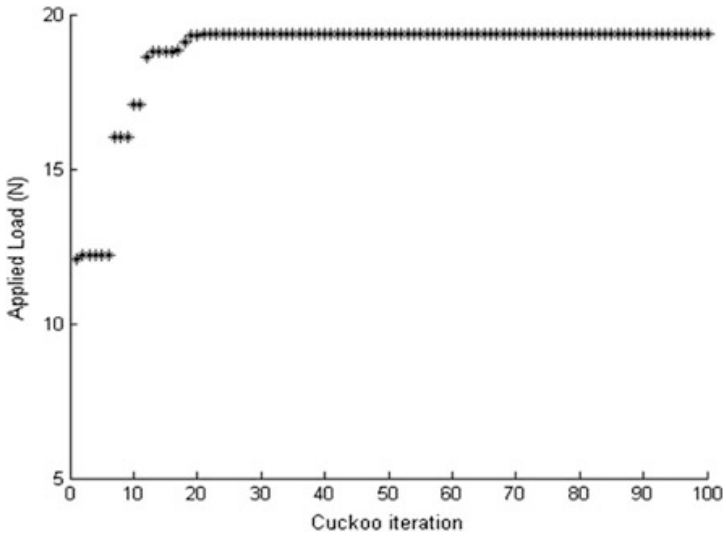


Fig. 60.11 Variation of load with iterations

Table 60.5 Measured responses of confirmation experiment for cuckoo search

	Optimal parameters	
	Prediction	Experiment
Parameter levels	$A_4B_4C_4$	$A_4B_4C_4$
COF (μ)		0.43
SWR (mm^3/Nm)		0.010
Grey relational grade (GRG)	0.988	0.967
Objective function = 1/(1 + GRG)	0.5030	0.5083

60.3.5 Worn Surface Morphology

Figure 60.12 shows a schematic illustration of abrasive wear mechanism. Figure 60.13a shows typical micrograph of unabraded counter-face of 320 grit size SiC abrasive paper. In case of low sliding distance, repetition of the pin over the wear track is less so clogging of abrasive wear debris is observed low as shown in Fig. 60.14a. However, in the case of 500-m sliding distance repetition of the pin over the wear track is more resulting in less effect of abrasiveness as shown in Figs. 60.13b and 60.14b. Consequently, decrease COF and SWR. Plastic deformation, micro-cracking and micro-cutting are the dominant wear mechanisms found in case of pure Nylon 6. This may be mainly attributed to more frictional heat generated at contact zone as shown in Fig. 60.15a. However, in the case of glass-filled Nylon 6, polymer softening decreased as depicted in Fig. 60.15b. This may be mainly attributed to the load shared by glass filled and Nylon 6 as shown in Fig. 60.4. Typical micrograph of 30 wt% GF Nylon 6 at 20 N shows more damage worn surface which results in de-bonding of the glass filled from the matrix. Consequently at 5 N, worn surface observed mild wear. Wear mechanism shows a good correlation with the trend of COF and SWR (Fig. 60.16).

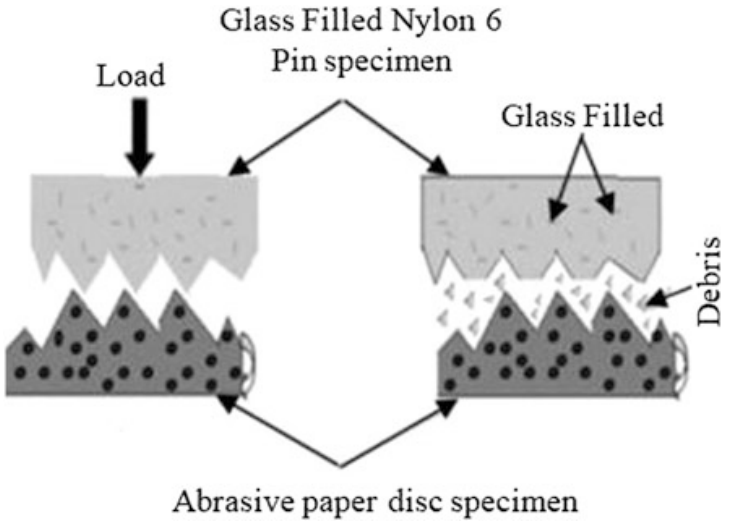


Fig. 60.12 Abrasive wear mechanism of GF Nylon 6 [8]

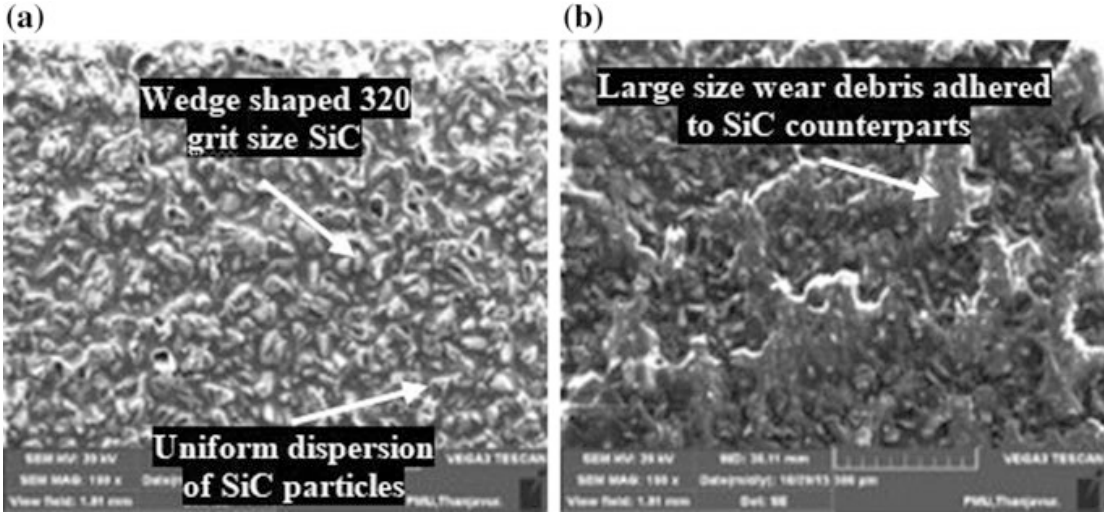


Fig. 60.13 Typical micrograph of **a** unabraded counter-face and **b** abraded counter-face at 500 m sliding distance

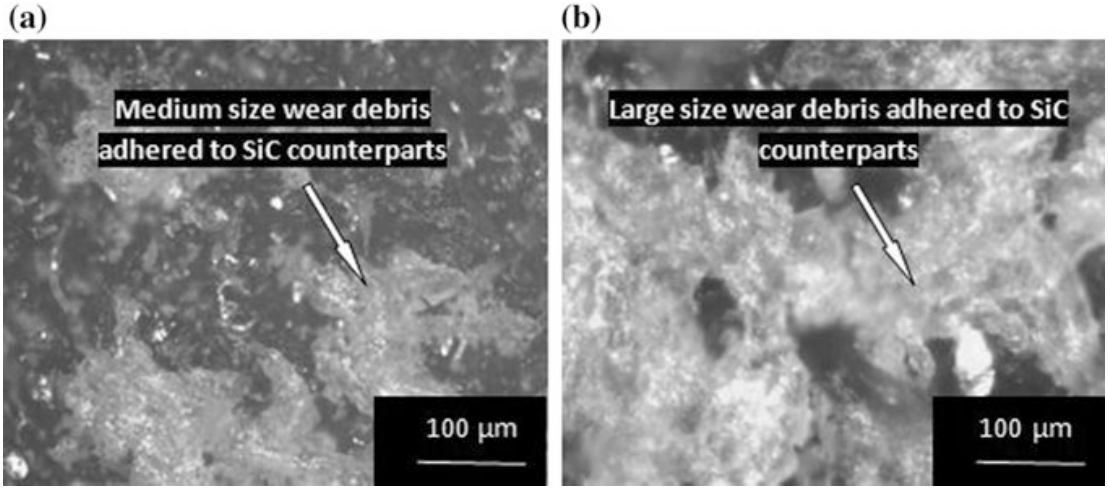


Fig. 60.14 Typical micrograph of abraded counter-face at sliding distance: **a** 125 m and **b** 500 m

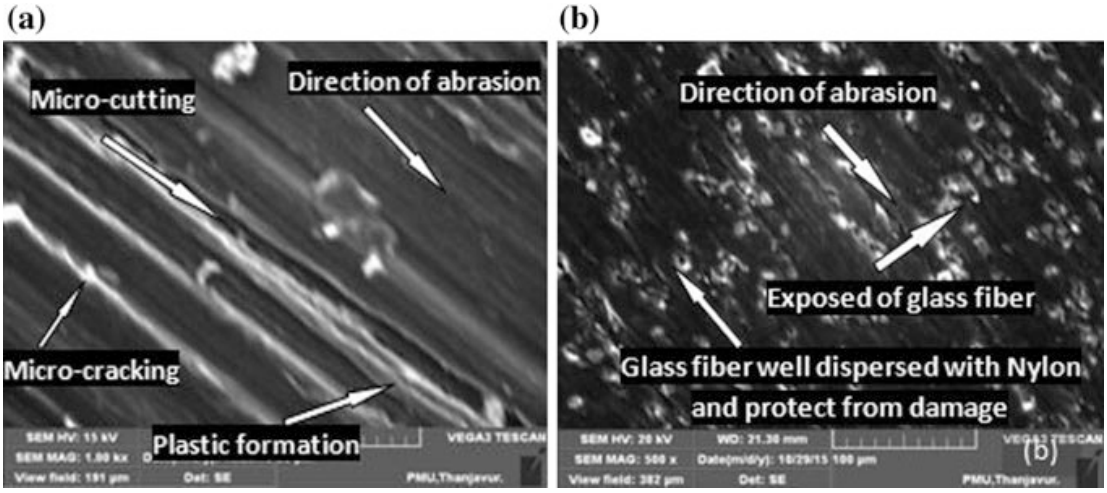


Fig. 60.15 Typical micrograph of worn surface at 20 N load and 500 sliding distance: **a** Nylon 6 and **b** 30 wt% GF Nylon 6

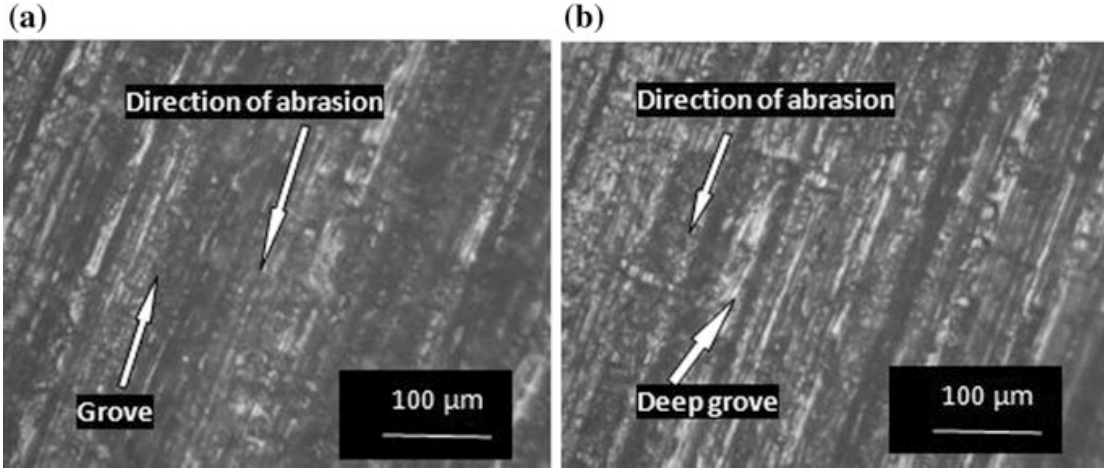


Fig. 60.16 Typical micrograph of worn surface of 30 wt% GF Nylon 6: **a** at 5 N load and **b** at 20 N load

60.4 Conclusions

From the experimental investigation, the following conclusions have been drawn:

- The present study demonstrated successful fabrication of glass-filled Nylon 6 composites using injection moulding.
- Glass-filled Nylon 6 improves the mechanical (hardness and tensile strength) and tribological properties (COF, weight loss and SWR).
- The weight loss of Nylon 6 composites increases as load increases; however, the addition of glass filled decreases the weight loss.
- The COF and SWR of glass filled Nylon 6 composites decreases as increase glass filled content and least achieved at 30 wt%.

- On the basis of ANOVA, glass filled content 57.92%, sliding distance 11.74% and load 29.02% exerts a significance influence on multiple responses.
- Cuckoo search shows that 30 wt% glass-filled content, 20 N load and 500 m sliding distance combinations of factor give the best solution.
- In case of Nylon 6 plastic deformation, micro-cracking and micro-cutting are dominant wear mechanisms; however, glass-filled Nylon 6 reduces these defects. The worn surface is characterized by the direction of abrasion.

References

1. Kumar, S., Panneerselvam, K.: Investigations on dry sliding wear behaviour of Nylon-6 and GFR Nylon-6 composites using Taguchi technique. Indian J. Eng. Mater. Sci. **25**, 89–97 (2018)
2. Zhang, S.W.: State-of-the-art of polymer tribology. Tribol. Int. **31**(1), 49–60 (1998)
3. Ramesh, B.N., Suresha, B.: Optimization of tribological parameters in abrasive wear mode of carbon-epoxy hybrid composites. Mater. Des. **59**, 38–49 (2014)
4. Kuo, Y., Yang, T., Huang, G.W.: The use of grey based Taguchi method for optimizing multi response optimization problems. Eng. Optim. **4**, 517–528 (2008)
5. Kumar, S., Panneerselvam, K.: Optimization of friction and wear of Nylon 6 and glass fiber reinforced (GFR) Nylon 6 composites against 30 wt.% GFR Nylon 6 Disc. J. Adv. Res. Mater. Sci. **19**, 14–32 (2016)
6. Senthilkumar, N., Tamizharasan, T., Anandakrishnan, V.: An hybrid Taguchi-grey relational technique and cuckoo search algorithm for multi-criteria optimization in hard turning of AISI D3 steel. J. Adv. Eng. Res. **1**(1), 16–31 (2014)
7. Yang, X.S., Deb, S.: Engineering optimisation by cuckoo search. Int. J. Math. Model. Numer. Optim. **1**, 330–343 (2010)
8. Kumar, S., Panneerselvam, K.: Two-body Abrasive wear behavior of Nylon 6 and glass fiber reinforced (GFR) Nylon 6 composite. Proc. Technol. **25**, 1129–1136 (2016)
9. Jafrey, D.D., Panneerselvam, K.: Abrasive wear of polypropylene/Cloisite 30B/Elvaloy AC 3427 nanocomposites. J. Compos. Mater. **52**(13), 1833–1843 (2018)

Chapter 61
Fabrication of Hexagonal Bar from Aluminum Alloy AA6063 Scrap by Frictional Stir Back Extrusion on Milling Machine

Gizaw Yohannes, Habtamu Beri and Perumalla Janaki Ramulu

Abstract Friction stir back extrusion (FSBE) is a novel and an efficient method of manufacturing sounded products by using specialized tool in one step (without melting of scrap). It can be saved about 40% of material, 26–31% of energy consumption, and 16–60% of labor costs compared with conventional extrusion. The main aim of this study is fabrication of hexagonal bar from aluminum alloy scrap by FSBE. The material used for this study is AA6063 as a scrap and H-13 tool steel as a die. Considered parameters were chamber rotational speed and vertical die transverse speed as major parameter on conventional milling machine. This work has proved the feasibility to extrude hexagonal-shaped bar by using FSBE by fixed die and rotates chamber. So far, none of the researchers carried out experiment by fixed die and rotate chamber in FSBE, and also no one applied FSBE to produce other than circular shape. Aluminum hexagonal-shaped bar is used as a supporter in medical and marine industry, as a raw material to the mass production of hexagonal head aluminum nut and bolt, Allen wrenches, etc. Extruded bar was defect-free, non-blurred grain boundary, and clear surface than the base material. The result showed that the compression strength of extruded sample was higher than the base material. The optimum rotational speed of chamber to produce hexagonal bar with the minimum twist angle is 250 rpm and transverse speed 6.25 mm/min. If the rotational speed increases, twist angle also increased.

G. Yohannes · H. Beri
Program of Mechanical Design and Manufacturing Engineering, School of Mechanical, Chemical and Materials Engineering, Adama Science and Technology University, Adama, Ethiopia
e-mail: gizawgotayohannes@gmail.com

H. Beri
e-mail: habtamu_beri@yahoo.com

P. J. Ramulu (✉)
Program of Mechanical Design and Manufacturing Engineering & Director of CoEAME, School of Mechanical, Chemical and Materials Engineering, Adama Science and Technology University, Adama, Ethiopia
e-mail: perumalla.janaki@astu.edu.et

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_61

Keywords Hexagonal bar · Friction stir back extrusion process · Aluminum alloy scrap · H-13 tool steel die and chamber

61.1 Introduction

Friction stir extrusion (FE) was developed by The Welding Institute (Cambridge, UK) in the early 1990s. It was derived from friction stir processing as a process of forming or reforming ceramic/metal composite material. Over the past two decades, significant amount of work had done on multiple aspect of friction stir process. Along with the rapid development of automotive, aerospace, and manufacturing industry in recent years, novel technologies like near-net shape manufacturing, additive manufacturing, solar energy jet, and fuel cell battery plays a critical role in producing safe, economical, and durable products. Friction stir extrusion process can meet this need as an innovative and economical material processing method as it directly converts metal wastes without melting to useful products. Compared with traditional manufacturing processes, it has more efficient, green, energy saving, and environmentally friendly technologies [1]. Friction stir extrusion relies on extensive plastic deformation and basic frictional heating, consolidates, and extrudes the small block, chips, and scrap directly into functional form. For example, direct recycling aluminum can save about 40% of material, 26–31% of energy consumption, and 16–60% of labor costs. The friction extrusion process falls into the category of direct recycling methods and it has all the advantages [2].

Aluminum and its alloys had received much attention in aerospace, automobile, and construction industries; due to their high strength-to-weight ratio, good formability, excellent corrosion resistance, low maintenance, good toughness, and high strength-to-weight ratio [3]. Tang and Reynolds [4] applied friction stir extrusion (FSE) technology and produced fully consolidated wires from aluminum alloy AA2050 and AA2195 machining chips. The fully equiaxed recrystallized microstructure in extruded wire was observed by optical microscopy. But, neither the trend of compression properties nor hardness test has been discovered. Abu-Farha [5] applied the derivative of FE which is called frictional stir back extrusion to produce sounded, defect-free aluminum alloy AA6063-T52 tube. A rotating tool is plunged into aluminum round bar specimens at a selected feed rate, forcing the processed material radially outward and thus forming tubes. Khorrami and Movahedi [6] fabricated the aluminum tubes by using a novel process of friction stir back extrusion and also investigated in detail the microstructural evolution as well as the mechanical properties of the formed tube. Behnagh et al. (2014), produced fully consolidated wires from aluminum alloy AA7277 machining chips by the friction stir extrusion. Components were stationary cartridge and a rotating plunger with a scroll-faced head. Hangai et al. [7] fabricated the AA1050 tubes filled with ADC12 foam by frictional stir back extrusion and foam able precursor was coated on the surface of the inner wall of the fabricated tube.

Babu et al. [8] carried out the hardness properties of friction stir processed AZ31 magnesium alloy wire at different processing conditions. The parameters were studied at the three different traveling speeds: 20, 30, and 40 mm/min with two different rotation speeds 600 and 800 rpm. The smallest precipitate size is obtained at spindle speed and feed rate combination of 800 rpm and 40 mm/min. Mechanical characterization and manufacturing of square bar was made of aluminum scraps through friction stir back extrusion process using a conventional lathe machine [9]. Baffari et al. [10] investigated the effect of process parameters and initial temper of the base metal on the wire quality by frictional stir extrusion. They considered rotational speed and vertical applied force. Extrusions were carried out with varying rotational speed (100, 200, and 300 rpm) and extrusion force (26.7, 35.6, 44.5, and 53.4 kN). Die and tool machined from H-13 tool steel and billet AA2050 chip and best sounded wire fabricated by 35.9 kN and 100 rpm are optimum. Although many researchers have done on frictional stir welding, frictional stir extrusion and channeling, some researchers experimentally proved frictional stir back extrusion by rotating the die and fixed chamber to produce tube, wire, rod, and c-channel. But no one applied frictional stir back extrusion to produce hexagonal-shaped bar which was proposed in this work. No work has been reported in the area of studying mechanical properties like hardness and compression strength of aluminum alloy AA6063 friction extruded bar, which was carried out in the present work.

61.2 Materials and Methodologies

The materials used to perform this work are as follows

- H-13 tool steel die and chamber
- Aluminum alloy billet
- 210 mm length M16 × 2 bolts
- Three-jaw chuck
- Rotary table
- T-headed bolt

The equipment to facilitate the friction stir back extrusion (FSBE) to produce hexagonal bar were adjustable wrench, chuck jaw Allen key, file and silicon carbide paper, small hammer, and box wrench. Table 61.1 shows the equipment and their specification used for the FSBE.

Universal Milling Machine

In the below Table 61.2, the maximum and minimum torque of milling machine was calculated by keeping other parameters constant. The specification of the milling machine was as indicated in Table 61.2.

Table 61.1 Machinery used for experiment and sample test

No.	Equipment	Tasks	Machine specification
1.	Universal milling machine	To fabricate hexagonal bar and produce key way in chamber	Excel vertical milling machine, FU-281
2.	CNC milling machine	Produce spiral in face of die	Fanuc series Oi Mate-MB
3.	Lathe machine	Turn, taper turn, face, threaded out, and bored of chamber and die	Excel XL-420
4.	Forging machine	Make hexagonal-shaped hole at center of die	Forge machine new 2010 version
5.	Drilling machine	To make hole in die	Universal drill TSAR-25
6.	Heat treatment machine	Heat treat both die and chamber up to HRC-55	Fours nagat 44550-montoir de Bretagne-France
7.	Universal test machine (UTM)	Compression strength of sample	GUNT WP 310 hydraulic
8.	Optical microscopy	Microstructural test	Magnification X50, X100, X200, X500, and X1000
9.	Vickers hardness tester	Hardness test	IS 1754/2002

Table 61.2 Specification of universal milling machine code F-281 (Excel machine manual)

No.	Spindle description	Units	Maximum range
1.	Main-driven range normally	RPM	40–2000
2.	Motor power	kW	5.5
3.	Maximum torque	Nm	1313.7
4.	Minimum torque	Nm	26.27
5.	Feed motor power	kW	1.5
6.	Feed motor rotational speed	RPM	920
Feed ranges			
	Longitudinal transverse	mm/min	12.5–630
		″	″
7.	Vertical, horizontal, and transverse.	″	5.2–262
8.	Gear transmission ratio	–	1:1

61.2.1 Design and Fabrication of Die and Chamber

Geometrical Design of Die and Chamber
The 3D modeling was made before manufacturing die and chamber. Figure 61.1a, b shows the surface of chamber and scroll-faced die with hexagonal-shaped center hole.

Chemical Composition of Aluminum Alloy Billet
The chemical composition of aluminum alloy billet was tested using spectrometer instrument. The obtained chemical composition of billet is shown in Table 61.3 and identified as AA 6063 alloy.

Figure 61.2 shows the scrap of aluminum alloy which was used for the FSBE to produce the hexagonal bars.

61.2.2 Process Parameters of Milling Machine to Fabricate Hexagonal-Shaped Bar

The process parameter like die rotational speed and plunge rate were varied by keeping die, chamber, plunge diameters, and contact land length of die parameters constant. The selected rotational speed was 200, 250, 315, 400, 500, and 630 rpm and with transverse speed 12.5 mm/min as shown Table 61.4. The range of rotational speed and plunge rate were considered based on previous research work [2]. The process parameters of frictional stir extrusion were optimized and decided the

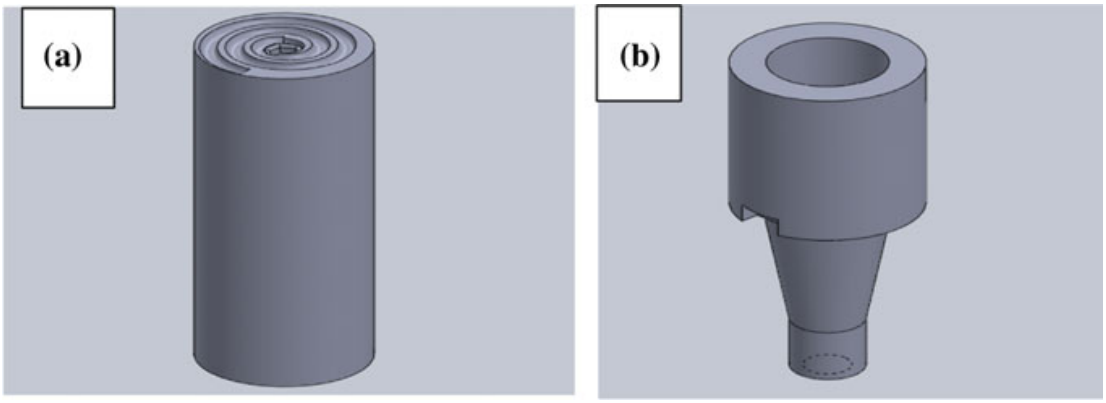


Fig. 61.1 Geometrical design of die (**a**) and chamber and (**b**) by solid work

Table 61.3 Chemical composition of base material (wt%)

Si	Fe	Cu	Mn	Mg	Zn	Sn	Cr	Pb	Ni	Al
0.58	1.35	0.017	0.014	0.7	0.092	0.039	0.033	0.015	0.019	97

Fig. 61.2 Aluminum alloy scrap chips

Table 61.4 Friction stir back extrusion process parameters

Test number	Die rotational speed (rpm)	Transverse speed of milling table (mm/min)
1	200	6.25
2	250	6.25
3	315	6.25
4	400	6.25
5	500	6.25
6	630	6.25

plunge rate 14 mm/min, hole size 5 mm, and initial contact land length was 2 mm to produced sounded wire [7].

Within the above given rotational speed and transverse speed, the experiment was performed.

61.2.3 Consolidation and Extrusion Period of Time

Friction consolidation is the period of time that compacted the scrap material with in cavities of chamber and maximized density of chips and scrap. The combination of high temperature and plastic deformation to be made the scrap material highly compacted. Friction extrusion is the whole process that includes consolidation, stirring, and extrusion by using scroll- or spiral-shaped die feed to the center plasticized material. The local temperature of extruded material increased up to 0.6–0.9 T_m [11]. In this experiment, the transverse feed was given by manually up to consolidate aluminum alloy scrap by above 0.6 T_m and engaged automatic feed

after consolidate of material. The extrusion period was taken 1.28 min to each sample using automatic feed.

61.2.4 Fabrication of Hexagonal Bar

Specimen Fabrication Procedure

To produce sample bar, the prepared scrap was compacted within cavities of chamber using small-sized ball pen hammer. This pre-compact process was used to increase the length of output hexagonal bar. Then the compacted chips and scrap were taped by a circular disc 40 mm diameter and 5 mm thickness. The taper disc was made from the same aluminum alloy the chips made of to hold the scrap and increases friction between scroll-faced die up to consolidate. It was turned by lathe machine at high-speed steel cutter. To produce hexagonal sample bars, the compacted chamber with scrap was installed on the spindle of milling machine and also die was fixed milling machine table with proper alignment as indicated in Fig. 61.3. The consolidation time depends on rotational speed of chamber. Then after chips are plasticized automatic transverse gear engaged. In this process the machine table moved up 6.25 mm distance per one minute. Then after the sample bar extruded, the die was removed by engaging the clash downward to change feed direction.

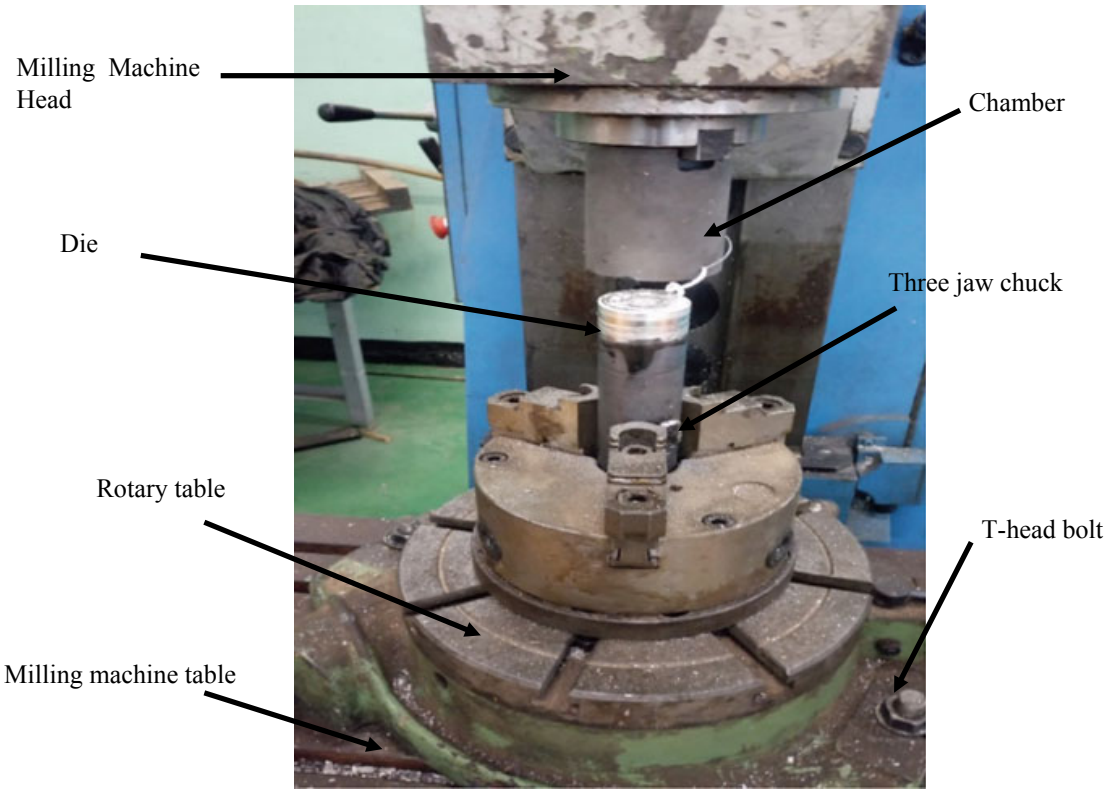


Fig. 61.3 Setup of friction stir back extrusion process on milling machine

61.3 Results and Discussion

61.3.1 The Fabricated Die and Chamber

The scroll-faced die and one side cavity chamber are the basic equipment in friction stir back extrusion (FSBE). The function of this die and chamber is to consolidate, stir, and extrude aluminum scrap. To resist wear occurrence in the process of stirring and hexagonal-shaped sample bar production, it should be heat treated. The color looks like black after heat treatment as Fig. 61.4.

61.3.2 Extruded Samples at Different Speeds and Discussion

Figure 61.5 shows the samples fabricated at different rotational speed and their shapes at each rotational speed. For example, sample one is extruded at 630 rpm and 6.25 mm/min, sample two 500 rpm and 6.25 mm/min, etc.

Figure 61.5 shows the possibilities of FSBE to fabricate hexagonal-shaped bar and other shapes from aluminum alloy scrap. Most of extrudes twisted and circular shapes. The reason of twist is, in FSBE die shape, machine power and extruded material properties affects the shape, length, and mechanical properties of extrudes. Sample one is performed at the rotational speed of 630 rpm and transverse speed of 6.25 mm/min. In this process parameter, the extruded product is resulted with large twisting angle almost like a circular rod shaped. This sample is longer than the other and is about 280 mm in length and circular in shape due to higher rotational speed. The rotational speed played an important role and contributed 71.6% to the overall extrusion parameters. If the rotational speed increases, the possibility of consolidated aluminum to be changed into liquid form is high. In this time, the die land length must be done longer than the other because it takes few second to solidify for melted aluminum. To this work, the die center hole depth (hexagonal shape former) was 2 mm. the much-plasticized aluminum extrude out immediately by not formed hexagonal shape. It made circular shape due to high rotational speed and shortage of depth of hexagonal shape formed hole.

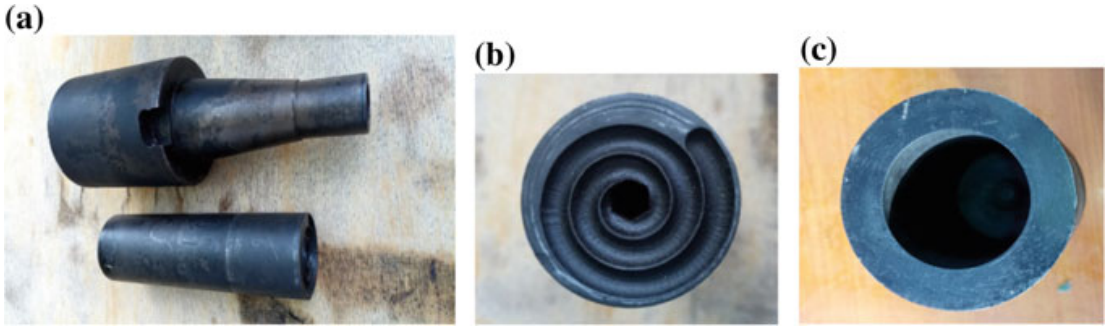


Fig. 61.4 Surface of die and chamber (**a**), spiral face of die, (**b**) and cavities of chamber (**c**)

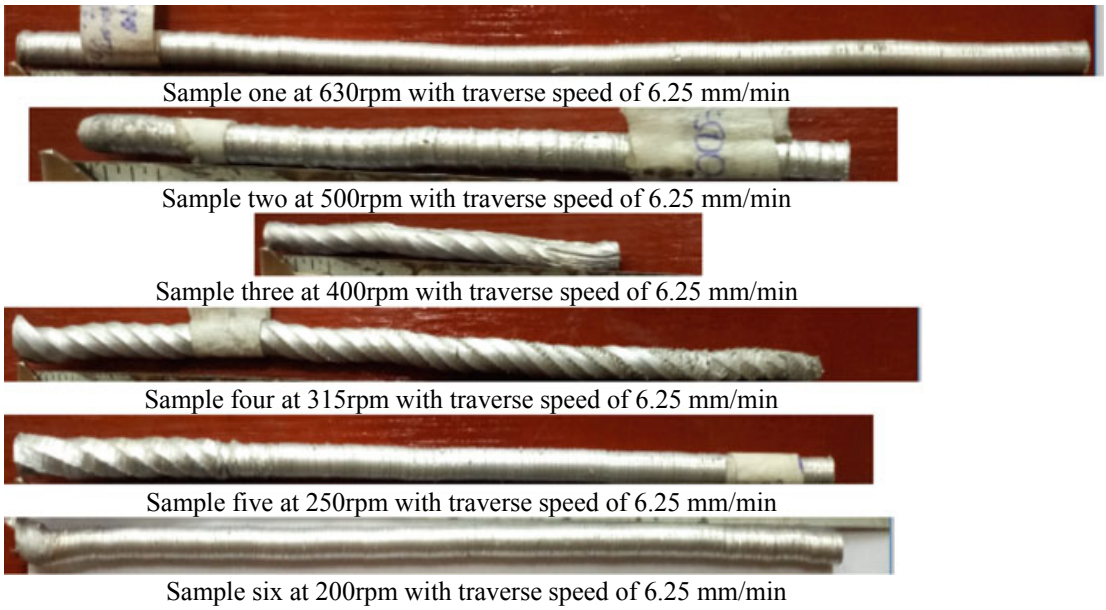


Fig. 61.5 Samples extruded by FSBE at different rotational speed

Sample two also almost the same shape to sample one, it is some hexagonal form related with sample one. It has been performed by rotational speed of 500 rpm and constant transverse speed 6.25 mm/min. The consolidate time took 5 min and extrusion time 1.28 min and the result is shown in Fig. 61.5 as 10.5 mm length and 6 mm diameter extrudes by friction stir extrusion. As explained above in sample one, rotational speed changed also resulted in shape change of the extruded products. The solution to this problem is design and manufacture die by deepened on each rotational speed of chamber. And also die center shape formed hole depth varied (means above 2 mm). Similarly, the other samples are manufactured at different process parameters as shown in Table 61.4.

Conclude to all the above discussion, the optimum rotational speed was 250 rotational speed and 6.25 mm/min transverse speed. In the case of surface smoothness, the optimum rotational speed was 325 rpm and the consolidation time also affects the physical properties of sample or hexa-bar (Table 61.5).

Table 61.5 Duration of scrap consolidation and extrusion time at each rotational speed

Sample No.	Rotational speed	Transverse speed (mm/min)	Consolidation time (min)	Extrusion period (min)
1.	630	6.25	5	1.28
2.	500	6.25	7	1.28
3.	400	6.25	7	1.28
4.	315	6.25	8	1.28
5.	250	6.25	9	1.28
6.	200	6.25	10	1.28

61.4 Conclusions

Defect-free sounded hexagonal-shaped bar has been fabricated by using FSBE process from aluminum alloy 6063 scrap. The considered process parameters are constant transverse speed 6.25 mm/min and variable rotational speed (200, 250, 315, 400, 500, and 630 rpm) of milling machine FU-281. From this interval of rotational speed, the accurate hexagonal-shaped bar obtained at 250, 315, and 400 rpm and all others are completely circle.

References

1. Li, X., Tang, W., Reynolds, A.P., Tayon, W.A., Brice, C.A.: Strain and texture in friction extrusion of aluminum wire. J. Mater. Process. Technol. **229**, 191–198 (2016)
2. Hosseini, A., Azarsa, E., Davoodi, B., Ardahani, Y.: Effect of process parameters on the physical properties of wires produced by friction extrusion method. Int. J. Adv. Eng. Technol. **3**(1), 592 (2012)
3. Behnagh, R.A., Mahdavinejad, R., Yavari, A., Abdollahi, M., Narvan, M.: Production of wire from AA7277 aluminum chips via friction-stir extrusion (FSE). Metall. Mater. Trans. B **45**(4), 1484–1489 (2014)
4. Tang, W., Reynolds, A.P.: Production of wire via friction extrusion of aluminum alloy machining chips. J. Mater. Process. Technol. **210**(15), 2231–2237 (2010)
5. Abu-Farha, F.: A preliminary study on the feasibility of friction stir back extrusion. Scr. Mater. **66**(9), 615–618 (2012)
6. Khorrami, M.S., Movahedi, M.: Microstructure evolutions and mechanical properties of tubular aluminum produced by friction stir back extrusion. Mater. Des. **1980–2015**(65), 74–79 (2015)
7. Ansari, M.A., Behnagh, R.A., Narvan, M., Naeini, E.S., Givi, M.K.B., Ding, H.: Optimization of friction stir extrusion (FSE) parameters through Taguchi technique. Trans. Indian Inst. Met. **69**(7), 1351–1357 (2016)
8. Babu, J., Anjaiah, M., Mathew, A.: Experimental studies on friction stir processing of AZ31 Magnesium alloy. Mater. Today Proc. **5**(2), 4515–4522 (2018)
9. Gelaw, M., Ramulu, P.J., Hailu, D., Desta, T.: Manufacturing and mechanical characterization of square bar made of aluminium scraps through friction stir back extrusion process. J. Eng. Des. Technol. **16**(4), 596–615 (2018)
10. Baffari, D., Reynolds, A.P., Li, X., Fratini, L.: Influence of processing parameters and initial temper on friction stir extrusion of 2050 aluminum alloy. J. Manufact. Proc. **28**, 319–325 (2017)
11. Khodabakhshi, F., Nosko, M., Gerlich, A.P.: Effects of graphene nano-platelets (GNPs) on the microstructural characteristics and textural development of an Al–Mg alloy during friction-stir processing. Surf. Coat. Technol. **335**, 288–305 (2018)

Chapter 62
Fabrication, Characterization and Evaluation of Mechanical Properties of Aluminium Hybrid Matrix Composite (Al6063/SiC-Bagasse Fly-Ash)

Getachew Gashaw, Perumalla Janaki Ramulu and Ch. Venkatesh

Abstract The fabrication and mechanical performance of Al-Mg-Si alloy matrix composites reinforced with silicon carbide (SiC) and bagasse fly-ash has been explored in this work. The fabrication and experimental investigation of mechanical properties of aluminium matrix (Al/SiC) composite with and without addition of bagasse fly-ash by varying wt% of reinforcements were investigated. Chemical composition of reinforcement, weighted required amount of matrix and reinforcements, powder mixing, stirring the slurry at 500 rpm speed and 5–10 min were the parameters used to develop the Al/SiC-bagasse fly-ash hybrid composite. Bagasse fly-ash particulates added with 7.5, 10, and 12 wt% SiC were utilized to prepare 10 and 12 wt% of the reinforcing phase with Al-Mg-Si alloy as matrix using stir casting method. The cast samples were machined for evaluating the properties by performing tensile testing, compression strength testing, and micro-hardness measurement. The results disclosed that the hardness of the hybrid composites increased with increase in bagasse fly-ash content with a maximum increment of 17% observed for the Al6063-7.5% SiC-10% bagasse ash composition (in comparison with the Al-10 wt% SiC-7.5 fly-ash reinforced composition). Tensile strength reduced of maximum value of 90–68 MPa at 7.5 SiC-7.5 bagasse ash compositions.

G. Gashaw · Ch.Venkatesh
Program of Mechanical Design and Manufacturing Engineering, School of Mechanical, Chemical and Materials Engineering, Adama Science and Technology University, Adama, Ethiopia
e-mail: gech21912@gmail.com

Ch.Venkatesh
e-mail: venkymech0607@gmail.com

P. J. Ramulu (✉)
Program of Mechanical Design and Manufacturing Engineering & Director of CoEAME, School of Mechanical, Chemical and Materials Engineering, Adama Science and Technology University, Adama, Ethiopia
e-mail: perumalla.janaki@astu.edu.et

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_62

Keywords SiC · Bagasse fly-ash · Stir casting · MMCs · Al/SiC

62.1 Introduction

Aluminium matrix composites (AMCs) have significant properties than other metal matrix composites with respect to high strength to weight ratio. The mechanical properties and wear resistance were improved by using SiC, TiC, TiB2 and fly-ash particulates to reinforce Al alloys [1]. Among all reinforcements, SiC was chemically compatible with aluminium and forms an adequate bond with the matrix without developing inter-metallic phase and had other advantages such as excellent thermal conductivity, good workability and low cost revealed in the literature [2]. In recent years, the use of fly-ash as reinforcement material in Al alloys was reported to be desirable from both environmental and economic significance due to its low cost. To overcome this, many research and development programmes were focused on Al metal matrix composites using low-cost industrial waste by-products as the reinforcement particulate [3]. Investigation of mechanical behaviour of aluminium alloy reinforced by nano- and micro-hard particles such as Al_2O_3 and SiC was a motivating area of research [4]. The present work, Aluminium hybrid matrix composite with reinforcements such as, SiC particle and waste of sugar cane bagasse fly-ash particle. Al/SiC-bagasse fly-ash composite provides more wear resistance, high coefficient of thermal expansion and high strength. Silicon carbide (SiC) is hard, wear resistance, high heat resistance ceramic material. With respect to low cost, less density and availability bagasse fly-ash is one of them which has inexpensive, low density reinforcement and available in large quantities by-product during combustion of coal in thermal power plants as solid waste. By which, composite material made with fly-ash as reinforcement are likely to overcome the cost barrier for extensive applications in automotive industries.

Metal matrix hybrid composites are the second-generation composite materials that have potential to substitute single-reinforced composite by improving properties [5, 6]. Vijayakumar et al. [7] studied the tensile strength and hardness of Al6061/SiC composite by varying the wt% SiC reinforcements 3–18%. The composite was manufactured by using stir casting of 600 rpm, 5 min stirring speed, and stirring time, respectively. The ultimate tensile strength and hardness results showed that the tensile strength and hardness of the developed metal matrix improved during increasing the wt% of SiC. Imran et al. [8] fabricated Al7075/bagasse ash–graphite of hybrid metal matrix composite (HMMC) by stir casting method. Reinforcements were added with three cases: in first case, 2% constant of bagasse ash with varying graphite 1, 3 and 5%; in second case, 4% bagasse ash and in third case, 6% constant with varying graphite percentage. Results showed that 6% bagasse ash with varying graphite 1, 3 and 5% more enhanced mechanical properties achieved as compared to case one and two of bagasse–ash combination. The hardness, ultimate tensile strength and yield strength of the developed composites were gradually increased with increasing the wt% of reinforcement, while

ductility decreased. Similar studies were also carried out on MMCs and their characterizations with respect to their mechanical behaviour [9, 10]. Failure prediction of FRP laminates was studied for a new generation of materials to an aerospace application [11]. The present work significance is natural fibres have good mechanical property, but a finite resource. The usage of their wastage, such as fly-ash and plant wastes from industries as reinforcement, is recommendable. Bagasse fly-ash is a waste by-product of coal combustion in thermal power plants which is available in large quantities in Ethiopian sugar factories. The usage of this waste fly-ash particle reinforced into aluminium alloys, which can effectively enhance the various mechanical properties of the aluminium alloy. Automotive heavy components fabricated from metals, Due to an increase in the weight of components, consumption of fuel increases. Increasing volume percentage of SiC reinforcement the weight of the composite increase. Bagasse fly-ash leads to decrease in the density of the composite when using some percentage with silicon carbide as reinforcement for developing Al/SiC/fly-ash composite.

62.2 Experimental Methodology

62.2.1 Matrix and Reinforcement Materials

Base material-AA 6063 alloy

The Al6xxx alloys were chosen because of its wide use in automotive components and architectural materials, including pistons, disk rotors, cylinder liners and decorative parts, rivets, etc. AA 6063 was the most widely used because of its extrudability and used for the automotive space frame parts [12].

Reinforcement material–Silicon carbide (SiC)

Silicon carbide material was chosen as a reinforcement material because it is used in high hardness and high thermal resistance applications, such as automotive brakes, cylinders, clutches, valves and bullet proof vests, semiconductor electronics etc. It is also applicable in abrasive cutters, cutting tools, grinding wheels, nuclear reactors due to high temperature resistance, high thermal conductivity and low nuclear activation [13]. The grain size of silicon carbide varies micro to nano sizes. One way of improving mechanical properties of composite is reducing grain size of reinforcements depends on processing methods. The very small sizes of powder particle of SiC create agglomeration (mud) during casting of poor stirring. To overcome such problems, proper stirring action is used from 5–10 min. Figure 62.1 shows the micron-level silicon carbide powders which are used as a reinforcement ceramic material to fabricate Al/SiC-bagasse fly-ash hybrid composite.

Bagasse fly-ash

Bagasse is produced from sugarcane plant during processing in sugar production. For the present work, the bagasse ash was taken from Wonji Shewa Ethiopian

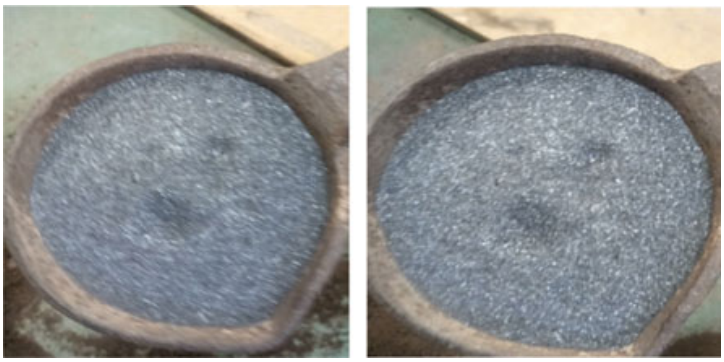


Fig. 62.1 Silicon carbide powders

Sugar Factory. Bagasse fly-ash was chosen due to low weight, density and increase in the strength. The ash was manually crushed and screened using sieves to make fine powder particles.

62.2.2 Methods

Physical Properties and chemical composition of bagasse fly-ash.

Density—1.96 g/cc and the specific gravity of fly-ash varies in the range of 0.6–2.8 gm/cc. Table 62.1 shows the chemical composition of the bagasse fly-ash by chemical analysis. The chemical composition analysis was performed on Dangote cement industries in Ethiopia.

The sugarcane bagasse ash is rich in SiO_2, which is good to replace SiC and used as reinforcement for metal matrix composites. In [14] research work, it was reported that alumina (Al_2O_3) also promise reinforcement material for production of metal matrix composite.

Powders mixing

Al-5% volume of SiC powders are mixed by stirrer on milling machine. The calculation for mass of aluminium, silicon carbide and bagasse fly-ash powder is needed for a sample of 130 mm length and 15 mm diameter. Pre-hearing of reinforcements from 600–700 °C of temperature before added in molten metal. Weight of powder added into matrix depends on the sample size. The blade of stirrer assembled on rotating part of milling machine speed of 500 rpm, 5–10 min stirring speed and time respectively, used to mix matrix and reinforcements.

Table 62.1 Chemical composition of bagasse ash in weight percentage

Composition wt%	SiO_2	Al_2O_3	Fe_2O_3	CaO	MgO
	%	%	%	%	%
	66.55	10.25	5.48	3.26	1.17

Mould preparation and fabrication the composite
To facilitate the easy removal of compacted specimen, tapered section has been used. The size of the mould depends on the size of samples which were used for tests. The required amount of reinforcements and matrix for each sample weighed and added in the crucible. Three samples were manufactured on one mould with the same composition of reinforcement. This is for the reputation of three experiments.

62.2.3 Sampling and Sample Method

In total, 8 number of experiments were conducted with three reputations for each type of mechanical testes. The wt% composition of reinforcements was taken as per available literatures, which has highest improvement in mechanical properties. The composition was 7.5, 10 wt% of SiC and 7.5, 10, 12 wt% of bagasse fly-ash and form eight combination of experiments. For specimen preparation with different percentage, composition of reinforcements manufactured was shown in Tables 62.2 and 62.3.

62.2.4 Equipment Used for This Work

Table 62.4 shows the equipment used for testing mechanical properties of the developed composite. Casting set-up was used to melt the matrix aluminium alloy during fabrication of Al/SiC–bagasse ash composite.

Table 62.2 Composite content in wt% of reinforcements

S. No	wt% of silicon carbide (SiC)	wt% of bagasse fly-ash	Al/SiC/bagasse fly-ash
1	7.5	7.5	Al6063-7.5% SiC-7.5% bagasse ash
2	7.5	10	Al6063-7.5% SiC-10% bagasse ash
3	7.5	12	Al6063-7.5% SiC-12% bagasse ash
4	10	7.5	Al6063-10% SiC-7.5% bagasse ash
5	10	10	Al6063-10% SiC-10% bagasse ash
6	10	12	Al6063-10% SiC-12% bagasse ash

Table 62.3 Composite content in wt% of silicon carbide only

S. No.	wt% of Al	wt% of SiC	Al/SiC
1	92.5	7.5	Al6063-7.5% SiC
2	90	10	Al6063-10% SiC

Table 62.4 Equipment used for this work

S. No.	Equipment	Task (Test name)	Machine model
1	Universal test machine (UTS)	Tensile test	WP 310 HYDROLIC UTM-Gunt Hamburg
2	Casting set-up	To melt and stir materials	

62.2.5 Mould Preparation and Fabrication

The sand mould was prepared for fabrication of Al/SiC/fly-ash hybrid composite based on the size of required specimens. 30 mm diameter and 150 mm length, 20 mm diameter and 60 mm length, specimen sizes were prepared for tensile test and compression test, respectively.

Fabrication

The matrix aluminium melted in a crucible which is placed inside furnace. It was melted at a temperature of 700 °C and reinforcements added and stirring with a molten metal for 5–10 min. The stirrer blade was assembled on hand drilling machine.

62.2.6 Tensile Test Specimen

The tensile test specimens of metal matrix composites were prepared based on the ASTM Standard D 3552. The flat and round configurations of metal matrix specimens are adopted. If the composite produced by diffusion bonding, flat shape of specimen would produce. In other case, metal matrix composite produced by various liquid infiltrations and other methods used for producing massive materials is better suited to circular cross-sectional shapes.

62.3 Results and Discussion

62.3.1 Fabrication of Al/SiC/Bagasse Ash Composite

Figure 62.2 shows the manufactured samples removed from moulds which were ready for machining and prepare samples.

Machining
Machining is performed on laze machine using carbide cutting tool. The specimens prepared based on ASME standard to conduct tensile, compression and hardness. Figure 62.3 shows the specimens machined and prepared for tensile tests.

Fig. 62.2 Manufacturing of Al/SiC-bagasse ash composite

Fig. 62.3 Specimens for tensile and compression test

62.3.2 Mechanical Tests of the Al/SiC-Bagasse Ash Composite

Tensile test

Tensile strength of a material is the ability of resist loads which applied in the longitudinal direction of sample. Circular specimens were machined with a specific gage length, specimen diameter (D), notch diameter (d) and fillet radius (r) of 100,

Fig. 62.4 Failure of tensile test specimens

12, 7.5, and 9 mm, respectively. The tensile test was conducted on WP 310 computerized hydraulic universal testing machine (UTM) and the test specimens were shown in Fig. 62.4. The test was carried out on WP-300 hydraulic testing machine of 50 kN capacities with at a constant piston speed of 5 mm/min. Table 62.5 presented the complete test results of tensile data. The compared results are between single-reinforced composite (Al/SiC) and hybrid metal matrix composite (Al/SiC-bagasse ash) in addition to the base metal (Al6063). The variation of average tensile strength in each single and hybrid composite is presented in Table 62.6. Tensile test results showed that increasing percentage of reinforcement (SiC) only in the matrix reduces average tensile strength while increasing reinforcements, SiC and bagasse ash improves the tensile strength of Al/SiC-bagasse ash composite. At 10% SiC-10% bagasse ash hybrid composite have better improvement than other single and hybrid composites.

Table 62.5 Tensile strength of the developed hybrid metal matrix composites

S. No.	Al/SiC/fly-ash composite	Ultimate tensile strength (MPa)			Average ultimate tensile strength (MPa)
		Test 1	Test 2	Test 3	
1	Al6063-7.5% SiC-7.5% bagasse ash	52.49	81.71	70.4	68.2
2	Al6063-7.5% SiC-10% bagasse ash	88.5	76.43	60.7	75.2
3	Al6063-7.5% SiC-12% bagasse ash	74.5	61.12	78.98	71.6
4	Al6063-10% SiC-7.5% bagasse ash	76.8	68.8	76.32	73.89
5	Al6063-10% SiC-10%bagasse ash	84.71	87.53	76.62	82.95
6	Al6063-10% SiC-12% bagasse ash	87.5	79.00	85.09	83.86

Table 62.6 Tensile strength of the developed Al/SiC matrix composites

S. No.	Al/SiC composite	Ultimate tensile strength (MPa)			Average ultimate tensile strength (MPa)
		Test 1	Test 2	Test 3	
1	Al6063-7.5% SiC	78.46	45.4	64.87	62.91
2	Al6063-10% SiC	74.24	65.34	47.5	62.36

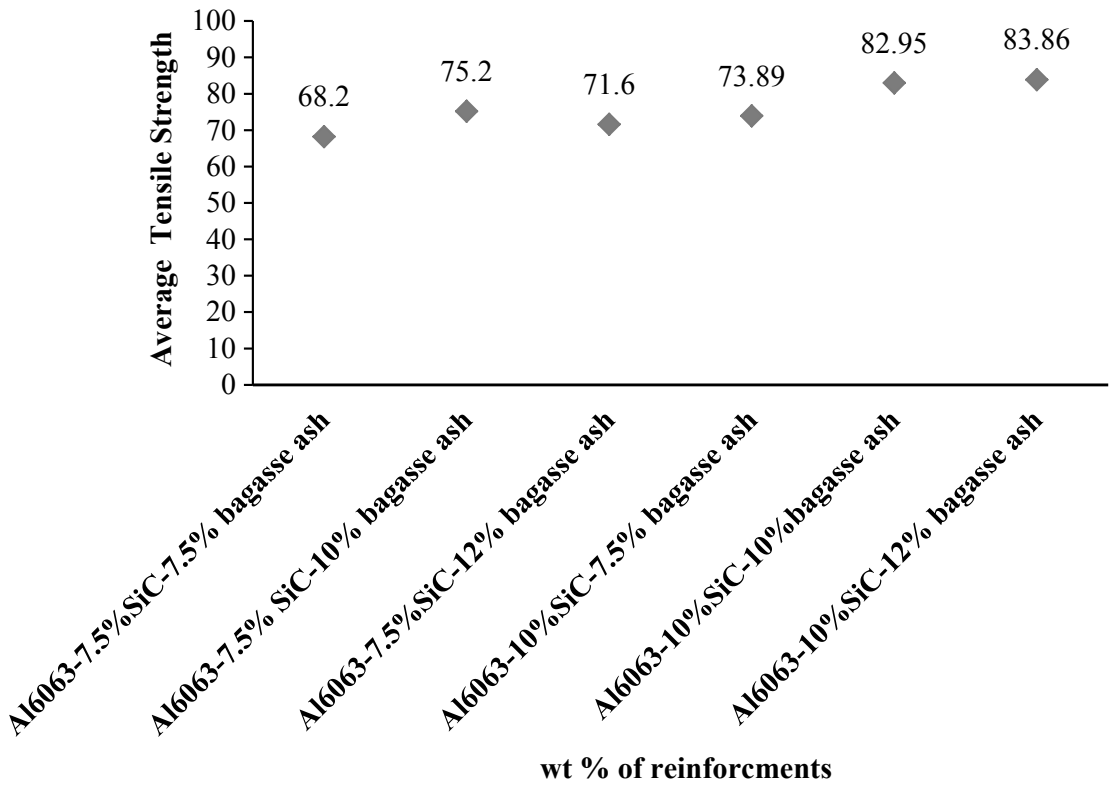


Fig. 62.5 Tensile strength and weight percentage of reinforcements

SiC and FA particles are very effective in improving the tensile strength of the composite. The addition of SiC and fly-ash particles in the matrix induces much strength to matrix alloy by offering more resistance to tensile stresses. It is well known that the thermal expansion coefficient of SiC and fly-ash particle is 3.25 × 106/°C and for aluminium alloy is 23 × 106/°C. The thermal mismatch between matrix and the reinforcement causes higher dislocation density in the matrix and load-bearing capacity of the hard particles which subsequently increases the composite strength. Figure 62.5 reveals the fracture surface of AA6063 10 wt% SiC 7.5 wt% fly-ash composite. It shows a network of dimples whose size is smaller compared to matrix alloy. The fly-ash and silicon carbide refined the grain size of matrix alloy and reduced the ductility which resulted in smaller size dimples.

62.4 Conclusions

The Al/SiC-bagasse fly-ash hybrid composite is successfully manufactured with different weight percentage of SiC and bagasse fly-ash reinforcements (7.5 and 10 wt% of SiC and 7.5, 10, and 12 wt% bagasse fly-ash). The mechanical properties of the developed composite are studied with conducting different tests.

The conclusions drawn from the present investigation are as follows:

- Bagasse fly-ash, the waste generated from sugarcane plant can be successfully used as a reinforcing material to produce metal matrix composite (MMC) component in aluminium matrix to be used in lightweight of automotive parts.
- The tensile strength, compression strength and hardness of the hybrid composite improved with increasing wt% of reinforcements.
- The hybrid Al/SiC-bagasse ash composite were better mechanical properties with compared single-reinforced Al/SiC composite.

References

1. Suresh, N., Venkateswaran, S., Seetharamu, S.: Influence of cenospheres of fly ash on the mechanical properties and wear of permanent moulded eutectic Al-Si alloys. Mater. Sci. Pol. **28**(1), 55–65 (2010)
2. James, S.J., et al.: Hybrid aluminium metal matrix composite reinforced with SiC and TiB2. Proced. Eng. **97**, 1018–1026 (2014)
3. Kala, H., Mer, K.K.S., Kumar, S.: A review on mechanical and tribological behaviors of stir cast aluminum matrix composites. Proced. Mater. Sci. **6,** 1951–1960 (2014)
4. Sajjadi, S.A., Ezatpour, H.R., Parizi, M.T.: Comparison of microstructure and mechanical properties of A356 aluminum alloy/Al_2O_3 composites fabricated by stir and compo-casting processes. Mater. Des. **34**, 106–111 (2012)
5. Singh, J., Chauhan, A.: Characterization of hybrid aluminum matrix composites for advanced applications—a review. J. Mater. Res. Technol. **5**(2), 159–169 (2016)
6. Bodunrin, M.O., Alaneme, k.k., Chown, L.H.: Aluminium matrix hybrid composites: a review of reinforcement philosophies; mechanical, corrosion and tribological characteristics. J. Mater. Res. Technol. **4**(4), 434–445 (2015)
7. Vijayakumar, S., et al.: Studies on mechanical properties of Al-SiC metal matrix composite. SSRG-IJMSE **2**(Nov–Dec), (2013)
8. Imran, M., et al.: Study of hardness and tensile strength of Aluminium-7075 percentage varying reinforced with graphite and bagasse-ash composites. Res. Effic. Technol. **2**(2), 81–88 (2016)
9. Alaneme, K.K., et al.: Fabrication characteristics and mechanical behaviour of rice husk ash—Alumina reinforced Al–Mg–Si alloy matrix hybrid composites. J. Mater. Res. Technol. **2**(1), 60–67 (2013)
10. Alaneme, K.K., Bodunrin, M.O., Awe, A.A.: Microstructure, mechanical and fracture properties of groundnut shell ash and silicon carbide dispersion strengthened aluminium matrix composites. J. K. Saud Univ. Eng. Sci. (2016)

11. Rao, H.J., Ramulu, P.J., Vardhan, M.V., Chandramouli, C.H.: Failure prediction in fiber metal laminates for next generation aero materials. In: IOP Conference Series: Materials Science and Engineering, vol. 149, No. 1, p. 012102. IOP Publishing, September 2016
12. Oyinbo, S.T., Ikumapayi, O.M., Ajiboye, J.S., Afolalu, S.A.: Numerical simulation of axisymmetric and asymmetric extrusion process using finite element method. Int. J. Sci. Eng. Res. **6**(6), 1246 (2015)
13. Lorrette, C., Réau, A., Briottet, L.: Mechanical properties of nanostructured silicon carbide consolidated by spark plasma sintering. J. Eur. Ceram. Soc. **33**(1), 147–156 (2013)
14. Baradeswaran, A., Perumal, A.E.: Study on mechanical and wear properties of Al 7075/Al_2O_3/graphite hybrid composites. Compos. Part B Eng. **56**, 464–471 (2014)

Chapter 63
Fuzzy-EDAS (Evaluation Based on Distance from Average Solution) for Material Selection Problems

Divya Zindani, Saikat Ranjan Maity and Sumit Bhowmik

Abstract This present work proposes a novel fuzzy-based multi-criteria decision-making (MCDM) approach for addressing material selection problem under uncertain environment. The integrated framework integrates fuzzy decision making with evaluation based on distance from average solution (EDAS) MCDM method. Proposed framework fulfills the gap of restricted exploration of fuzzy decision-making tools in the domain of material design. The applicability of the proposed framework has been evidenced through a case study of a material selection of the casting plate used in sand casting. The selection was done considering five material alternatives and seven criteria. The presented fuzzy-based approach considers the objective weights while calculating the weights of the considered criteria.

Keywords Soft computing · Fuzzy theory · Material selection · Multi-criteria decision making

63.1 Introduction

Decision-making problems of the real world are complex and have to be solved under the influence of conflicting criteria. The decision-making problems when solved considering single criterion results in unrealistic solutions [1]. Therefore, in order to obtain suitable solutions, models for such problems are constructed considering multiple criteria. Systematic and scientific tools in the form of multi-criteria

D. Zindani (✉) · S. R. Maity · S. Bhowmik
Department of Mechanical Engineering, National Institute of Technology Silchar, Silchar 7888010, Assam, India
e-mail: divyazindani@gmail.com

S. R. Maity
e-mail: saikat.jumtech@gmail.com

S. Bhowmik
e-mail: bhowmiksumit04@yahoo.co.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_63

decision-making (MCDM) tools have been developed to deal with real-world problems considering multiple factors and criteria [2].

Material selection is a decision making problem that involves the selection of the best material from amongst the number of material alternatives. The rankings are obtained under the influence of multiple criteria that are often conflicting in nature. Therefore, material selection problems have been addressed with different MCDM tools. As for instance, the applicability of different preference ranking-based MCDM methods, i.e., operational competitiveness rating analysis (OCRA), organization, arrangement Et synthesize de donors relationnelles (ORESTE), grey system theory-based COPRAS (COPRAS-G) and extended preference ranking organization method for enrichment evaluations (EXPROM2) was evidenced by Chatterjee and Chakraborty [3] through the gear material selection problem. Rastogi et al. [4] employed the TOPSIS method for obtaining the final rankings of material for the phase change material. Material selection of brake friction material was addressed by Singh et al. [4] using combined VIKOR and AHP methodology. TODIM was applied by Zindani et al. [5] for the material selection problems of flywheel and spur gear. Evaluation based on distance from average solution (EDAS) with the design of experiments (DOE) was used recently by Chatterjee et al. [6] and the applicability was justified using gear material selection problem.

However, the problem becomes relatively more complex because of the associated uncertainty. The decision-making process then becomes relatively more challenging. In order to deal with qualitative criteria in material selection problems, fuzzy logic has been employed by the decisionmakers. Simplified fuzzy logic was used for dealing with qualitative properties of materials by Khabbaz et al. [7]. The applicability of the proposed methodology was demonstrated using material selection problems of the liquid nitrogen storage tank, the spar of an aircraft wing, and mast for a sailing boat. To deal with problems involving both the qualitative and quantitative properties, fuzzy-based MCDM models have been proposed and applied by the decisionmakers. Some of the examples include fuzzy-TOPSIS [8], fuzzy-VIKOR [9], etc.

EDAS method has been extended to solve under uncertain environments such as interval type-2 fuzzy sets [10], interval grey numbers [11], neutrosophic soft sets [12], interval-valued fuzzy soft sets [13], interval-valued neutrosophic sets [14], and intuitionistic fuzzy sets [15].

Although EDAS methodology has been evidenced for material selection problem, its applicability to address material selection problem under uncertain environment is yet to be explored. Therefore, the present work is a modest effort to fulfill this gap by proposing a fuzzy-EDAS methodology for material selection problems wherein triangular fuzzy numbers are used to represent linguistic information. The applicability of the proposed material selection framework has been demonstrated using material selection problem of casting plates used in sand casting. Rest of the paper has been organized into the following sections: basic concepts and arithmetic related to fuzzy set have been elaborated in Sect. 63.2. Section 63.3 presents a fuzzy-EDAS method to deal with material selection

problems under an environment of uncertainty. An illustrative example of material selection problem has been demonstrated in Sect. 63.4 of the paper. Paper finally terminates with concluding remarks in Sect. 63.5.

63.2 Fuzzy Theory and Some Definitions

Zadeh [16] developed a fuzzy set theory to handle the problems comprising vague information. Some of the basic definitions pertaining to the fuzzy theory that will be essential in the present work have been delineated in this section.

63.2.1 Definition Number 1

Let $\tilde{b}$ denote a fuzzy set. A membership function $\mu_{\tilde{b}}(x)$ characterizes the fuzzy set which maps x to a real number lying in the interval [0, 1]. Value of the function $\mu_{\tilde{b}}(x)$ is known as the grade of membership of x in $\tilde{b}$.

63.2.2 Definition Number 2

A fuzzy set $\tilde{a}$ is a triangular fuzzy number when $\tilde{b} = (b_1, b_2, b_3)$. Membership function $\mu_{\tilde{b}}(x)$ for a triangular fuzzy number $\tilde{b}$ is denoted by Eq. (63.1):

$$\mu_{\tilde{b}}(x) = \begin{cases} 0, x \leq b_1, \\ \frac{x-b_1}{b_2-b_3}, b_1 \leq x \leq b_2, \\ \frac{b_3-x}{b_3-b_2}, b_2 \leq x \leq b_3, \\ 0, x > b_3, \end{cases} \tag{63.1}$$

where, b_1, b_2, and b_3 denote real numbers such that $b_1 < b_2 < b_3$.

63.2.3 Definition Number 3

Main operations between two triangular fuzzy numbers $\tilde{b} = (b_1, b_2, b_3)$ and $\tilde{c} = (c_1, c_2, c_3)$ can be expressed as follows:

Addition $\tilde{b}$ (+) $\tilde{c}$ = $(b_1 + c_1, b_2 + c_2, b_3 + c_3)$
Subtraction $\tilde{b}$ (−) $\tilde{c}$ = $(b_1 - c_3, b_2 - c_2, b_3 - c_1)$

Multiplication $\tilde{b}\ (\times)\ \tilde{c} = (b_1 \times c_1, b_2 \times c_2, b_3 \times c_3)$
Division $\tilde{b}\ (/)\ \tilde{c} = (b_1 \times c_1, b_2 \times c_2, b_3 \times c_3)$

63.2.4 Definition Number 4

Operations between a triangular fuzzy number and a real number k are described as below:

Multiplication $k\ (\times)\ \tilde{b} = (kb_1, kb_2, kb_3)$
Division $\tilde{b}/k = (b_1/k, b_2/k, b_3/k)$
Division $k/\tilde{b} = (k/b_1, k/b_2, k/b_3)$

63.2.5 Definition Number 5

The crisp value of a triangular fuzzy number $\tilde{b} = (b_1, b_2, b_3)$ is defined as by Eq. (63.2):

$$k(\tilde{b}) = \frac{1}{6}(b_1 + 4b_2 + b_3) \tag{63.2}$$

63.2.6 Definition Number 6

The maximum of zero and a triangular fuzzy number can be established using Eq. (63.3):

$$\psi(\tilde{b}) = \begin{cases} \tilde{b} \text{ if } k(\tilde{b}) > 0 \\ \tilde{0} \text{ if } k(\tilde{b}) \leq 0 \end{cases} \tag{63.3}$$

63.3 The Conceptual Framework of Fuzzy-EDAS Methodology

EDAS methodology was developed by Keshavarz Ghorabaee et al. [17] and the method ranks material on the basis of the average solution. The average solution is arrived at by calculation of the positive distance from average (PDA) and negative

distance from average (PDA). The alternative that has higher PDA values and lower NDA values are the best-ranked material alternative.

In the case of fuzzy-EDAS methodology, alternatives are ranked in accordance with the decreasing value of defuzzified appraisal score. The steps involved in the fuzzy-EDAS method have been delineated in the ensuing discussion. Let $\tilde{A}$ denote the fuzzy decision matrix, i.e.,

$$\tilde{A} = \left[\tilde{a}_{ij}\right]_{n\times m} = \begin{bmatrix} \tilde{a}_{11} & \tilde{a}_{12} & \cdots & \tilde{a}_{1m} \\ \tilde{a}_{21} & \tilde{a}_{22} & \cdots & \tilde{a}_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{a}_{n1} & \tilde{a}_{n2} & \cdots & \tilde{a}_{nm} \end{bmatrix} \tag{63.4}$$

where j is the number of alternatives and i signify the criteria.

Following steps are traced in order to solve a decision-making problem using fuzzy-EDAS methodology.

63.3.1 Step 1

In this step of the proposed framework, objective weights are determined for each of the criteria Ci using Eqs. (63.5) and (63.6). Objective weights are determined for the decision matrix supplied by all the decisionmakers.

$$w_j^* = \frac{\sum_{i=1}^{m}\sum_{r=1}^{m}\left|p_{ij} - p_{rj}\right|}{\sqrt{\sum_{j=1}^{m}\left[\sum_{i=1}^{m}\sum_{r=1}^{m}\left|p_{ij} - p_{rj}\right|\right]^2}} \tag{63.5}$$

$$w_j^o = \frac{w_j^*}{\sum_{j=1}^{m} w_j^*} \tag{63.6}$$

63.3.2 Step 3

A fuzzy average decision matrix is developed with respect to all the considered criteria using Table 63.1 and Eq. (63.7):

$$\mathrm{AV}_j = \frac{\sum_{i=1}^{n}\tilde{a}_{ij}}{k} \tag{63.7}$$

Table 63.1 Linguistic terms for alternatives ratings

Linguistic term	Membership function
Very low (VL)	(0,0,0.1)
Medium low (ML)	(0, 0.1, 0.3)
Low (L)	(0.1,0.3, 0.5)
Medium (F)	(0.3, 0.5, 0.75)
High (H)	(0.5,0.75,0.9)
Medium High (MH)	(0.75,0.9,1)
Very high (VH)	(0.9, 1, 1)

63.3.3 Step 4

The distance of optimal solution from negative feasible solutions must be maximum whereas it should be minimum from the negative feasible solution. In this step of the fuzzy-EDAS methodology, matrices for fuzzy positive distance from average (PDA) and fuzzy negative distance from average (NDA) are calculated using Eqs. (63.8) and (63.9), respectively.

$$\mathrm{PDA} = \left[\mathrm{pda}_{ij}\right]_{n\times m} \tag{63.8}$$

$$\mathrm{NDA} = \left[\mathrm{nda}_{ij}\right]_{n\times m} \tag{63.9}$$

where for beneficial criteria

$$\mathrm{pda}_{ij} = \begin{cases} \dfrac{\psi\left(\tilde{a}_{ij} - \mathrm{AV}_j\right)}{k\left(\mathrm{AV}_j\right)} \end{cases} \tag{63.10}$$

and

$$\mathrm{nda}_{ij} = \begin{cases} \dfrac{\psi\left(\mathrm{AV}_j - \tilde{a}_{ij}\right)}{k\left(\mathrm{AV}_j\right)} \end{cases} \tag{63.11}$$

and for non-beneficial criteria

$$\mathrm{pda}_{ij} = \begin{cases} \dfrac{\psi\left(\mathrm{AV}_j - \tilde{a}_{ij}\right)}{k\left(\mathrm{AV}_j\right)} \end{cases} \tag{63.12}$$

and

$$\mathrm{nda}_{ij} = \begin{cases} \dfrac{\psi\left(\tilde{a}_{ij} - \mathrm{AV}_j\right)}{k\left(\mathrm{AV}_j\right)} \end{cases} \tag{63.13}$$

63.3.4 Step 5

In this step of the methodology, matrices are developed for fuzzy weighted positive and fuzzy weighted negative distances. Equations (63.14) and (63.15) are used for this purpose

$$\mathrm{sp}_i = \sum_{j=1}^{m} \left(\tilde{w}_j + \mathrm{pda}_{ij}\right) \tag{63.14}$$

and

$$\mathrm{sn}_i = \sum_{j=1}^{m} \left(\tilde{w}_j + \mathrm{nda}_{ij}\right) \tag{63.15}$$

63.3.5 Step 6

Fuzzy normalized values for fuzzy weighted positive and fuzzy weighted negative distances are computed using Eqs. (63.16) and (63.17)

$$\mathrm{nsp}_i = \frac{\mathrm{sp}_i}{\max_i(k(\mathrm{sp}_i))} \tag{63.16}$$

and

$$\mathrm{nsp}_i = 1 - \frac{\mathrm{sn}_i}{\max_i(k(\mathrm{sn}_i))} \tag{63.17}$$

63.3.6 Step 7

In the penultimate step, fuzzy appraisal score for different alternatives is calculated using Eq. (63.18)

$$\mathrm{as}_i = \frac{\mathrm{nsp}_i + \mathrm{nsn}_i}{2} \tag{63.18}$$

63.3.7 Step 8

In the last step, alternative materials are ranked in accordance with the decreasing value of defuzzified appraisal score. The best choice of the alternative material is the one with the highest value of appraisal score.

63.4 Illustrative Example

In this section, a case study on material selection of pattern used in sand casting has been solved using the above-proposed methodology. The material alternatives identified from literature survey [19] were wood (A_1), aluminum (A_2), steel (A_3), plastic (A_4), and cast iron (A_5). The criteria considered for the problem under consideration were machinability (C_1), wear resistance (C_2), strength (C_3), weight (C_4), reparability (C_5), resistance to corrosion (C_6), and resistance to swelling (C_7). Excluding weight, all the other criteria are benefited criteria. Three decisionmakers were asked to provide their performance rating. Table 63.2 depicts the performance rating for decisionmakers 1, 2, and 3. Each of the decision matrices was then converted to a triangular fuzzy number using information from Table 63.1. Table 63.3 depicts the fuzzy converted matrix for all the decisionmakers considered. Equations (63.4) and (63.5) were then used to obtain the objective weights for

Table 63.2 Performance ratings as given by decisionmakers

DM		C_1	C_2	C_3	C_4	C_5	C_6	C_7
1	A_1	VH	ML	M	VH	VH	VH	ML
	A_2	MH	MH	MH	MH	ML	M	VH
	A_3	M	VH	VH	ML	MH	ML	VH
	A_4	MH	M	MH	MH	M	VH	VH
	A_5	MH	VH	MH	ML	MH	ML	VH
2	A_1	VH	ML	M	VH	VH	VH	ML
	A_2	H	H	H	H	ML	M	VH
	A_3	M	VH	VH	ML	H	ML	VH
	A_4	H	M	H	H	M	VH	VH
	A_5	H	VH	H	ML	H	ML	VH
3	A_1	H	ML	M	H	VH	H	ML
	A_2	MH	MH	MH	MH	ML	M	VH
	A_3	M	VH	VH	ML	MH	ML	VH
	A_4	MH	M	MH	MH	M	VH	VH
	A_5	MH	H	MH	ML	MH	ML	H

Table 63.3 Triangular fuzzy numbers for performance ratings as obtained from different decisionmakers

		C_1	C_2	C_3	C_4	C_5	C_6	C_7
D_1	A_1	(0.9,1,1)	(0.1,0.3,0.5)	(0.3,0.5,0.75)	(0.9,1,1)	(0.9,1,1)	(0.9,1,1)	(0.1,0.3,0.5)
	A_2	(0.5,0.75,0.9)	(0.5,0.75,0.9)	(0.5,0.75,0.9)	(0.5,0.75,0.9)	(0.1,0.3,0.5)	(0.3,0.5,0.75)	(0.9,1,1)
	A_3	(0.3,0.5,0.75)	(0.9,1,1)	(0.9,1,1)	(0.1,0.3,0.5)	(0.5,0.75,0.9)	(0.1,0.3,0.5)	(0.9,1,1)
	A_4	(0.5,0.75,0.9)	(0.3,0.5,0.75)	(0.5,0.75,0.9)	(0.5,0.75,0.9)	(0.3,0.5,0.75)	(0.9,1,1)	(0.9,1,1)
	A_5	(0.5,0.75,0.9)	(0.9,1,1)	(0.5,0.75,0.9)	(0.1,0.3,0.5)	(0.5,0.75,0.9)	(0.1,0.3,0.5)	(0.9,1,1)
D_2	A_1	(0.9,1,1)	(0,0.1,0.3)	(0.3,0.5,0.75)	(0.9,1,1)	(0.9,1,1)	(0.9,1,1)	(0.1,0.3,0.5)
	A_2	(0.75,0.9,1)	(0.75,0.9,1)	(0.75,0.9,1)	(0.75,0.9,1)	(0,0.1,0.3)	(0.3,0.5,0.75)	(0.9,1,1)
	A_3	(0.3,0.5,0.75)	(0.9,1,1)	(0.9,1,1)	(0,0.1,0.3)	(0.75,0.9,1)	(0,0.1,0.3)	(0.9,1,1)
	A_4	(0.75,0.9,1)	(0.3,0.5,0.75)	(0.75,0.9,1)	(0.75,0.9,1)	(0.3,0.5,0.75)	(0.9,1,1)	(0.9,1,1)
	A_5	(0.75,0.9,1)	(0.9,1,1)	(0.75,0.9,1)	(0,0.1,0.3)	(0.75,0.9,1)	(0,0.1,0.3)	(0.9,1,1)
D_3	A_1	(0.9,1,1)	(0,0.1,0.3)	(0.3,0.5,0.75)	(0.9,1,1)	(0.9,1,1)	(0.9,1,1)	(0.1,0.3,0.5)
	A_2	(0.75,0.9,1)	(0.75,0.9,1)	(0.75,0.9,1)	(0.75,0.9,1)	(0,0.1,0.3)	(0.3,0.5,0.75)	(0.9,1,1)
	A_3	(0.3,0.5,0.75)	(0.9,1,1)	(0.9,1,1)	(0,0.1,0.3)	(0.75,0.9,1)	(0,0.1,0.3)	(0.9,1,1)
	A_4	(0.75,0.9,1)	(0.3,0.5,0.75)	(0.75,0.9,1)	(0.75,0.9,1)	(0.3,0.5,0.75)	(0.9,1,1)	(0.9,1,1)
	A_5	(0.75,0.9,1)	(0.9,1,1)	(0.75,0.9,1)	(0,0.1,0.3)	(0.75,0.9,1)	(0,0.1,0.3)	(0.9,1,1)

Table 63.4 Objective weights derived for different decisionmakers and the aggregated weights

	C_1	C_2	C_3	C_4	C_5	C_6	C_7
w_o (DM_1)	0.088	0.158	0.089	0.182	0.162	0.209	0.112
w_o (DM_2)	0.064	0.171	0.064	0.205	0.168	0.237	0.091
w_o (DM_3)	0.089	0.158	0.088	0.182	0.162	0.209	0.112
Aggregated weight	0.081	0.162	0.081	0.190	0.164	0.218	0.105

all the decision matrix obtained from the three decisionmakers. The aggregate weights were finally obtained by taking the average of all the objective weights for particular criteria. Table 63.4 depicts the individual objective weights and the aggregated objective weights. In the next step, the average decision matrix was formulated and has been tabulated in Table 63.5. Equation (63.7) was then used to arrive at the average solution which has been depicted in Table 63.6 which also shows the crisp value of the average solution obtained using Eq. (63.2). Positive distance from average (PDA) and negative distance from average (NDA) values were obtained using Eqs. (63.8–63.13) and Tables 63.7 and 63.8 represent these values. Fuzzy normalized values for fuzzy weighted positive and fuzzy weighted negative distances were computed using Eqs. (63.16) and (63.17). In the penultimate step, fuzzy appraisal score for different alternatives is calculated using Eq. (63.18). Finally, the alternatives were ranked on the basis of the defuzzified appraisal score. Table 63.9 portrays all these values. Aluminum was found to be the best alternative material.

63.5 Result and Discussion

Present work proposes a fuzzy-based integrated methodology in the form of fuzzy-EDAS for solving material selection problem. The proposed methodology aids the decisionmaker in addressing the linguistic performance rating of material alternatives with respect to the criteria. The presented material selection framework, therefore, aids in handling the complexity associated with the uncertain environment. However, the presented methodology only considers objective weights and does not take into consideration the subjective importance. To justify the applicability of the proposed method, material selection problem of casting pattern plate was solved. A total of five material alternatives and seven criteria were considered. Aluminum was found to be the best-ranked material alternative, whereas wood was the worst performing of all. As a future scope of the work, subjective weights could also be considered while arriving at the aggregated weights for each criterion. Further, quantitative properties in the decision matrix could also be given due consideration.

Table 63.5 Elements of the average decision matrix

	C_1	C_2	C_3	C_4	C_5	C_6	C_7
A_1	(0.85,0.967,1)	(0.067,0.233,0.433)	(0.3,0.5,0.75)	(0.85,0.967,1)	(0.9,1,1)	(0.850,0.967,1)	(1,0.3,0.5)
A_2	(0.583,0.8,0.933)	(0.583,0.80,0.933)	(0.583,0.80,0.933)	(0.583,0.80,0.933)	(0.067,0.233,0.433)	(0.3,0.5,0.75)	(0.9,1,1)
A_3	(0.3,0.5,0.75)	(0.90,1,1)	(0.9,1,1)	(0.067,0.233,0.433)	(0.583,0.783,0.933)	(0.067,0.233,0.433)	(0.9,1,1)
A_4	(0.583,0.80,0.933)	(0.3,0.5,0.75)	(0.583,0.80,0.933)	(0.583,0.8,0.933)	(0.3,0.5,0.75)	(0.9,1,1)	(0.9,1,1)
A_5	(0.583,0.80,0.933)	(0.85,0.967,1)	(0.583,0.80,0.933)	(0.067,0.233,0.433)	(0.583,0.783,0.933)	(0.067,0.233,0.433)	(0.833,0.967,1)

Table 63.6 Average solution matrix

	C_1	C_2	C_3	C_4	C_5	C_6	C_7
AV	(0.58,0.77,0.91)	(0.54,0.7,0.82)	(0.59,0.78,0.91)	(0.43,0.61,0.75)	(0.487,0.66,0.81)	(0.437, 0.59,0.72)	(0.72,0.85,0.9)
k (AV)	0.76	0.69	0.77	0.61	0.66	0.58	0.84

Table 63.7 Values of PDA

	C_1	C_2	C_3	C_4	C_5	C_6	C_7
A_1	(0.353,0.253,0.118)	(0,0,0)	(0,0,0)	(0.699,0.599,0.172)	(0.631,0.519,0.295)	(0.707,0.650,0.473)	(0.0,0)
A_2	(0.004,0.035,0.031)	(0.062,0.144,0.159)	(−0.009,0.026,0.030)	(1.415,1.610,1.415)	(0,0,0)	(0,0,0)	(0.206,0.175,0.119)
A_3	(0,0,0)	(0.519,0.432,0.255)	(0.403,0.286,0.117)	(0.415,1.610,1.415)	(0.147,0.188,0.193)	(0,0,0)	(0.206,0.175,0.119)
A_4	(0.004,0.035,0.031)	(0,0,0)	(−0.009,0.026,0.030)	(1.415,1.610,1.415)	(0,0,0)	(0.793,0.707,0.473)	(0.206,0.175,0.119)
A_5	(0.004,0.035,0.031)	(0.447,0.384,0.255)	(−0.009,0.026,0.030)	(1.415,1.610,1.415)	(0.147,0.188,0.193)	(0,0,0)	(0.127,0.135,0.119)

Table 63.8 Values of NDA

	C_1	C_2	C_3	C_4	C_5	C_6	C_7
A_1	(0,0,0)	(0.682,0.673,0.562)	(0.377,0.364,0.208)	(0.699,0.599,0.172)	(0,0,0)	(0,0,0)	(0.746.0.659,0.476)
A_2	(0,0,0)	(0,0,0)	(0,0,0)	(1.415,1.610,1.415)	(0.641,0.651,0.569)	(0.234,0.148,0.−046)	(0,0,0)
A_3	(0.367,00.358,0.209)	(0,0,0)	(0,0,0)	(0,0,0)	(0,0,0)	(0.633,0.605,0.496)	(0,0,0)
A_4	(0,0,0)	(0.346,0.288,0.106)	(0,0,0)	(1.415,1.610,1.415)	(0.285,0.244,0.112)	(0,0,0)	(0,0,0)
A_5	(0,0,0)	(0,0,0)	(0,0,0)	(0,0,0)	(0,0,0)	(0.633,0.605,0.496)	(0,0,0)

Table 63.9 Weighted sum of distances, their normalized values, and the appraisal scores

	sp_i	nsp_i	sn_i	nsn_i	as_i	k (as_i)	Rank
A_1	(0.42,0.36,0.19)	(1.69,1.46,0.78)	(0.35,0.32,0.19)	(−0.05, 0.04,0.43)	(0.82,0.75,0.61)	0.74	5
A_2	(0.30,0.35,0.31)	(1.21,1.42,1.26)	(0.42,0.20,0.14)	(−0.27,0.40,0.57)	(0.47,0.91,0.92)	1.17	**1**
A_3	(0.24, 0.45, 0.36)	(0.98,1.81,1.47)	(0.17,0.16,0.13)	(0.50,0.52,0.63)	(0.74,1.17,1.05)	1.08	3
A_4	(0.46, 0.48, 0.39)	(1.87,1.95,1.57)	(0.37,0.15,0.09)	(−0.11,0.56,0.72)	(0.88,1.26,1.15)	0.84	4
A_5	(0.38, 0.42, 0.36)	(1.53,1.69,1.45)	(0.14,0.13,0.11)	(0.59,0.61,0.68)	(1.06,1.15,1.06)	1.12	2

Acknowledgements The authors of the work would like to express their gratitude toward the anonymous reviewers in giving their valuable feedback that has aided in enhancing the overall quality of the present work. The authors would like to thank TEQIP-III at National Institute of Technology Silchar in providing financial assistance to attend the conference.

References

1. Chakraborty, S., Zavadskas, E.K., Antucheviciene, J.: Applications of WASPAS method as a multi-criteria decision-making tool. Econ. Comput. Econ. Cybern. Stud. Res. 1 **49**(1) (2015)
2. Kosareva, N., Zavadskas, E.K., Krylovas, A., Dadelo, S.: Personnel ranking and selection problem solution by application of KEMIRA method. Int. J. Comput. Commun. Control **11** (1), 51–66 (2016)
3. Chatterjee, P., Chakraborty, S.: Material selection using preferential ranking methods. Mater. Des. **35**, 384–393 (2012)
4. Rastogi, M., Chauhan, A., Vaish, R., Kishan, A.: Selection and performance assessment of Phase Change Materials for heating, ventilation and air-conditioning applications. Energy Convers. Manag. **89**, 260–269 (2015)
5. Zindani, D., Maity, S.R., Bhowmik, S., Chakraborty, S.: A material selection approach using the TODIM (TOmada de Decisao Interativa Multicriterio) method and its analysis. Int. J. Mater. Res. **108**(5), 345–354 (2017)
6. Chatterjee, P., Banerjee, A., Mondal, S., Boral, S., Chakraborty, S.: Development of a hybrid meta-model for material selection using design of experiments and EDAS method. Eng Trans (2018)
7. Khabbaz, R.S., Manshadi, B.D., Abedian, A., Mahmudi, R.: A simplified fuzzy logic approach for materials selection in mechanical engineering design. Mater. Des. **30**(3), 687–697 (2009)
8. Rathod, M.K., Kanzaria, H.V.: A methodological concept for phase change material selection based on multiple criteria decision analysis with and without fuzzy environment. Mater. Des. **32**(6), 3578–3585 (2011)
9. Girubha, R.J., Vinodh, S.: Application of fuzzy VIKOR and environmental impact analysis for material selection of an automotive component. Mater. Des. **37**, 478–486 (2012)
10. Ghorabaee, M.K., Amiri, M., Zavadskas, E.K., Turskis, Z.: Multi-criteria group decision-making using an extended EDAS method with interval type-2 fuzzy sets
11. Stanujkic, D., Zavadskas, E.K., Ghorabaee, M.K., Turskis, Z.: An extension of the EDAS method based on the use of interval grey numbers. Stud. Inf. Control **26**(1), 5–12 (2017)
12. Peng, X., Liu, C.: Algorithms for neutrosophic soft decision making based on EDAS, new similarity measure and level soft set. J. Intell. Fuzzy Syst. **32**(1), 955–968 (2017)
13. Peng, X., Dai, J., Yuan, H.: Interval-valued fuzzy soft decision making methods based on MABAC, similarity measure and EDAS. Fundamenta Informaticae **152**(4), 373–396 (2017)
14. Karaşan, A., Kahraman, C.: Interval-valued neutrosophic extension of EDAS method. In: Advances in Fuzzy Logic and Technology 2017, 13 Sept 2017, pp. 343–357. Springer, Cham
15. Stanujkic, D., Zavadskas, E.K.: A modified weighted sum method based on the decision-maker's preferred levels of performances. Stud. Inf. Control **24**(4), 461–470 (2015)
16. Kahraman, C., Keshavarz Ghorabaee, M., Zavadskas, E.K., Cevik Onar, S., Yazdani, M., Oztaysi, B.: Intuitionistic fuzzy EDAS method: an application to solid waste disposal site selection. J. Environ. Eng. Landscape Manag. **25**(1), 1–2 (2017)

17. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**, 3 (1965)
18. Keshavarz Ghorabaee, M., Zavadskas, E.K., Olfat, L., Turskis, Z.: Multi-criteria inventory classification using a new method of evaluation based on distance from average solution (EDAS). Informatica **26**(3), 435–451 (2015)
19. Ekey, D.C., Winter, W.P.: Introduction to Foundry Technology. McGraw-Hill (1958)

Chapter 64
Tensile and Flexural Property Analyses of Glass/Carbon Embedded Hybrid Natural Fibre Composites

Jobin G. Jose, E. Sivakumar, Jibin T. Philip, Kore Mahesh and Basil Kuriachen

Abstract The composites developed from natural fibre have brought indisputable advantages over synthetic materials in terms of nontoxicity, efficient waste disposability, comparable strength, etc. Some automotive structures such as in aerospace and motor racing vehicles, strength-to-weight ratio is precisely imperative and fibre-reinforced composite materials are used to develop such structures with critical stiffness and weight sensitivities. The environmental effect and the subsequent failure of the fibre-reinforced polymers (FRP) have led to the study of its different mechanical properties. In this work, the fibre-reinforced composite material was fabricated with hybrid jute–coir fibre and polyester resin matrix embedded with glass/carbon fibres. Tensile and flexural strength of the developed natural fibre composites are tested using universal testing machine (UTM) as per the ASTM standards. The sample C2 [jute 75% + coir 25% and carbon 100% of 30% and 15% of total composite composition in weight percentage (wt%)] with yield, ultimate and flexural strength of 118, 122 and 81.04 MPa in-errantly proved to possess an upper edge over jute–coir–glass composites upon testing.

Keywords FRP · Natural Fibre · Composites

64.1 Introduction

Fibre-reinforced composite materials have made a drastic influence on the technology of part designing and manufacturing. Natural fibres are being used as reinforcements due to their good mechanical properties, cost-effectiveness,

J. G. Jose
Saintgits College of Engineering, Kottayam 686532, Kerala, India

E. Sivakumar
Centre for Bio-Polymer Science and Technology, Eloor 683501, Kerala, India

J. T. Philip (✉) · K. Mahesh · B. Kuriachen
National Institute of Technology Mizoram, Aizawl 796012, Mizoram, India
e-mail: philip.jibin07@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_64

biodegradability and lightweight [1, 2]. Moreover, the hybridization of natural fibre increases its mechanical, thermal and water absorption properties [3–5]. Fibre composites have various industrial as well as commercial applications. Nowadays, the metals/metal matrix composites are being replaced by polymers in various fields of their utilization. The orientation and stacking of fibres effect greatly on the properties of any developed composite [6]. Further, in continuous fibre composites, the reinforcements are arranged in the form of layers in different orientations to obtain the required properties with fibre volume up to 60–70% [7].

E-glass fibres have good properties such as high tensile strength high dimensional stability and high chemical resistance and thus it is preferred as the reinforcement material in most of the industries [8]. Carbon fibre is mostly used where lightness and high stiffness are required. Synthetic fibres are not eco-friendly and there comes the importance of using natural fibres to develop environment-friendly composite materials. Generally, these natural fibres do not possess any substantial properties to replace synthetic fibres when used as such. Hence, the natural fibres are hybridized with synthetic fibres to develop good property composites and reduce their impact on the environment. Mostly used natural fibres are jute, sisal, kenaf, hemp, etc. due to their desirable properties and availability in nature. E-glass fibres are often used for hybridization with jute fibre. In some cases, the natural fibre composites have properties better than glass fibre composites [9]. Coir natural fibres are found to have poor mechanical properties. Precisely, coir fibre composites showed only average values for tensile, flexural and impact strength which are poor when compared to glass fibre-reinforced polymer (GFRP) composites [10]. Nevertheless, natural fibres have good mechanical properties with polyester resin matrix [11].

The present work investigates the tensile and flexural strength of hybrid jute–coir fibre composite material compared with glass and carbon fibre at varying compositional ratios (wt%) with polyester resin as matrix. Methyl ethyl ketone peroxide and cobalt naphthenate are the catalyst and accelerator, respectively. Compression moulding technique was used to develop the samples and then subjected to property analyses on universal testing machine (UTM).

64.2 Experimental Methodology

64.2.1 Materials

In this work, the jute, coir, glass and carbon fibres are used as raw materials for the composite specimen preparation. Polyester resin was the matrix used, methyl ethyl ketone peroxide and cobalt naphthenate are the catalysts and accelerators, respectively. Jute fibre was imported from Jaipur, Rajasthan, India; coir fibre has been purchased from Alappuzha, Kerala, India; glass and carbon fibres are obtained from Sharon Enterprises, Ernakulam, Kerala, India. Polyester resin along with its

catalysts and accelerators are also procured from Sharon Enterprises, Ernakulam, Kerala, India.

64.2.2 Composite Preparation

The specimens were prepared by compression moulding process. Stainless steel mould of dimension (200 × 200 × 3) mm was used for moulding the composite sheet specimens. Specimens were prepared according to proportion as described in Table 64.1. Calculation was done to find the required volume for moulding. The matrix and the reinforcement are taken in the ratio 55:45. A calculated amount of polyester resin along with its accelerator and catalyst was mixed thoroughly. Mixing of the matrix is done gently to avoid formation of voids and instantaneously poured onto the reinforcement in order to avoid curing in room temperature. Separator sheets were placed at the inner surface of the mould for quick and easy removal of the moulded sheet. A thin layer of matrix mixture was poured on the separator sheet and the fibres were kept on top. The fibres were arranged in a layer by layer manner in which natural fibres were kept outside and glass/carbon fibres were kept inside. Polyester resin mixture was poured on top of the fibres and spread all around. Separator sheet was placed on it and the mould was closed. Care was taken to avoid formation of air bubbles. Mould was kept in compression moulding machine and pressure was applied. The mould was allowed to cure at room temperature for 48 h. Further, the specimens were taken out of the mould and sides were shaped as shown in Fig. 64.1. The specimen was then cut into standard shapes required for tensile and flexural tests using counter cutting machine.

Table 64.1 Reinforced composite proportions

Sample	Jute/coir (30%)	Glass/carbon (15%)
G1	Jute 100%	Glass 100%
G2	Jute 75% Coir 25%	Glass 100%
G3	Jute 50% Coir 50%	Glass 100%
G4	Jute 25% Coir 75%	Glass 100%
G5	Coir 100%	Glass 100%
C1	Jute 100%	Carbon 100%
C2	Jute 75% Coir 25%	Carbon 100%
C3	Jute 50% Coir 50%	Carbon 100%
C4	Jute 25% Coir 75%	Carbon 100%
C5	Coir 100%	Carbon 100%

Fig. 64.1 Composite sheet specimen

64.2.3 Mechanical Property Testing

The tensile and flexure tests were carried out on the prepared specimens using standard instruments to study their mechanical properties. The data values such as displacement, force stress and strain at different time intervals were obtained from the system connected to the UTM, from which the mechanical property trend graphs of different specimens were plotted. The detailed methodology followed for various test methods are as follows.

Tensile Test. The tensile testing samples were cutted using a counter cutting machine according the ASTM standard D638 [12]. The edges of the samples were finished for tensile testing as shown in Fig. 64.2. The sample was fixed vertically on the UTM by clamping both ends on to the holder and load was applied gradually

Fig. 64.2 Sample for tensile testing as per ASTM D638

until it fractures. The test was repeated for all the samples and its elongation value at the gauge section was recorded.

Flexural Test. Flexural test samples were prepared as per the ASTM standard D790 [13] as shown in Fig. 64.3. Three-point flexure test procedure was employed to carry out the entire sequence of methodology using UTM. The specimen was placed horizontally on the UTM as per standard procedural specifications by providing support at the end points. The force was applied gradually at the midsection of the sample until fracture occurs. The procedure was repeated for all the test specimens developed using different compositions of the fibres. The flexural strength values were recorded for all the test samples.

64.3 Results and Discussion

64.3.1 Tensile Properties

The tensile test results for all the samples are obtained and their yield stress and ultimate stress of glass and carbon fibre composite specimens are tabulated separately in Tables 64.2 and 64.3. It is observed that jute–coir composite samples are having less tensile strength when compared with jute composite samples. This is due to the hybridization of coir fibres in it. These samples show ductile behaviour. Comparing Tables 64.2 and 64.3, it is observed that the carbon composite specimens are having better tensile strength than glass composite specimens. It can also be observed that G3 and C3 samples are having less tensile properties. This is due to lack of proper adhesion between jute–coir–glass/carbon specimens as they are of equal fibre volume.

The force–displacement graph for jute–coir–glass and jute–coir–carbon fibre for tensile testing is as shown in Figs. 64.4 and 64.5, respectively. Comparing both the plots it can be interpreted that the samples G2 and C2 with jute 75% and coir 25% possess better properties than those samples with higher percentages of coir. Moreover, jute–coir–carbon fibre (C2) has greater withstanding capacity of 5990 N before breakdown than the jute–coir–glass fibre (G2) which has a value of 4160 N. So, the general trend is that the increase in percentage of coir with corresponding decrease in percentage of jute causes a reduction in tensile properties of the composite fibre. Hence, optimum value near to the sample compositions of G2 and C2 must be chosen for significant results.

Fig. 64.3 Sample for flexural testing as per ASTM D790

Table 64.2 Tensile strength of jute–coir–glass fibre composites

Sample	Yield stress (MPa)	Ultimate stress (MPa)
G1	90.5	100.5
G2	81	87.5
G3	45	47.1
G4	46.9	49.7
G5	48	50.5

Table 64.3 Tensile strength of jute–coir–carbon fibre composites

Sample	Yield stress (MPa)	Ultimate stress (MPa)
C1	133	140
C2	118	122
C3	21.1	33
C4	69.8	110
C5	71	80

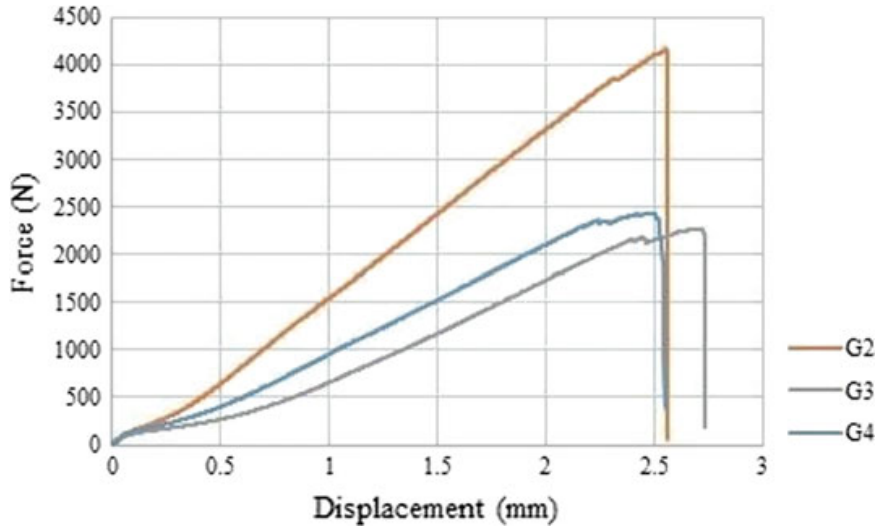


Fig. 64.4 Force–displacement plot for jute–coir–glass fibre for tensile test

64.3.2 Flexural Properties

The three-point flexural test values are obtained and are tabulated in Tables 64.4 and 64.5 separately for glass and carbon fibre composites. It is observed that flexural strength for jute–glass composite is less than jute–carbon composite. On investigation of hybridized jute–coir composite samples, it was observed that G3 is having greater flexural strength as 98.16 MPa and C3 is having the lowest as 54.05 MPa. Comparatively, jute-coir–carbon composite samples have good flexural strength than jute–coir–glass composite samples. The force–displacement graph for jute–coir–glass and jute–coir–carbon fibre for flexural test is as shown in Figs. 64.6

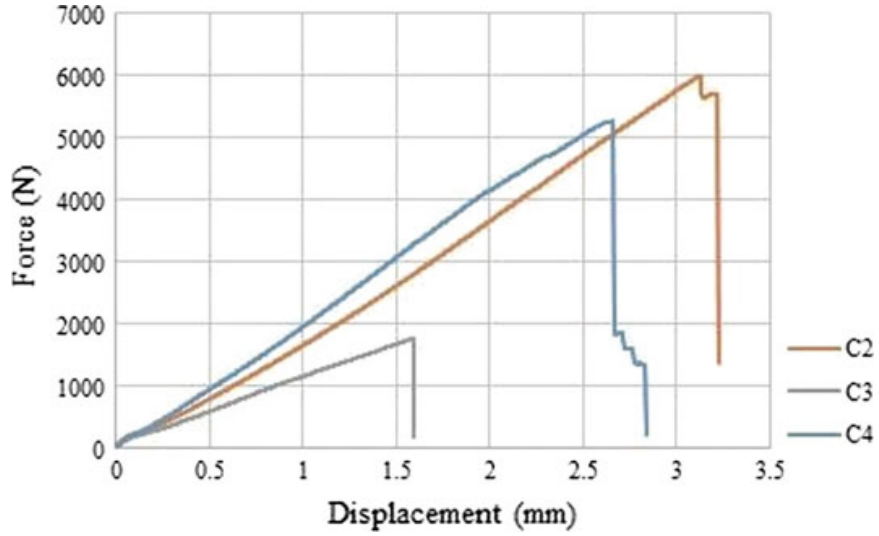


Fig. 64.5 Force–displacement plot for jute–coir–carbon fibre for tensile test

Table 64.4 Flexural strength of jute–coir–glass fibre composites

Sample	Flexural strength (MPa)
G1	100.22
G2	82.88
G3	98.16
G4	86.51
G5	77

Table 64.5 Flexural strength of jute–coir–carbon fibre composites

Sample	Flexural strength (MPa)
C1	125.8
C2	81.04
C3	54.05
C4	90.90
C5	80

and 64.7, respectively. Comparing both the plots, it can be noted that the samples G3, G4, C2 and C4 possess better withstanding capabilities to bending force. When considering the tensile properties discussed in the previous section along with the flexural properties, the samples G3 and G4 can be eliminated due to inconsistent behaviour in properties. Although these samples could withstand flexural loads of 170 and 198.2 N, their tensile load bearing capabilities are insignificant. Interestingly, the samples C2 and C4 sustained its consistency in both tensile and flexural property tests. On considering Fig. 64.6, it can be precisely noted that the samples C2 and C4 were able to withstand the flexural loads up to 145 N and 163.3 N, respectively. The depleting effect of the coir fibre with increase in its

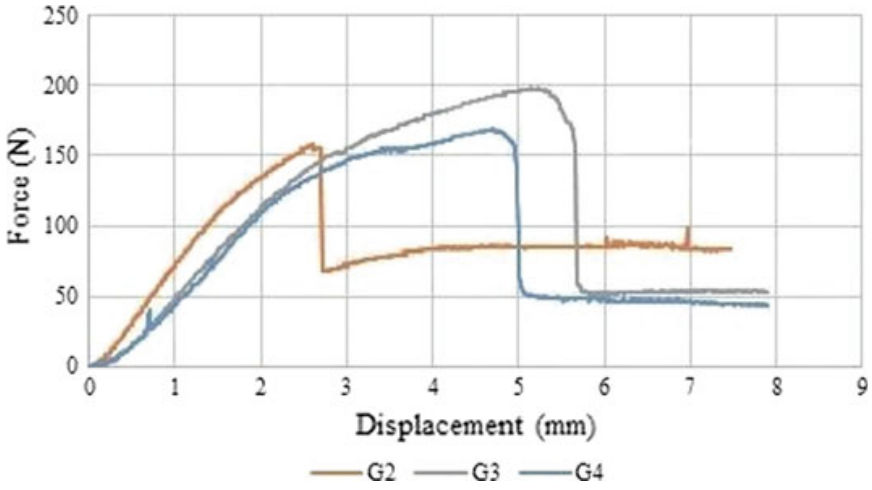


Fig. 64.6 Force–displacement graph for jute–coir–glass fibre for flexural test

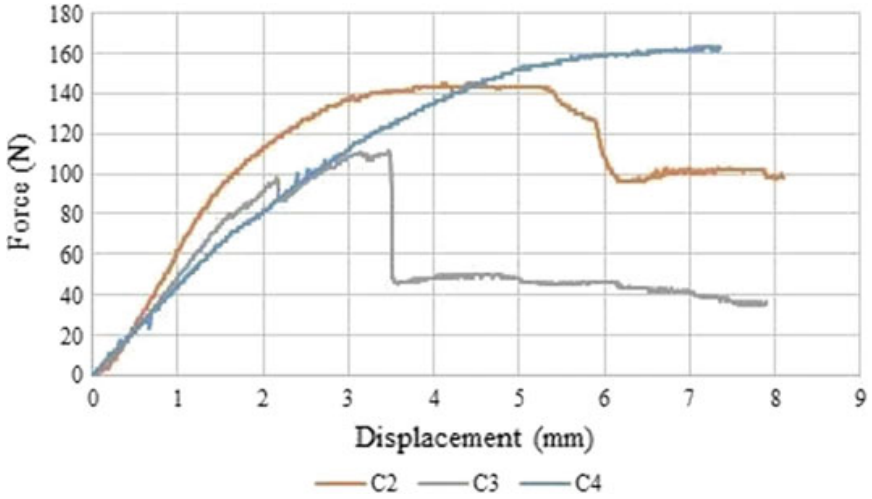


Fig. 64.7 Force–displacement graph for jute–coir–carbon fibre for flexural test

percentage is not much evident in the case of flexural tests. Hence, the jute–coir–carbon fibre is found to have an upper hand compared to the jute–coir–glass fibre.

64.4 Conclusions

The various proportions of jute–coir–glass fibre polyester resin composites and jute–coir–carbon fibre polyester resin composites were prepared and subjected to tensile and flexural tests for the comparative study. The following conclusions were derived from the experiment.

- Jute–carbon composites showed good tensile and flexural properties in comparison with jute–glass composites. Addition of coir into the composites

decreased the mechanical properties of both jute–glass and jute–carbon fibre composites.

- Jute–coir–glass/carbon at equal volume proportion showed the lowest tensile strength due to lack of adhesion between fibres and the resin.
- The samples C2 and C3 showed the maximum and minimum tensile strengths among hybrid jute–coir fibre composites. Flexural strength was highest for G3 (98.16 MPa) and lowest for C3 (54.05 MPa) samples of the jute–coir fibre composites. Further, all the composites investigated showed ductile behaviour.
- The samples C2 and C4 of the jute–coir–carbon fibre composites possess better properties compared to the other samples tested for its mechanical properties. Specifically, the sample C2 was found to have consistent property of all the samples considered. The samples G1, G2 and G3 of the jute–coir–glass fibre composites were eliminated due to inconsistency in properties tested under various conditions.
- Finally, jute–coir–carbon fibre composites had greater tensile and flexural properties over jute–coir–glass fibre composites, comparatively.

Acknowledgements The authors would like to extend their sincere gratitude to Centre for Bio-Polymer Science and Technology (CBPST), Cochin, A unit of CIPET for providing the technical support for this initiative.

References

1. Ku, H., Wang, H., Pattarachaiyakoop, N., Trada, M.: A review on the tensile properties of natural fibre reinforced polymer composites. Compos. B Eng. **42**(4), 856–873 (2011)
2. Thakur, V.K., Thakur, M.K., Gupta, R.K.: Review: raw natural fibre–based polymer composites. Int. J. Polym. Anal. Charact. **19**(3), 256–271 (2014)
3. Lenda, T.A., Mridha, S.: Influence of moisture absorption on impact strength and failure behavior of hybrid jute-carbon/epoxy composite. In: Advanced Materials Research, vol. 264, pp. 457–462. Trans Tech Publications (2011)
4. Sreekala, M.S., George, J., Kumaran, M.G., Thomas, S.: The mechanical performance of hybrid phenol-formaldehyde-based composites reinforced with glass and oil palm fibres. Compos. Sci. Technol. **62**(3), 339–353 (2002)
5. Rajkumar, G., Srinivasan, J., Suvitha, L.: Natural protein fibre hybrid composites: effects of fibre content and fibre orientation on mechanical, thermal conductivity and water absorption properties. J. Ind. Text. **44**(5), 709–724 (2015)
6. Campbell, F.C.: Structural Composite Materials. ASM International (2010)
7. Sezgin, H., Berkalp, O.B.: The effect of hybridization on significant characteristics of jute/glass and jute/carbon-reinforced composites. J. Ind. Text. (2016). Doi: 1528083716644290
8. Bhagat, V.K., Biswas, S., Dehury, J.: Physical, mechanical, and water absorption behavior of coir/glass fibre reinforced epoxy-based hybrid composites. Polym. Compos. **35**(5), 925–930 (2014)
9. Wambua, P., Ivens, J., Verpoest, I.: Natural fibres: can they replace glass in fibre reinforced plastics. Compos. Sci. Technol. **63**(9), 1259–1264 (2003)
10. Harish, S., Michael, D.P., Bensely, A., Lal, D.M., Rajadurai, A.: Mechanical property evaluation of natural fibre coir composite. Mater. Charact. **60**(1), 44–49 (2009)

11. Hasan, K.F., Islam, M., Morshed, M.N., Sultana, M.Z., Mueeid, M.A., Wu, J.H.: Influence of naturally originated jute on the fabrication & mechanical properties of jute/polyester hybrid composite. Am. J. Polym. Sci. Eng. **4**(1), 82–90 (2016)
12. Standard, A.S.T.M. D638: Standard Test Method for Tensile Properties of Plastics. ASTM International, West Conshohocken (2010)
13. Standard, A.S.T.M. D790: Standard Test Methods for Flexural Properties of Unreinforced and Reinforced Plastics and Electrical Insulating Materials. ASTM International, West Conshohocken (PA) (2010)

Chapter 65
Mechanical Characterization of Hybrid Carbon–Glass-Reinforced Polypropylene Composites

Ansuman Padhi and Debiprasada Sahoo

Abstract In the present work, an attempt has been made to investigate the effect of varying wt% of short carbon fiber on the hybrid carbon–glass fiber/polypropylene (PP) composites. Composites of PP reinforced with short glass fibers and carbon fibers were prepared using two-roll milling machine and compression molding techniques with different vol. fraction and wt. percentage. The wt% and vol. fraction of glass fibers were kept constant, by varying the other two materials. Different mechanical properties such as tensile and impact properties were investigated by using suitable experimental setup. In tensile test, with the increase in carbon fiber wt% the strength (MPa) initially decreases but afterward it increases and is giving highest value for maximum wt% of carbon fiber. But for the impact test, the energy absorbed (J) gradually increases by increasing the carbon fiber wt%. From the test results, it shows that the material properties have been improved significantly compared to the polypropylene with only glass fiber, which encourages to use and opt for hybrid composite.

Keywords Hybrid composites · Carbon fiber · Glass fiber · Polypropylene · Mechanical properties

65.1 Introduction

Any single material may not be able to satisfy all the stringent requirements of properties to be used in high-end structural applications. This leads to the development and use of multi-phase integrated composite material systems. Composite materials are formed by macro-scale combination of two or more chemically distinct phases whose performances are superior to their constituent materials and often exhibit qualities that neither constituent possesses. FRP composites consist of

A. Padhi (✉) · D. Sahoo
Indira Gandhi Institute of Technology, Sarang, 759146 Dhenkanal, Odisha, India
e-mail: Anshmech25@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_65

fibers of high strength and modulus embedded or dispersed in matrix (polymer) with distinct interfaces between them.

In the past few years, the thermoplastic polymers have become most preferred materials for various applications due to several attractive properties including less weight, easy to process and cost-effectiveness. Hence, efforts have been made significantly to utilize polymers in various industrial applications, using different types of reinforcements including various kinds of fibers that are integrated into the polymers to increase their physical and mechanical properties. Thus, fiber-reinforced polymer matrix composites are comprehensively used due to their lightweight, high strength, biodegradability ability to be processed rapidly, high stiffness and good corrosion resistance. Among the synthetic fibers, carbon fiber is having high strength and stiffness with significantly light weight compared to same volume of glass fiber. But due to its cost, it can not be used in day-to-day applications. So in this paper, an attempt has been taken to manufacture a hybrid composite in which carbon fiber is added with different weight percentage ratios to the glass fiber polypropylene composite.

However, a number of studies have been carried out on hybrid composites with thermoplastic material (polypropylene) as matrix. Some of the significant findings are discussed below. Fu et al. [1] prepared and investigated the tensile properties of polypropylene (PP) composites reinforced with short carbon fibers (SCF) and short glass fibers (SGF) with extrusion compounding and injection molding techniques. Karsli et al. [2] studied the hybrid reinforcement effect of surface treated or untreated glass fiber (GF)/carbon nanotube (CNT) on the mechanical and electrical properties as well as morphology of polypropylene matrix composites. Gamstedt et al. [3] have studied the strength properties of polypropylene (PP) and polypropylene modified with maleic anhydride (MA-PP) reinforced by continuous longitudinal glass fibers. Ferreira et al. [4] have concerned studies of fatigue on polypropylene/glass fiber-thermoplastic composites produced from a bi-directional woven cloth of co-mingled E-glass fibers and polypropylene. Suresh et al. [5] discussed the influence of the forming pressure and coupler concentration on the mechanical behavior of polypropylene composite laminates reinforced with glass fiber. Ashori et al. [6] worked on improvement in the mechanical and thermo-mechanical properties of short carbon fiber (SCF)/polypropylene (PP) composite, resulting from coating of the SCFs with exfoliated grapheme nanoplatelets (xGnPs). The adhesion enhancement of xGnP–SCF/PP interface was clearly shown by scanning electron microscopy (SEM) images of the surfaces from tensile failure. Rezaei et al. [7] studied the thermal stability of short carbon fiber-reinforced polypropylene (SCF/PP) composite were prepared with melt blending and hot-pressing techniques. Swolfs et al. [8] prepared composite hybrids by taking bonding between carbon fiber prepregs and PP tapes. For a weak bonding, high penetration impacts resistance and the 20% ultimate tensile failure strain of self-reinforced PP were maintained. For a strong bonding, the ultimate tensile failure strain was reduced marginally, but the flexural performance was improved. Unterweger et al. [9] evaluated the mechanical properties of injection-molded CF-reinforced polypropylene (PP) composites under the influence of carbon fibers

(CF) surface properties and the amount of a coupling agent. Khan et al. [10] found that the damping ratio of the hybrid composites is enhanced with the addition of CNTs, which is attributed to sliding at the CNT–matrix interfaces. Arikan et al. [11] present the impact response of E-glass fiber-reinforced composites that have been manufactured with two types of resin, polypropylene and epoxy (thermoplastic and thermoset). Swolfs et al. [8] found that in the prepregs for a homopolymer PP matrix, the weak bonding among fiber and matrix affected the penetration impact resistance according to a linear rule-of-mixtures. Zhang et al. [12] investigated the influences of stacking sequence on the strength of hybrid composites comprising materials with differing stiffness and strength. Sullins et al. [13] worked on the effects of material treatment(s) on the mechanical behaviors of hemp fiber-reinforced polypropylene (PP) composites.

In this work, the composites have been prepared by taking polypropylene as matrix reinforced with short glass and carbon fiber, where the wt% of glass fiber kept constant in order to visualize the effect of CF. The above materials are prepared by using two-roll mill and compression molding techniques. This study mainly aimed to investigate the effect of carbon fiber hybridization on the glass-reinforced PP composites. The tensile properties as well as the energy absorption during impact were primarily considered in the experimentation.

65.2 Experimental Details

Materials used

The materials employed in this investigation were polypropylene, E-glass fiber roving and carbon fiber rovings. The mechanical and physical properties of these materials are listed in Table 65.1. The composition and volume fraction of each specimen are given in Table 65.2. The wt% of glass fiber was kept constant, while the wt% of other two materials is varied.

Specimen Preparation

The glass fiber and carbon fiber are cut into 3–5 mm in length by help of a scissor and make it to chopped fiber form. After that the polypropylene which is already procured is mixed with the above two fibers in the specified ratios with the help of a two-roll mill machine. These machines are extensively used for mixing and

Table 65.1 Mechanical properties of materials

Materials	Tensile strength (MPa)	Young's modulus (GPa)	Density (g/cm^3)
Glass fiber	2450	81.3	2.65
Carbon fiber	3665	230	1.75
Polypropylene	23.9	1.28	0.904

Table 65.2 Volume fraction and weight composition in different specimens

Specimen code	Polypropylene (PP) wt %/(Mass in gm.)	Carbon fiber (CF) wt%/ (Mass in gm.)	Glass fiber (GF) wt%/ (Mass in gm.)
A	0.8/400	0.05/25	0.15/75
B	0.75/375	0.1/50	0.15/75
C	0.7/350	0.15/75	0.15/75

kneading raw rubber, thermoplastics or EVA with chemicals, synthetic rubber into uniform materials. The uniform materials can be fed to calendar, hot presses or other processing machinery for preparing rubber or plastic products. Two-roller mills are the most simple in operation, where the material is crushed between two rollers before it continues on to its final destination. The spacing can be adjusted in between these two rollers by the operator. Smaller pieces are produced when the material is crushed in thinner spacing. The front roller in this study was set to 170 °C, 15 rpm and rear roller to 175 °C, 10 rpm and a gap of 2 mm kept in between the two rollers, after that the mixing process will start (Fig. 65.1).

By setting the rollers quite tightly to each other, refined rubber in thin-sheet form (∼0.10 mm) is produced. A smooth, uniform and free of grain or lumps sheet is produced. The impurities contained in reclaimed rubber are pressed by rolls to the both sides of rubber sheet which can be removed. Consequently, the purity of the reclaimed rubber is increased in refiner mills. Earlier, the finished sheet was pulled toward a wind-up attachment and layered few times to increase the thickness of the sheet (approx. 25 mm), after which they are cut using a hand knife. The cut sheets are then cleaned and stacked (Fig. 65.2).

Compression molding The final product is manufactured using the compression molding machine. Compression molding is a high-pressure and high-volume method suitable for molding intricate and high-strength fiberglass reinforcements. Advanced composite thermoplastics can also be compression molded with

Fig. 65.1 Mixing of fibers with polypropylene in two-roll mill

Fig. 65.2 Sheets obtained after two-roll mill operation

unidirectional tapes, woven fabrics, randomly oriented fiber mat or chopped strand. Compression molding is having ability to mold large and fairly complex parts. Also, it is one of the cost-effective molding methods compared with other methods such as transfer molding and injection molding; moreover, it trashes reasonably little material, giving it an advantage when working with expensive compounds.

The molding material is positioned in the mold cavity and the heated platens are closed by a hydraulic ram. The mold is subjected to pressure and heated until the curing reaction occurs. Temperature history during molding process is shown in Fig. 65.3. Holding time is set for 90 s for maximum temperature of the mold surface at 200 °C. To heat up from 50 to 200 °C, it took only 50 and 130 s for the cooling process from 200 to 50 °C. Altogether, molding was completed in 270 s. In this way, the three different compositions of composite sheets are prepared by compression molding as shown below. For quick and easy removal of composite sheets, polyester film applied at the inner surface of the mold. The dimension of the die is 160 × 160 × 3 mm^3. Care is taken to avoid formation of air bubbles. The die was allowed to cool at room temperature. Then the samples were taken out of the die, cut by hacksaw into different sizes for further experimentation as per standard (Table 65.3; Fig. 65.4).

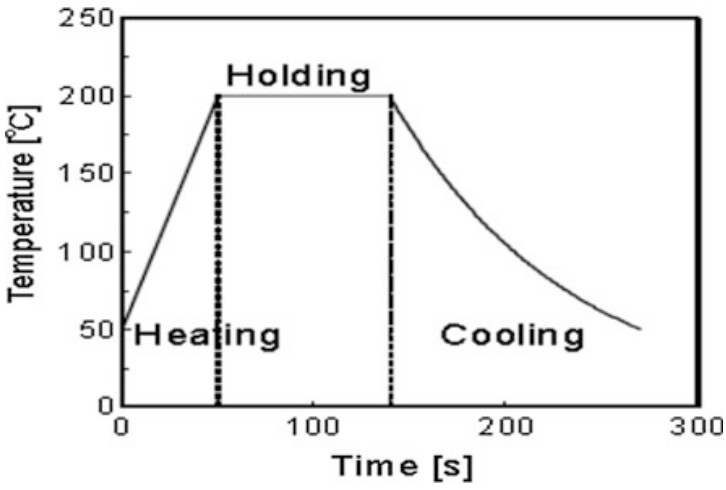


Fig. 65.3 Temperature versus time during a typical compression molding process

Table 65.3 Specimen dimensions for Tensile & Impact test

Test	Length (mm)	Width (mm)	Thickness (mm)
Tensile	165	12.7	3.2
Impact	63.5	12.7	3.2

Fig. 65.4 Compression molding machine

65.3 Results and Discussion

65.3.1 Tensile Test

Tensile tests were conducted on Instron universal testing machine (model 3382) in accordance with ASTM D638. The testing speed was 1 mm/min with at least three samples of each specimen were tested (Figs. 65.5 and 65.6; Table 65.4).

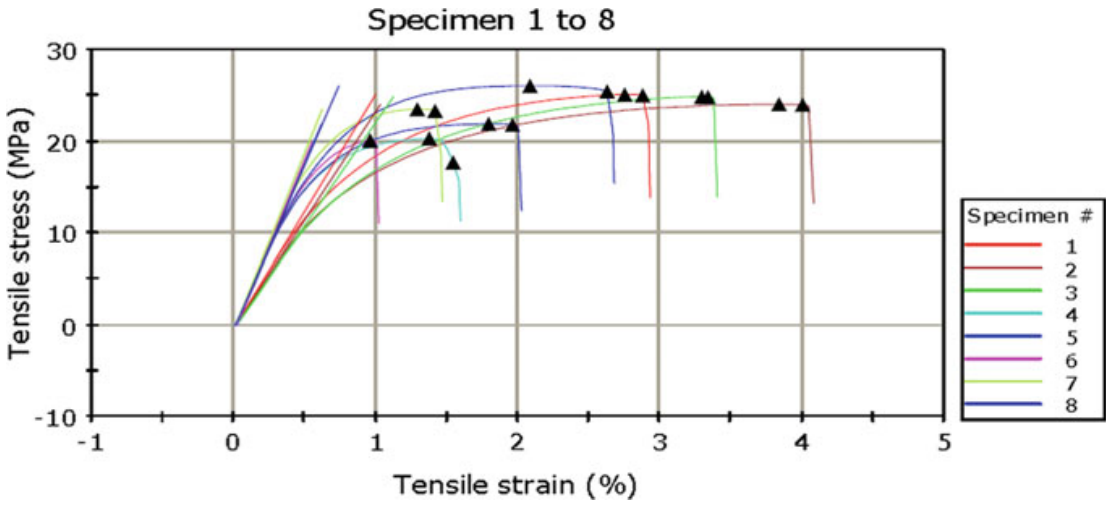


Fig. 65.5 Stress–strain curve of 8 specimens of three different compositions

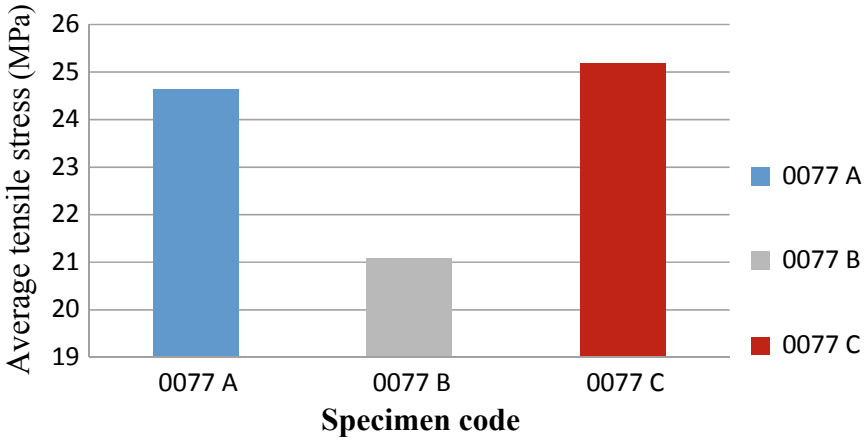


Fig. 65.6 Average ultimate tensile strength of three Specimens

65.3.2 Impact Test

The impact properties of the three specimens were determined using three samples for each composition with the dimensions as mentioned in Table 65.3. The Charpy impact test is a standardized high strain-rate test which measures the amount of energy absorbed by a material during fracture. This absorbed energy is a degree of a given material's notch toughness. The thickness of the specimen and the dimensions of the un-notched length were measured. The pendulum was raised to the left until it indicates the maximum energy range on the upper indicator unit. The specimen was placed horizontally across supports with the notch away from the pendulum. The pendulum was released to measure the energy absorbed during the impact. The energy absorbed during the impact test by different specimens was shown in Fig. 65.7. The test is done in accordance with the ASTM D 256.

It is observed that, the tensile test results shown in Fig. 65.6, specimen 8 have a better strength compared to all other seven specimens. The specimen 8 is having the composition of PP 70%, CF 15%, GF 15%. It has been noted that 5% CF content composition has more tensile strength compare to 10% CF content composition, but when 15% CF is added to PP and GF, it gives the maximum value of tensile strength and this value is 26.04 MPa.

Impact test results are shown in Fig. 65.7. It can be seen that impact strength of code "A" composition is 51.67 J/m. After that when CF% changes from 5 to 10%, the impact energy is increased to 66.67 J/m. But finally when the CF% changes from 10 to 15%, it gives a maximum value of impact energy and that value is 116.33 J/m.

Table 65.4 Tensile strength data sheet of all the 8 specimens

S. No.	Specimen code	Rate 1 (mm/min)	Width (mm)	Thickness (mm)	Maximum load (N)	Tensile stress at maximum load (MPa)	Average tensile stress (MPa)	Tensile strain at break (Standard) (%)	Tensile stress at break (MPa)
1	0077(A)	50.00	15.62	3.2	1254.2	25.09	**24.65**	2.88	24.99
2	0077(A)	50.00	15.54	3.16	1180.18	24.03		4	23.96
3	0077(A)	50.00	15.6	3.2	1239.94	24.84		3.33	24.81
4	0077(B)	50.00	15.2	3.28	1011.48	20.29	**21.08**	1.54	17.68
5	0077(B)	50.00	15.3	3.24	1084.28	21.87		1.96	21.78
6	0077(C)	50.00	15.3	3.2	981.74	24.05	**25.18**	0.96	20.05
7	0077(C)	50.00	15.3	3.2	1149.05	25.47		1.42	23.29
8	0077(C)	50.00	15.3	3.2	1274.99	26.04		2.62	25.43

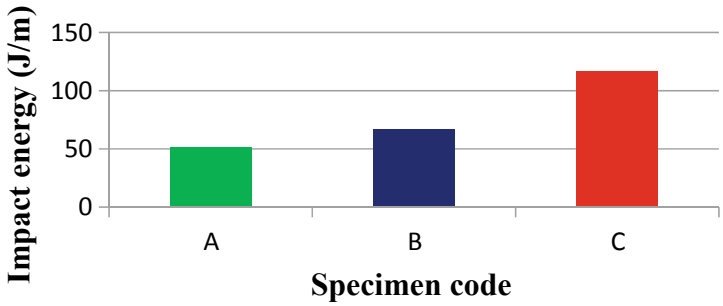


Fig. 65.7 Impact energy (J) of three different compositions

65.4 Conclusion

After conducting the tensile test and impact test of these three compositions, the following conclusions are derived.

- In tensile test the composition of 80% polypropylene, 5% carbon fiber and 15% glass fiber gives the maximum value as compared to 75% polypropylene, 10% carbon fiber and 15% glass fiber. But when the carbon fiber is changed from 10 to 15%, it gives a maximum value.
- In impact test, the value of the impact energy is gradually increased by increasing the carbon fiber %.
- So from the above two test, out of the three combinations the specimen code "C" showed the best result in terms of tensile strength and shock absorption properties.
- Hence, it is the best-suited mixture combination.
- The addition of carbon fiber plays a vital role in case of the glass fiber polypropylene composite. In order to make the specimen cost-friendly, less weight percentage of carbon fiber is taken over glass fiber.
- The test results show that the material properties have been improved significantly compared to the polypropylene with only glass fiber, which encourages to use and opt for hybrid composite.

References

1. Fu, S.Y., Lauke, B., Mäder, E., Yue, C.Y., Hu, X.: Tensile properties of short-glass-fiber-and short-carbon-fiber-reinforced polypropylene composites. Compos. A Appl. Sci. Manuf. **31** (10), 1117–1125 (2000)
2. Karsli, N.G., Yesil, S., Aytac, A.: Effect of hybrid carbon nanotube/short glass fiber reinforcement on the properties of polypropylene composites. Compos. B Eng. **63**, 154–160 (2014)
3. Gamstedt, E.K., Berglund, L.A., Peijs, T.: Fatigue mechanisms in unidirectional glass-fibre-reinforced polypropylene. Compos. Sci. Technol. **59**(5), 759–768 (1999)
4. Ferreira, J.A.M., Costa, J.D.M., Reis, P.N.B., Richardson, M.O.W.: Analysis of fatigue and damage in glass-fibre-reinforced polypropylene composite materials. Compos. Sci. Technol. **59**(10), 1461–1467 (1999)

5. Suresh, S., Kumar, V.S.: Experimental determination of the mechanical behavior of glass fiber reinforced polypropylene composites. Procedia Eng. **97**, 632–641 (2014)
6. Ashori, A., Menbari, S., Bahrami, R.: Mechanical and thermo-mechanical properties of short carbon fiber reinforced polypropylene composites using exfoliated graphene nanoplatelets coating. J. Ind. Eng. Chem. **38**, 37–42 (2016)
7. Rezaei, F., Yunus, R., Ibrahim, N.A.: Effect of fiber length on thermomechanical properties of short carbon fiber reinforced polypropylene composites. Mater. Des. **30**(2), 260–263 (2009)
8. Swolfs, Y., Shi, J., Meerten, Y., Hine, P., Ward, I., Verpoest, I., Gorbatikh, L.: The importance of bonding in intralayer carbon fibre/self-reinforced polypropylene hybrid composites. Compos. A Appl. Sci. Manuf. **76**, 299–308 (2015)
9. Unterweger, C., Duchoslav, J., Stifter, D., Fürst, C.: Characterization of carbon fiber surfaces and their impact on the mechanical properties of short carbon fiber reinforced polypropylene composites. Compos. Sci. Technol. **108**, 41–47 (2015)
10. Khan, S.U., Li, C.Y., Siddiqui, N.A., Kim, J.K.: Vibration damping characteristics of carbon fiber-reinforced composites containing multi-walled carbon nanotubes. Compos. Sci. Technol. **71**(12), 1486–1494 (2011)
11. Arikan, V., Sayman, O.: Comparative study on repeated impact response of E-glass fiber reinforced polypropylene & epoxy matrix composites. Compos. B Eng. **83**, 1–6 (2015)
12. Zhang, J., Chaisombat, K., He, S., Wang, C.H.: Hybrid composite laminates reinforced with glass/carbon woven fabrics for lightweight load bearing structures. Mat. Des. (1980–2015) **36**, 75–80 (2012)
13. Sullins, T., Pillay, S., Komus, A., Ning, H.: Hemp-fiber reinforced polypropylene composites: the effects of material treatments. Compos. B Eng. **114**, 15–22 (2017)

Chapter 66
Effect of Temperature on the Fracture Strength of Perfect and Defective MonoLayered Graphene

Kritesh Kumar Gupta and Sudip Dey

Abstract Graphene is one of the carbon allotropes which have one atom thickness. It is a monolayer of carbon atoms, which are bonded in a two-dimensional hexagonal lattice. It is a unit layer of graphite; stacking the monolayer graphene one over another forms the graphite. The exceptional electromechanical properties of graphene attracted the researchers and scientists to explore the wide application areas so that the properties of graphene could be harnessed. In this regard, the present paper analyses the perfect single-layer graphene sheet (SLGS) and defective graphene sheet by depicting the effect of temperature variations on the fracture strength of both of the sheets. To this end, the molecular dynamics (MD) simulations based on AIREBO interatomic potential field and Nose–Hoover thermostat technique are carried out. It is concluded by the study that with the increasing temperature, the fracture strength of the graphene reduces remarkably. It is also evident that introducing random vacancy of 2.5% in perfect graphene causes the yielding phenomenon at lower strains which remains absent in the perfect graphene. Also, the defective graphene shows the ductile fracture, confirming considerable yielding before complete fracture.

Keywords Fracture strength · Graphene · Molecular dynamics · AIREBO potential · LAMMPS

66.1 Introduction

The extraordinary characteristics of graphene have been used to develop nanoresonators [1], biosensors [2], nanocomposites [3] and many more nanoelectromechanical systems (NEMS). So far research has been conducted in order to determine the exact mechanical properties and strength of the monolayered graphene sheet.

K. K. Gupta (✉) · S. Dey
Department of Mechanical Engineering, National Institute of Technology Silchar, Silchar, India
e-mail: kriteshamie@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_66

Graphene is found to be one of the lightest (0.77 mgm^{-2}) and still stiffest material known. Its fracture strength and Young modulus are recorded as 130 GPa and 1 TPa, respectively. Finding the mechanical properties of graphene through experimental procedures are limited due to the fact that setting up the experimentation for the nanoscale material has its own complications and limitations. Experimentally, one of the common methods to extract graphene from graphite is chemical vapour decomposition (CVD), which is performed at elevated temperatures, it is obvious that during extraction some defects may be introduced to the graphene structure by vacancy, chemical reactions with foreign elements or by a change in the bond structure at elevated temperatures. This leads to getting incorrect desired results from the experimentation. The defect such as vacancies remarkably reduces the strength of graphene, which in return affects the performance of the NEMS.

In the past, researchers have carried out a vast study to accurately characterize the SLGS. To this end, Tsai et al. [4] explored the elastic properties of the graphene along with graphite flakes through molecular dynamics simulation and concluded that the single graphene layer has the higher value of Young's modulus of elasticity and shear modulus in comparison with graphite flakes. The effective characterization of graphene has been a challenge so far as there are a lot of discrepancies recorded regarding its exact characteristics, to address this Thomas and Ajith [5] performed the atomistic simulations using Tersoff interatomic potential and determined the mechanical, thermal and structural properties of the SLGS. The defects in graphene are prone to induced while its extraction, in this contrast Ansari et al. [6] performed the atomistic simulation using Tersoff–Brenner potential energy function to determine the mechanical properties of perfect and defective single-layered graphene, while Dewapriya and Rajapakse [7] investigated the rate of strain and temperature dependence of the failure criteria of single-layer graphene sheet and proposed a fracture mechanics model based on the continuum to characterize the toughness of defected graphene sheet. To explore the effect of Stone–Wales defect, Wang et al. [8] studied the fracture strength of defective graphene sheet using molecular dynamic simulation, the study was carried out at various temperature and it was concluded that the strength of SLGS is affected by the presence of defects, loading directions and temperature, whereas Wang et al. [9] investigated the mechanical responses of graphene membrane with various orientations of 5-8-5 defect using molecular dynamics simulation, and it was concluded from the study that the introduction of 5-8-5 defect in graphene sheet reduces its fracture strength and strain. To address the irregularities based on interatomic potential field, Anastasi et al. [10] presented the MD simulation of single-layer graphene under tensile loading. In order to carry out the simulation, several potential functions such as the bending angle, Morse, Lennard-Jones and torsion potential functions have been used from the mdFOAM library of openFOAM software. It was concluded from the study that graphene is anisotropic, and its fracture strength depends on the sheet size, while Singh et al. [11] predicted the elastic properties of graphene sheet at infinitesimal and finite strains by combining MM3 potential with the Cauchy–Born rule and compared the results with the outcome of a study carried out using

Tersoff–Brenner potential field and second-generation REBO potential field. The MD simulation utilizes different ensembles such as NPT, NVT and NVE for averaging the atomistic properties, and to differentiate the usage of NPT and NVT in the MD simulations, Clavier et al. [12] computed the elastic stiffness tensor at 1 bar and 60 K of FCC argon and compared the result obtained in both the cases. The graphene could be the potential reinforcement material for strengthening the nanocomposites in this context, and Papageorgio et al. [13] reviewed the mechanical properties of graphene and described the strategies to prepare the bulk graphene-based nanocomposites. It was concluded from the study that graphene as reinforcement in composites may improve the strength of the composite up to the next level.

Performing experimentation on graphene is a challenging task due to obvious limitations in experimental set-up for nanoscale analysis. Therefore, using MD simulation as the analysis tool gives the edge over understanding the complex nanomaterials in a better way. In the past, a lot of research is executed in order to determine the exact characteristics of pristine and defected SLGS. However, there is limited study available dealing with concentrations of defect present in the monolayer graphene. In this study, the mechanical properties and fracture strength of the perfect graphene and the graphene sheet with 2.5% vacancy concentration are determined using MD simulation. In order to observe the effect of temperature variation over the fracture strength of graphene, the analysis is carried out at various temperatures, i.e. at 1, 300, 900 and 1200 K. Young's modulus, fracture strength and failure strain are determined for both the directions (armchair and zigzag) of graphene, and results are compared.

66.2 Molecular Dynamics Simulation

66.2.1 Molecular Dynamics

Nanomechanical modelling is one of the best methods to analyse the properties of nanomaterials. Nanomechanics model may be differentiated into three main categories—the first principle approaches such as DFT, semi-empirical methods such as molecular dynamics simulation and the third is the modified continuum model.

The molecular dynamics is computationally affordable in comparison with the first principle approaches. Although it relies on several assumptions associated with the interatomic potential function. In the molecular simulations, materials are viewed as the collection of individual atoms which interact by exerting forces on each other and follow Newton's law of motion. In order to perform molecular dynamics simulations, three important ingredients are required; first, an initial system configuration such as position and velocities of the atoms [$r\ (t = 0)$, $v\ (t = 0)$], interaction potential for the system and a way to integrate Newton's equation of motion. Simulation procedure in MD initiates by determining the forces

acting on an individual atom by the neighbouring atoms. Force acting on an individual atom i may be represented as

$$f_i = -\frac{\partial E_i}{\partial r_i} \tag{66.1}$$

where E_i refers to the potential energy of the atom which is approximated by the interatomic potential field. After the determination of forces, the accelerations are determined using Newton's equation of motion

$$f_i = m_i \frac{\mathrm{d}^2 r_i}{\mathrm{d}t^2} = m_i a_i \tag{66.2}$$

Further atoms are allowed to accelerate with the determined acceleration for a certain duration known as a time step. At the end of each time step, the new positions and velocities are recorded and integrated using numerical integration algorithms such as velocity Verlet algorithm. The use of statistical ensembles such as NPT, NVT and NVE allows controlling the temperature and pressure of the system [14].

66.2.2 Modelling and Simulation

The graphene sheet of 5 nm × 5 nm with 1008 carbon atoms is modelled using visual molecular dynamics (VMD) [15], and the output file received from VMD is fed to Avogadro (an open-source atom manipulator) [16] in order to introduce a 2.5% vacancy in the perfect graphene sheet. Both the sheets (perfect and defected) were uniaxially deformed in x-direction using LAMMPS. LAMMPS [17] is an open-source code package by Sandia National Laboratories, USA. To avoid free edges, all three directions were applied with periodic boundary conditions. The sheets were firstly equilibrated to minimize the energy of the system and relax the graphene structure, over 30 ps prior to applying strain using a time step of 0.5 fs. In the relaxation period, the pressure component of the x- and y-directions was maintained zero using NPT ensemble. The NPT ensemble holds the no. of atoms, pressure and temperature of the system constant, and it uses Nose–Hoover thermostat to control the temperature of the system. Furthermore, to uniaxially deform the sheets a strain was imposed by elongation of the simulation box in both (armchair and zigzag) dimension and the applied strain rate was 10^9 s^{-1}. The stress components in y- and z-directions were kept at zero during simulating the tensile deformation of graphene by keeping in check the pressure component in the said directions using NPT ensemble. Figure 66.1 shows the topology of the perfect graphene sheet and graphene introduced to 2.5% vacancy concentration.

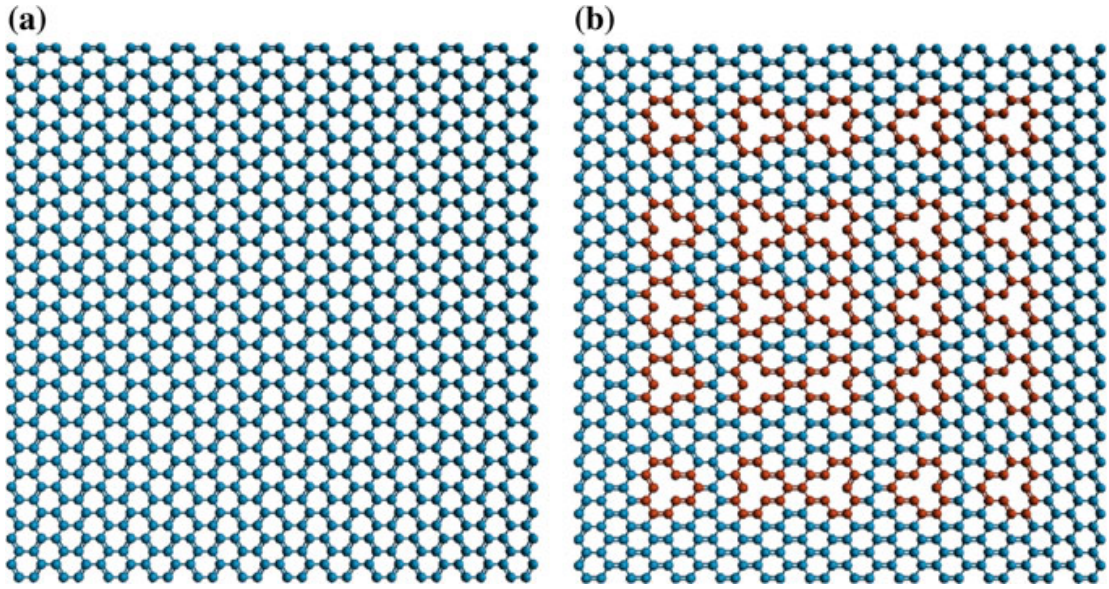


Fig. 66.1 Graphene sheet **a** perfect, **b** with 2.5% vacancy concentration

66.2.3 Potential Field

In molecular dynamics simulation, the potential field is an element which signifies the energy interactions between the atoms. The adaptive intermolecular reactive empirical bond order (AIREBO) potential is employed in the present study. The AIREBO potential is made by combining three different potential fields first, the original REBO potential, second the Lennard-Jones (LJ) and third the torsion potential. The original REBO term imparts the interatomic attractive and repulsive energy, the Lennard-Jones to account for the intermolecular interaction that includes dispersion and short-range repulsion and torsion component to signify the torsional interactions between the carbon–carbon bonds.

The general form to represent the AIREBO potential can be expressed as

$$E = \frac{1}{2}\sum_{i=1}\sum_{i\neq j}\left[E_{ij}^{\text{REBO}} + E_{ij}^{LJ} + \sum_{k\neq i}\sum_{l\neq i,j,k} E_{ijkl}^{\text{TORSION}}\right] \tag{66.3}$$

where E refers to the total potential energy with atoms i, j, k and l. The REBO term which refers the interatomic attraction and repulsion is represented as:

$$E_{ij}^{\text{REBO}} = V_{ij}^{R}\left(r_{ij}\right) + b_{ij}V_{ij}^{A}\left(r_{ij}\right) \tag{66.4}$$

where V^A and V^R refer the attractive and repulsive parts of the potential, respectively; r_{ij} signifies the distance between i and j atom up to the cut-off distance, and b_{ij} refers the bond order which allows for a change in the strength of the bond depending on the environment.

The LJ term of the potential can be described as:

$$E_{ij}^{\mathrm{LJ}} = f\left(r_{ij}\right) V_{ij}^{\mathrm{LJ}} \tag{66.5}$$

$$V_{ij}^{\mathrm{LJ}} = 4\varepsilon_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] \tag{66.6}$$

where $f\left(r_{ij}\right)$ represents the set of cut-off functions σ_{ij}, and ε_{ij} are the distance at which interparticle potential is zero and the potential barrier, respectively.

Finally, the torsion term can be represented as:

$$E_{ijkl}^{\mathrm{TORSION}} = f\left(r_{ij}, r_{jk}, r_{kl}\right) V^{\mathrm{TORSION}}\left(w_{ijkl}\right) \tag{66.7}$$

$$V^{\mathrm{TORSION}}\left(w_{ijkl}\right) = \varepsilon\left[\frac{256}{405}\cos^{10}\left(\frac{w}{2}\right) - \frac{1}{10}\right] \tag{66.8}$$

where r_{ij}, r_{jk} and r_{kl} are the bond lengths and w_{ijkl} is the dihedral angle and ε is the barrier height of the potential [18].

66.2.4 Calculation of Stress

In the molecular dynamics simulation, stress may be interpreted as Cauchy stress or virial stress. When it comes to the computational efficiency, the Cauchy stress is proved to be more considerable than virial stress. However, some amount of initial stress (when strain is zero) is induced by the Cauchy stress at higher temperatures. In this study, the computation of virial stress is carried out through atomistic simulation. Virial stress can be defined as:

$$\sigma_{ij} = \frac{1}{V}\sum_{\alpha}\left[\frac{1}{2}\sum_{\beta=1}^{N}\left(R_i^{\beta} - R_i^{\alpha}\right)F_j^{\alpha\beta} - m^{\alpha}v_i^{\alpha}v_j^{\alpha}\right] \tag{66.9}$$

where α and β refer to the serial number allocated to an atom and the number allocated to the neighbouring atom, respectively, while i and j signify the directional indices, and m^{α} and v_i^{α} are the mass and velocity of the atom α, respectively. R_i^{β} is the location of an atom β in the i direction; $F_j^{\alpha\beta}$ is the force due to an atom β on atom α along the j direction and V refers to the total volume of the material system. The virial stress represents a product of stress and volume, i.e. the stress value computed by LAMMPS would be in the units of pressure*volume, in order to extract the stress from the computed value it is essential to divide the value calculated by the LAMMPS with the total volume. Furthermore, in order to calculate the virial stress, the instantaneous volume is used, i.e. when the imposed strain is

'ε', the value of volume at the given instant would be $V_0(1+\varepsilon)$, where V_0 represents the initial volume of the system [19].

66.2.5 Constant Strain Method

In order to determine the elastic constants of graphene, it is subjected to a constant strain in the desired direction. As in this study, a strain rate of 10^{-9}/s is applied and the simulation box is allowed to deform in the applied direction. At the end of the time step, the resultant atomistic stresses are recorded in each direction. The relation between in-plane stress and stiffness matrix is written in terms of [4]

$$\begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \tau_{12} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & 0 \\ C_{21} & C_{22} & 0 \\ 0 & 0 & C_{66} \end{pmatrix} \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \gamma_{12} \end{pmatrix} \tag{66.10}$$

Due to symmetry in the stiffness matrix, the values of elastic constants are as [4]

$$\begin{aligned} C_{11} &= C_{22}, \\ C_{12} &= C_{21}, \\ C_{66} &= \frac{C_{11} - C_{12}}{2}, \end{aligned} \tag{66.11}$$

As the applied strain is known and the atomistic stresses can be recorded, using these values the elastic constants can be determined. Once the elastic constants are known the value of Young's modulus and Poisson's ratio can be determined as [4]

$$E = \frac{C_{11}C_{22} - C_{12}C_{21}}{C_{22}} \tag{66.12}$$

$$\upsilon_{12} = \frac{C_{12}}{C_{22}} \tag{66.13}$$

66.3 Effect of Temperature Variation

The value of Young's modulus, elastics constants, Poisson's ratio and ultimate strength are determined, which shows a good agreement with the literature values. The said values are determined for the perfect graphene, which is as follows (Table 66.1)

Table 66.1 Validation of mechanical properties of graphene

Properties	Present study	Literature values
C_{11} (in GPa)	887.2	1094 (Thomas and Ajith [5])
C_{12} (in GPa)	153.1	136 (Thomas and Ajith [5])
Young's modulus (in TPa)	0.860	1.020 (Lee et al. [20])
Poisson's ratio	0.172	0.1248 (Thomas and Ajith [5])
fracture strength (in GPa)	145.4	130 (Lee et al. [20])

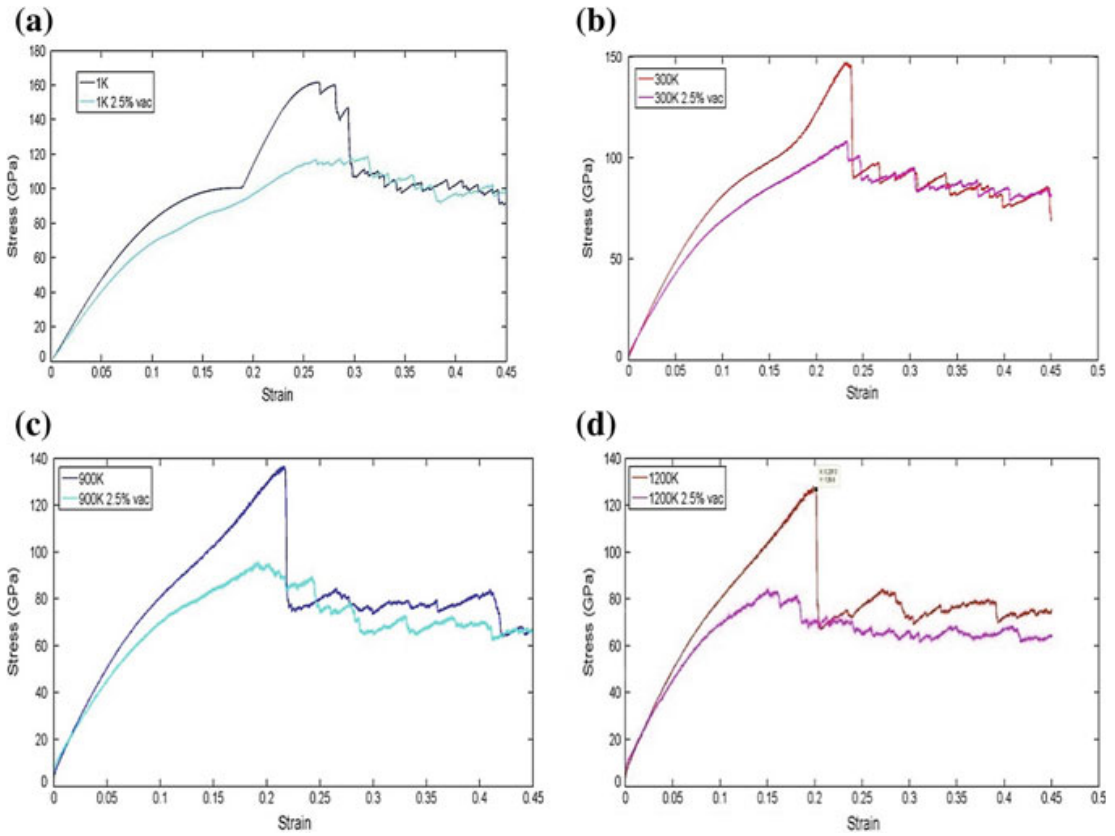


Fig. 66.2 Stress–strain behaviour of perfect and defective graphene loaded in an armchair direction **a** at 1 K, **b** at 300 K, **c** at 900 K, **d** at 1200 K. Defective sheets have 2.5% of randomly distributed vacancies

The stress–strain behaviour of perfect graphene and defected graphene at different temperatures, i.e. 1, 300, 900 and 1200 K is recorded in both armchair and zigzag directions, which is shown in Figs. 66.2 and 66.3, respectively. It is evident by the recorded values that the increment in temperature drastically reduces the fracture strength of the graphene in both the directions and failure occurs at fewer strain values. However, it is observed that the graphene depicts more resistance to failure in the transverse direction (zigzag) as compared to its longitudinal direction (armchair). It can also be seen that on uniaxially loading graphene in both the

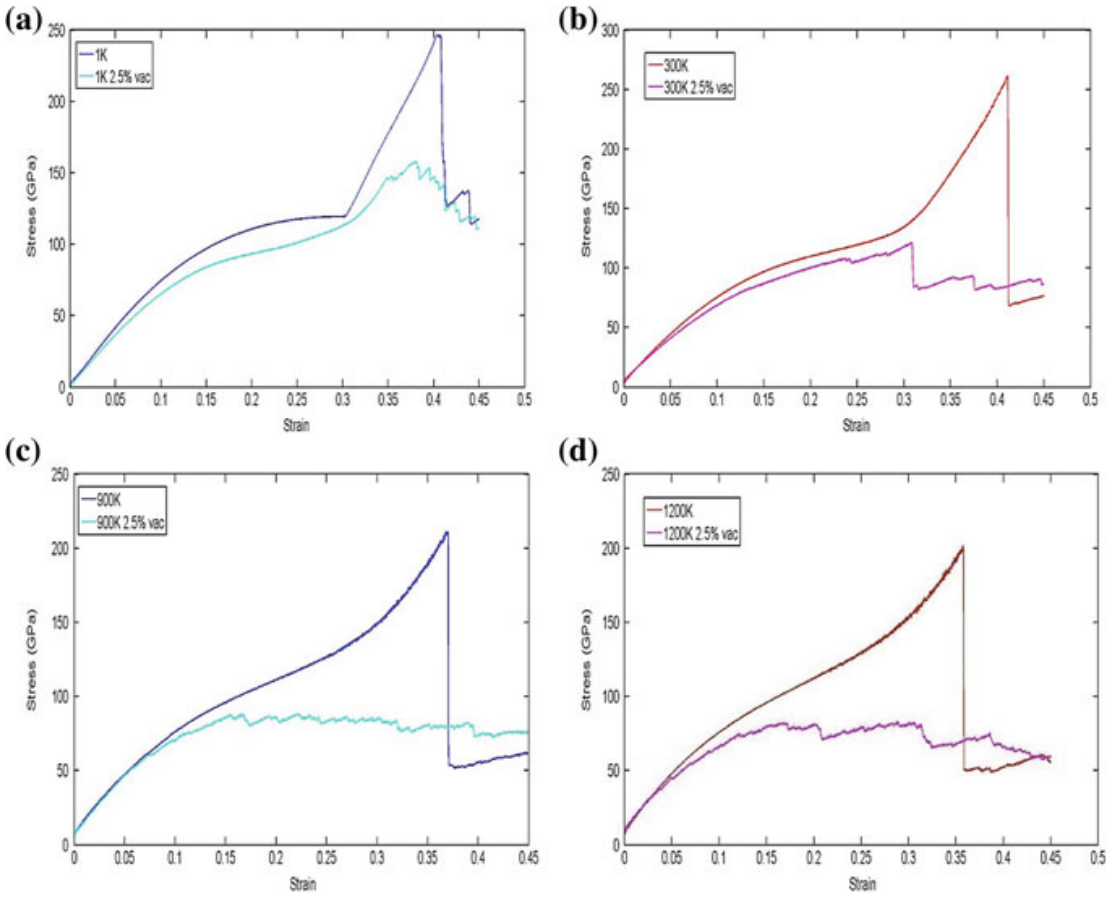


Fig. 66.3 Stress–strain behaviour of perfect and defective graphene loaded in a zigzag direction **a** at 1 K, **b** at 300 K, **c** at 900 K, **d** at 1200 K. Defective sheets have 2.5% of randomly distributed vacancies

directions (armchair and zigzag) at 1 K, it shows the sharp yielding point before fracture.

In both the cases, i.e. loading in an armchair and zigzag direction, the defected graphene does not show the instant fracture, and it is observed that on introducing the 2.5% vacancy to the graphene sheet, it shows the ductile fracture instead of brittle fracture. However, after some strain, it starts showing the yielding.

The given Tables 66.2 and 66.3 represent the recorded values of Young's modulus (E), fracture strength and failure strain of perfect graphene sheet at various temperatures in an armchair and zigzag directions, respectively.

Table 66.2 Young's modulus (E), fracture strength and failure strain of perfect graphene sheet loaded in an armchair direction due to variation in temperature

Temperature (K)	E (in TPa)	Fracture strength (in GPa)	Failure strain
1	0.814	146.9	0.293
300	0.812	145.4	0.236
900	0.809	135.7	0.217
1200	0.799	126.6	0.201

Table 66.3 Young's modulus (E), fracture strength and failure strain of perfect graphene sheet loaded in a zigzag direction due to variation in temperature

Temperature (K)	E (in TPa)	Fracture strength (in GPa)	Failure strain
1	0.745	246.3	0.407
300	0.756	261.9	0.411
900	0.764	211.4	0.370
1200	0.760	200.9	0.357

66.4 Conclusion

In summary, an MD simulation of uniaxial loading the perfect graphene sheet and defective graphene sheet (2.5% vacancy concentration) is carried out at various temperatures (1, 300, 900 and 1200 K) in both armchair and zigzag directions. The AIREBO potential field is used to incorporate the interatomic interactions between the constituent carbon atoms of the graphene sheet and to control the temperature of the system Nose–Hoover thermostat is employed, using NPT ensemble. The fracture strength of perfect and defective graphene sheet is observed at various temperatures, and it is concluded that the temperature has a profound impact over the fracture strength of the graphene sheet. It has been pointed in the study that how sometimes it is obvious that the extracted graphene sheet may have some defects, inherited by the way it is produced. The stress–strain behaviour of the perfect and defective graphene sheet is recorded at various temperatures by loading the graphene in both the directions, and it is concluded that graphene sheet possesses higher fracture strength in the zigzag direction as compared to its armchair direction. Furthermore, it is also been observed that the defective graphene sheet does not depict the brittle fracture, although the defected graphene starts yielding at lower strain than a perfect graphene sheet. The findings of the study provide the insights of variability in the fracture strength of graphene with respect to increasing temperature and with vacancy defects, using which the usability of graphene could be decided for the applications at elevated temperatures and where defects are unavoidable.

Acknowledgements The author(s) would like to acknowledge the financial support provided by the TEQIP-III to attend the conference ICCMM-2019 held at IIT Guwahati.

References

1. Dai, M.D., Kim, C.W., Eom, K.: Nonlinear vibration behaviour of graphene resonators and their applications in sensitive mass detection. Nanoscale Res. Lett. Springer Open J. **7**(1) (2012)
2. Kuila, T., Bose, S., Khanra, P., Mishra, A.K., Kim, N.H., Lee, J.H.: Recent advances in graphene-based biosensors. Biosens. Bioelectron. **26**, 4637–4648 (2011)
3. Wang, G., Wang, B., Wang, X., Park, J., Dou, S., Ahn, H., Kim, K.: Sn/graphene nanocomposite with 3D architecture for enhanced reversible lithium storage in lithium ion batteries. J. Mater. Chem. **19**, 8378–8384 (2009)
4. Tsai, J.L., Tu, J.F.: Characterizing mechanical properties of graphite using molecular dynamics simulation. Mater. Des. **31**, 194–199 (2010)
5. Thomas, S., Ajith, K.M.: Molecular dynamics simulation of the thermomechanical properties of the monolayer graphene sheet. Procedia Mater. Sci. **5**, 489–498 (2014)
6. Ansari, R., Ajori, S., Motevalli, B.: Mechanical properties of defective single-layered graphene sheets via molecular dynamics simulation. Superlattices Microstruct. **51**, 274–289 (2012)
7. Dewapriya, M.A.N., Rajapakse, R.K.N.D.: Molecular dynamics simulations and continuum modelling of temperature and strain rate dependent fracture strength of graphene with vacancy defects. J. Appl. Mech. **81** (2014)
8. Wang, M.C., Yan, C., Ma, L., Hu, N., Chen, M.W.: Effect of defects on the fracture strength of graphene sheets. Comput. Mater. Sci. **54**, 236–239 (2012)
9. Wang, S., Yang, B., Yuan, J., Si, Y., Chen, H.: Large-Scale molecular simulations on the mechanical response and failure behaviour of a defective graphene: cases of 5-8-5 defects. Sci. Rep. (2015)
10. Anastasi, A.A., Ritos, K., Cassar, G., Borg, M.K.: Mechanical properties of pristine and nanoporous graphene. Mol. Simul. (2016)
11. Singh, S., Patel, B.P.: Nonlinear elastic properties of graphene sheet using MM3 potential under finite deformation. Compos. B **136**, 81–91 (2018)
12. Clavier, G., Desbiens, N., Bourasseau, E., Lachet, V., Brussel-Dupend, N., Rousseau, B.: Computation of elastic constants of solids using molecular simulation: comparison of constant volume and constant pressure ensemble methods. Mol. Simul. **43**, 1413–1422 (2017)
13. Papageorgio, D.G., Kinloch, I.A., Young, R.J.: Mechanical properties of graphene and graphene-based nanocomposites. Prog. Mater. Sci. **90**, 75–127 (2017)
14. Cai, W.: Introduction to Molecular Simulations. Stanford University—Winter (2007)
15. Humphrey, W., Dalke, A., Schulten, K.: VMD—visual molecular dynamics. J. Molec. Graph. **14**, 33–38 (1996)
16. Hanwell, M.D., Curtis, D.E., Lonie, D.C., Vandermeersch, T., Zurek, E., Hutchison, G.R.: Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. J. Cheminform. **4**, 17 (2012)
17. Plimpton, S.: Fast parallel algorithms for short-range molecular dynamics. J. Comput. Phys. **117**, 1–19 (1995)
18. Stuart, S.J., Tutein, A.B., Harrison, J.A.: A reactive potential for hydrocarbons with intermolecular interactions. J. Chem. Phys. **112**(14) (2000)

19. Thompson, A.P., Plimpton, S.J., Mattson, W.: General formulation of pressure and stress tensor for arbitrary many-body interaction potentials under periodic boundary conditions. J. Chem. Phys. **131** (2009)
20. Lee, C., Wei, X., Kysar, J.W., Hone, J.: Measurement of the elastic properties and intrinsic strength of monolayer graphene. Science **321**, 385–388 (2008)

Chapter 67
Effect of Process Parameters of Plasma Arc Preheating on Low Carbon Steel

Abhishek Kumar and Pankaj Biswas

Abstract In this study, both experimental and numerical investigations were performed for plasma preheating of AISI 1018 low carbon steel plates. The finite element (FE) software package ABAQUS 6.14 was used to perform the transient thermal analysis using Gaussian distribution as a heat source for plasma preheating. In this FE model, temperature-dependent thermal material properties of low carbon steel (AISI 1018) were used. The influence of different parameters, i.e., traverse speed, plasma arc power on thermal history was determined. Experimentally obtained transient temperature distribution and peak temperature compared fairly well with FE results with a maximum percentage error of around 6.8% and 7.53%, respectively.

Keywords Plasma heating · Finite element analysis · Transient thermal analysis · Heat generation · Process parameters

67.1 Introduction

Plasma heating has been come into the picture since the late 1800s. Around 30 years ago, NASA simulated the very high temperatures of reentry into the atmosphere of earth through plasma heating. The applications of plasma heating are widespread, especially in welding and preheating of the materials. The plasma can achieve 11,100 °C temperatures with very little combustion. The finite element method (FEM) has been the most common technique for the analysis of preheating

A. Kumar (✉)
Department of Mechanical Engineering, Jorhat Engineering College, Jorhat 785007, India
e-mail: abhisheksharma3344@gmail.com

P. Biswas
Department of Mechanical Engineering, IIT Guwahati, Guwahati 781039, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_67

problems which involve thermal, metallurgical and mechanical phenomena simultaneously. The plasma arc preheating system offers better temperature control, quicker reaction time, better control on process, efficient use of energy and lesser capital costs as compared to the other conventional methods.

Nandan et al. [1] observed the influence of a supplementary heat source in front of the friction stir welding (FSW) tool. This reduces the heat input for welding required from the FSW tool and thus the plunging loads of the FSW tool. Wu et al. [2] developed a modified three-dimensional conical heat source model and quasi-steady state for plasma arc heat source to reflect the thermomechanical process of plasma arc welding. Cabage et al. [3] implemented laser as a heat source in front of the FSW tool and reported 50% drops in the friction stir welding tool plunging forces. Kou et al. [4] patented a tungsten inert gas arc preheating system, which reduces FSW tool wear, extends the capability of FSW tool to weld harder materials and also helps weld dissimilar metals. Deshpande et al. [5] developed a conical heat source model to simulate the relationship between welding parameters and welding efficiency. They proposed that the relationship is useful for selecting a combination of weld parameters and keyhole welding.

From the above published literatures, it was observed that only a few number of researchers studied plasma arc preheating process. In this paper, a finite element (FE)-based 3-D heat transfer model was developed to calculate the effects of plasma preheating on temperature history over the AISI 1018 steel plates and validated it with the experimental results.

67.2 Experimental Details

The experiment was conducted for preheating AISI 1018 low carbon steel plate, and the temperature was measured during the preheating of the butt-joint configuration shown in Fig. 67.1. The dimension of the workpiece was 200 mm × 100 mm × 4 mm as shown in Fig. 67.2a. The edges of the workpieces were machined to obtain proper contact of butt joint and clamped to the bed without root opening. The plasma arc torch was clamped at 2 mm standoff distance in the designed fixture, and the steel plates were fixed on the horizontal moveable bed. Transient temperature was measured by using K-type thermocouples over the top surface of the plate at 10, 20, 30 and 40 mm far from the weld line as shown in Fig. 67.2b.

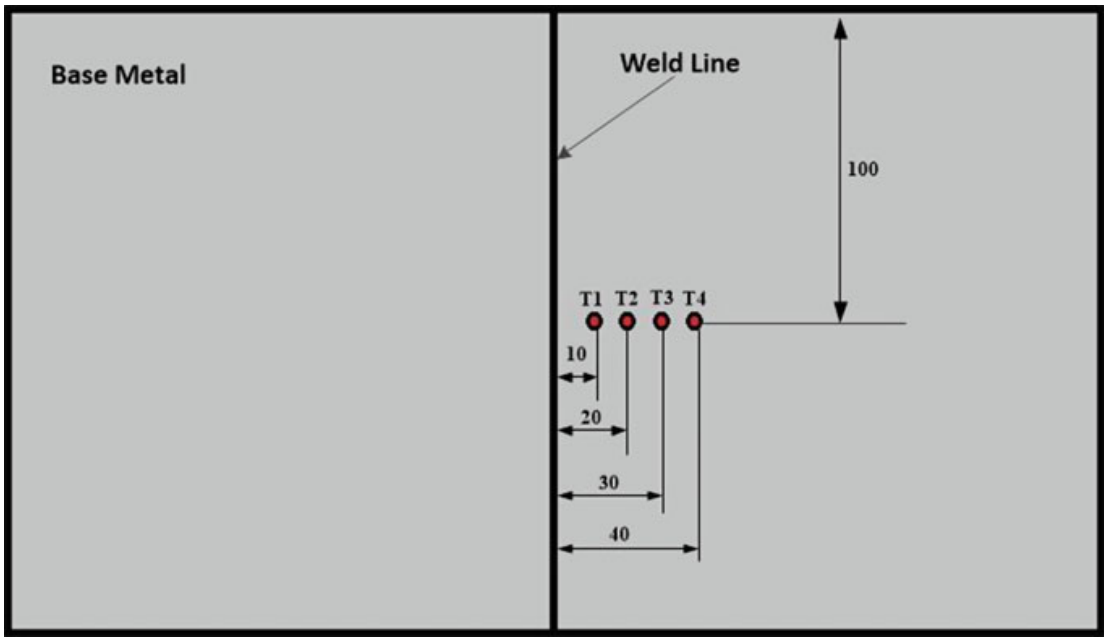


Fig. 67.1 Layout of the steel plate for plasma preheating

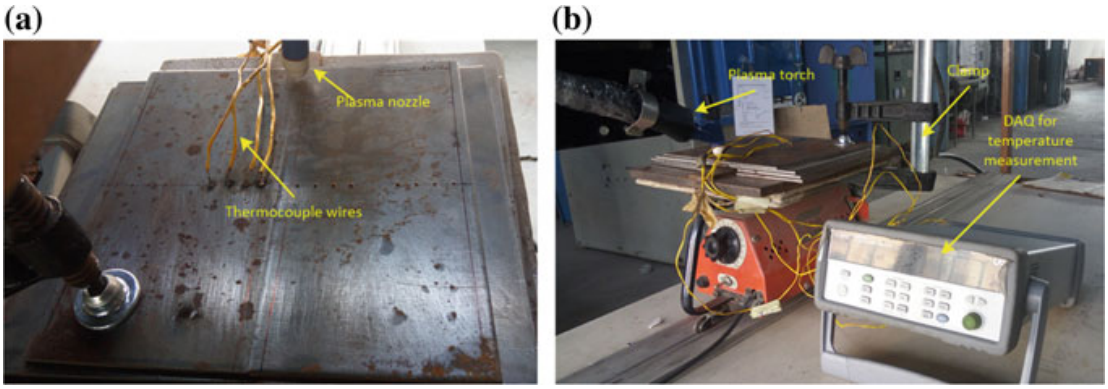


Fig. 67.2 **a** Butt-joint configuration for plasma heating; **b** experimental setup

67.3 Finite Element Analysis and Model Details

A numerical investigation by physical modeling gives a better understanding of physics involved in the process. A 3-D finite element transient thermal model was developed to analyze the heat flow and temperature distribution in preheating. The quality of the numerical model majorly depends on type, size and number of the element. A coarse mesh generally gives the wrong results; whereas, too much fine mesh increases the computational time. Therefore, to get reliable results, the

optimum choice of the mesh should be present. The commercial finite element ABAQUS code was used to develop thermal model. In the thermal model, the actual welding conditions were taken as far as possible. The following assumptions were made to develop the thermal model:

- Heat flux was considered as a heat load.
- The plasma arc heat source was moving with a constant speed over the work piece.
- The material was homogeneous and isotropic.

In this numerical model x-coordinate was considered for plasma arc moving direction. The differential equation for conduction of heat without heat generation in the rectangular coordinate system is expressed as

$$\frac{\partial\left[K(T)\frac{\partial T}{\partial x}\right]}{\partial x}+\frac{\partial\left[K(T)\frac{\partial T}{\partial y}\right]}{\partial y}+\frac{\partial\left[K(T)\frac{\partial T}{\partial z}\right]}{\partial z}=\rho C(T)\frac{\partial T}{\partial t} \tag{1}$$

where ρ, C, K and T stands for density, specific heat, thermal conductivity and temperature of the material, respectively. The initial conditions for the FE model that covers all the elements of the workpieces are expressed as

$$T=T_{\mathrm{a}} \quad \text{for } t=0 \tag{2}$$

where T_{a} is the atmospheric temperature for the present study. A convective heat transfer coefficient 25 W/m^2 °C for natural convection and ambient temperature 25 °C was used as a boundary condition for thermal analysis. At the bottom surface, a high value of heat transfer coefficient was assigned for considering the impact of the backing plate. The mathematical expression for combined convection and radiation heat loss is given as

$$k\frac{\partial T}{\partial n}=h(T-T_{\mathrm{a}})+\epsilon\sigma\left(T^4-T_{\mathrm{a}}^4\right) \tag{3}$$

where n, h, ε and σ are the normal direction vector of the boundary, convection coefficient, emissivity and Stefan-Boltzmann constant, respectively.

67.3.1 Heat Source Model

A three-dimensional thermal model for the preheating process was developed by using the temperature-dependent thermal properties and Newtonian convective heat loss for all surfaces open to the air. The plasma arc Gaussian distribution [6] heat flux is expressed as

$$q(r) = \frac{3Q_{\max}}{\pi R^2} e^{-3\left(\frac{r}{R}\right)^2} \tag{4}$$

where $q(r)$ indicates the heat flux at radius r, R indicates the radius over which maximum arc energy falls, and $Q_{\max}$ is the heat input to the plasma preheating source and given by

$$Q_{\max} = \eta \times V \times I \tag{5}$$

where η, V and I indicate power efficiency, voltage and heating current, respectively. In three-dimensional modeling, DC3D8 or 8-node linear heat transfer brick element was employed. The total number of elements and nodes used in this numerical model was 62,400 and 79,395, respectively. The workpiece was divided into three different mesh zones shown in Fig. 67.3. In the central region up to 15 mm away from the welding line, the element size was kept fine with 1 mm in size while out of this zone, gradually increasing mesh was assigned.

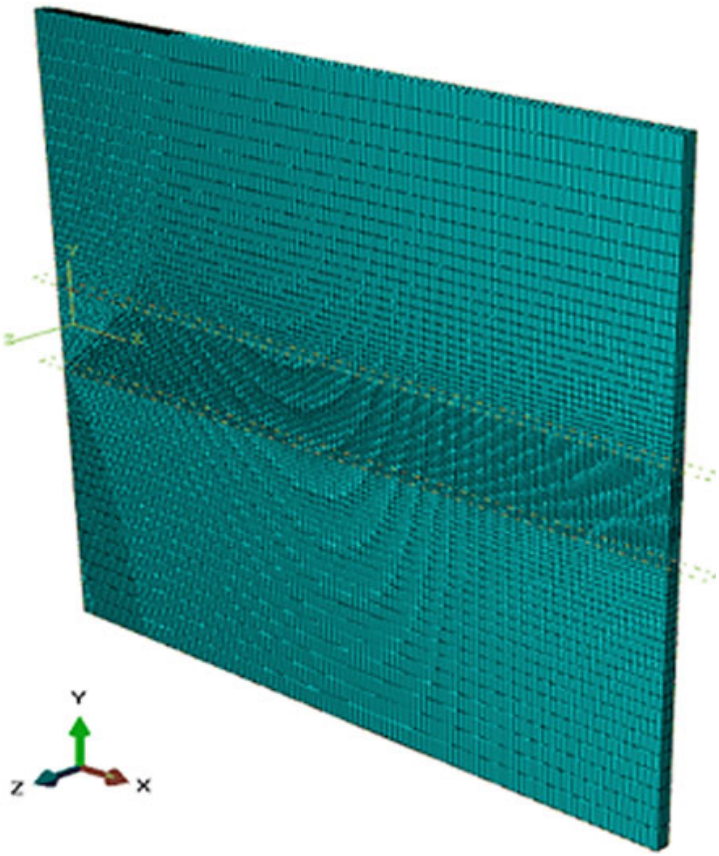


Fig. 67.3 Meshed zone for plasma heating

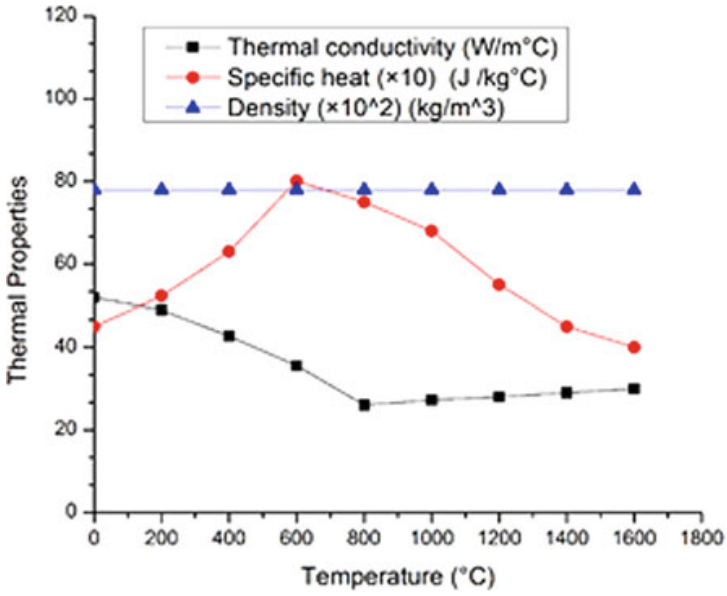


Fig. 67.4 Thermal properties of AISI 1018 low carbon steel

Table 67.1 Chemical composition [8] of AISI 1018 low carbon steel

C	Fe	Mn	P	S
0.14–0.20%	98.81–99.26%	0.60–0.90%	$\leq$ 0.04%	$\leq$ 0.05%

67.4 Material Properties

Thermal properties [7] of AISI 1018 low carbon steel used in the transient thermal analysis are shown in Fig. 67.4. Table 67.1 shows the Chemical composition [8] of AISI 1018 low carbon steel.

67.5 Results and Discussion

In this present FE model, experimental validation and effect of variation in parameters for plasma heating have been addressed. The 3-D numerical model results were developed and compared with the experimental data. The temperature distribution was computed using finite element simulation as shown in Fig. 67.5. The simulation was carried out on ABAQUS using DFLUX subroutine and validated it through experimental investigation. The parameters considered for performing numerical calculation and experimental investigation are presented in Table 67.2.

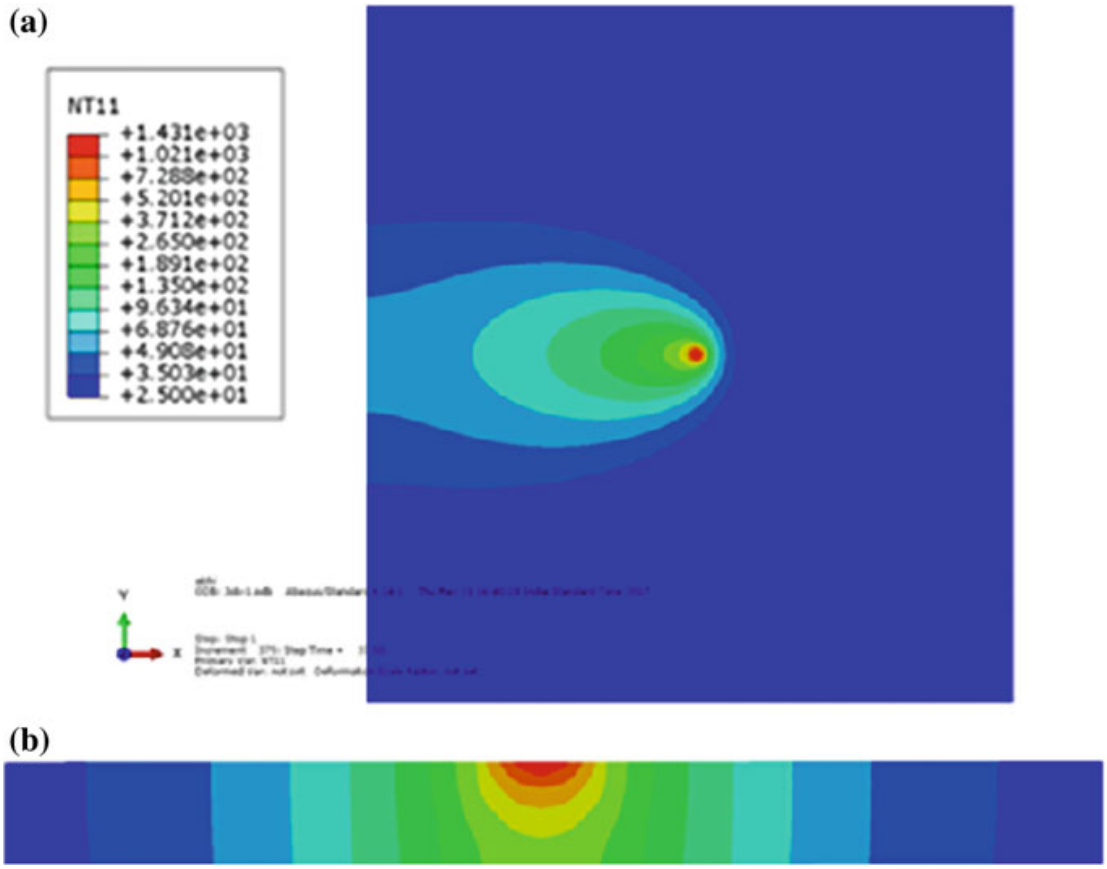


Fig. 67.5 **a** Simulated temperature distribution over the plate; **b** Temperature distribution along the thickness of the plate

Table 67.2 Process parameters for plasma heating

Workpiece length (x dir)	200 mm
Workpiece half-width (y dir)	100 mm
Voltage	13.2 V
Current	20 A
Arc radius	3 mm
Thermal efficiency [9]	0.7
Traverse speed	163.8 mm/min

67.5.1 Transient Temperature Profiles of Plasma

The comparison between the thermal model and data collected from K-type thermocouple for plasma preheating is shown in Fig. 67.6. From the combined numerical and experimental thermal profiles, it is observed that experimental data is much closer to the simulated data. The transient temperature curve plotted at different locations away from the center of weld line over the surface of the plates.

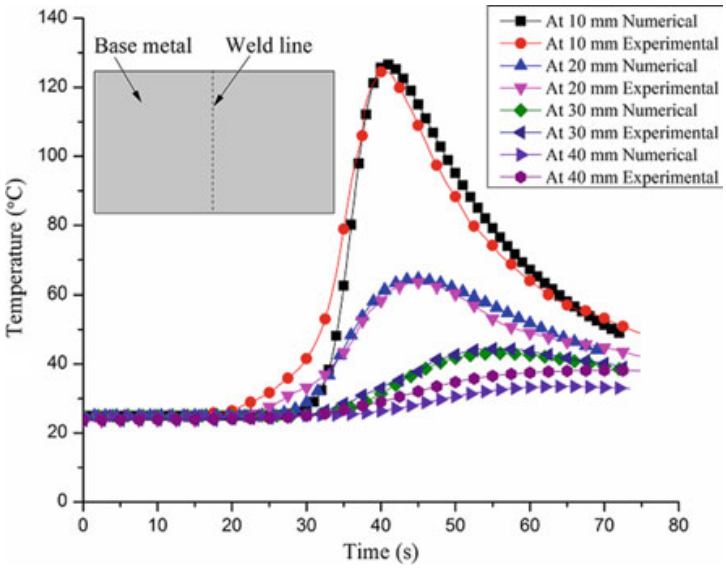


Fig. 67.6 Comparison between thermal model and experimental data at various distances from the weld line

Table 67.3 Percentage error between simulated and experimental data at various distances for preheating

Distance from the center of the plate (mm)	Average percentage error
10	10.82
20	4.77
30	3.39
40	8.35

The following Table 67.3 is representing the average percentage error between modeling results and experimental results for plasma preheating at 10, 20, 30 and 40 mm away from the weld line.

67.5.2 Peak Temperature–Distance Profiles on Top Surface

For comparing the results of temperature with respect to the perpendicular distance away from the center of weld line toward the traverse direction, the temperature is given by K-type thermocouple at four distinct locations 10, 20, 30 and 40 mm away from the weld line. Figure 67.7 shows the comparison between the experimental results and the results obtained from the numerical model for temperature distribution in the traverse direction of the weld line. The present numerical model is found to be in well agreement with the experimental data with an average error of 7.53% for peak temperature.

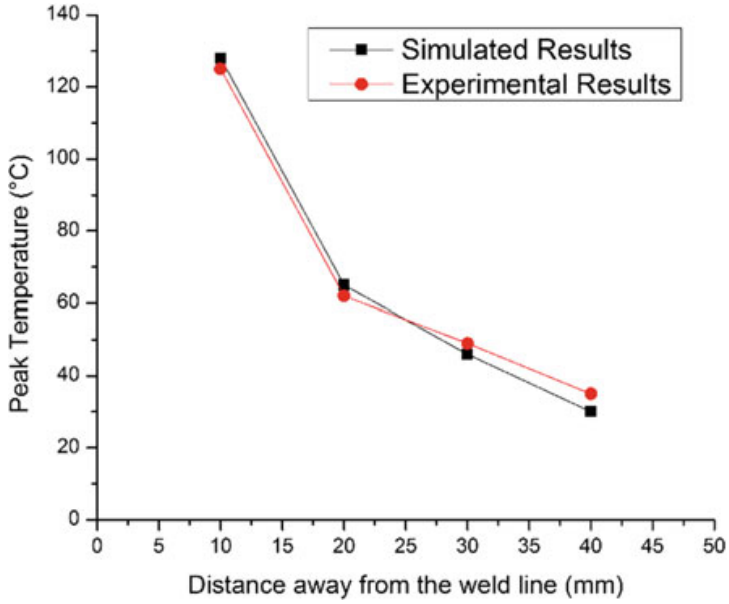


Fig. 67.7 Comparison of peak temperature profile for plasma preheating

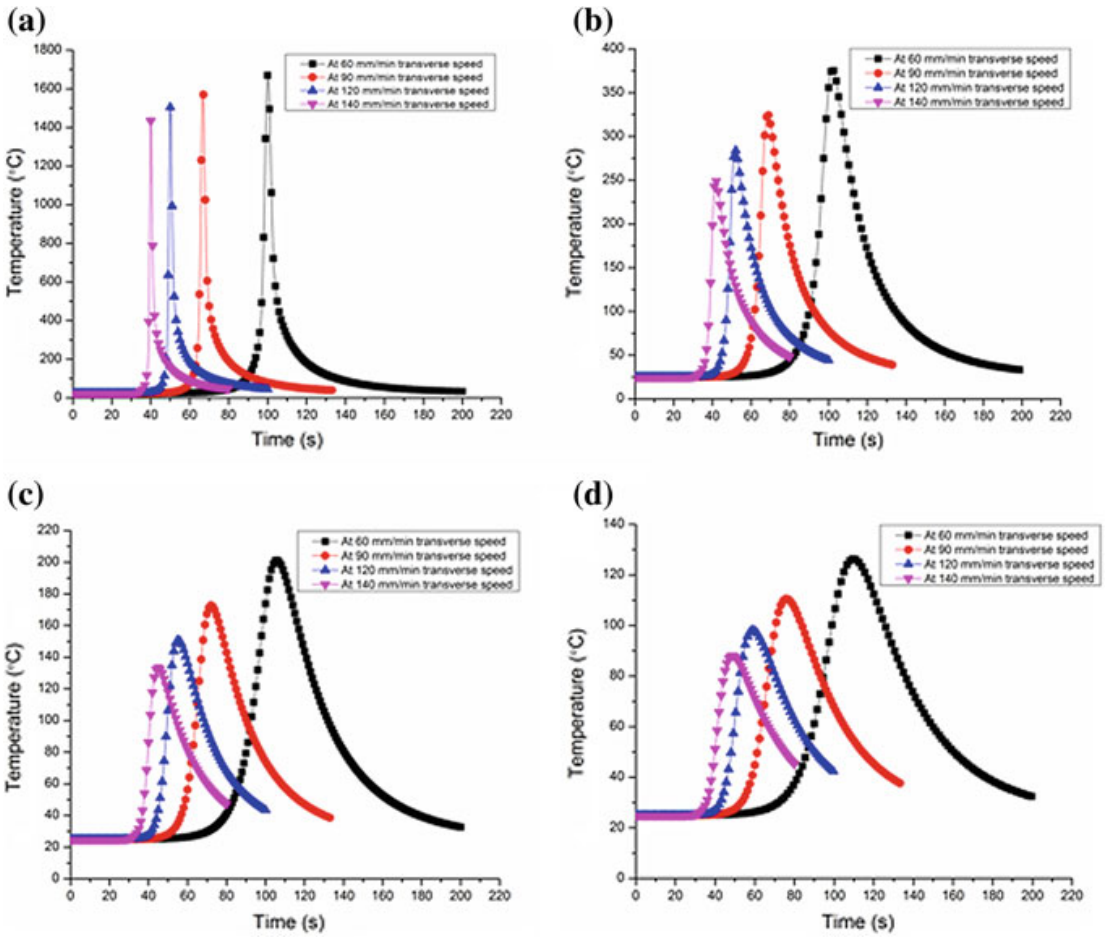


Fig. 67.8 Temperature profile at **a** center; **b** 5 mm; **c** 10 mm; **d** 15 mm away from the weld line at 30 A and 13.5 V

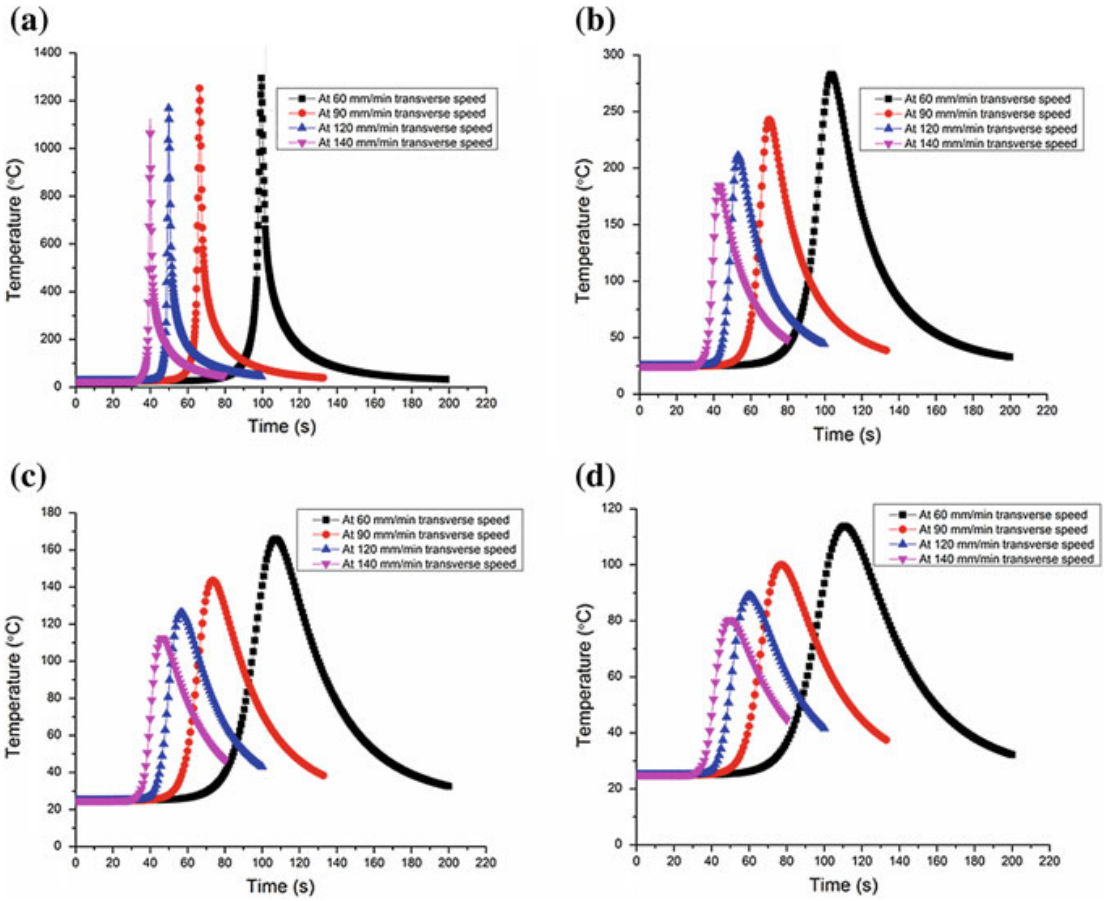


Fig. 67.9 Temperature profile at **a** center; **b** 5 mm; **c** 10 mm; **d** 15 mm away from the weld line at 20 A and 13.2 V

67.5.3 Effect of Variation in Traverse Speed on Transient Temperature Profile at Different Power Supply

The effect of variation in traverse speed on plasma at different power supply is shown in Figs. 67.8 and 67.9. The transient temperature curves were obtained from the numerically developed model at various locations showing that increase in traverse speed leads to a decrease in temperature of the plate.

67.6 Conclusion

In the present study, the plasma was used to preheat two AISI 1018 mild steel plates. The impact of plasma preheating on the thermal profile has been examined experimentally and numerically. The major conclusions are made from the analysis;

- A finite element thermal model was successfully developed for plasma preheating for AISI 1018 low carbon steel plate.
- It was found that finite element modeling of plasma preheating validated with the experimental results with around 6.8% and 7.53% error in transient profile and in peak temperature, respectively, which specifies that the developed thermal model can successfully implement in investigating the thermal histories during the plasma preheating.
- The influence of the power supply of plasma arc on temperature profile has been successfully observed over AISI 1018 low carbon steel plate. It was found that beyond 30 A current, the AISI 1018 steel plate starts melting. Hence, for preheating, the supply current should be less than 30 A.

References

1. Nandan, R., DebRoy, T., Bhadeshia, H.K.D.H.: Recent advances in friction-stir welding–process, weldment structure and properties. Prog. Mater Sci. **53**(6), 980–1023 (2008)
2. Wu, C.S., Wang, H.G., Zhang, Y.M.: A new heat source model for keyhole plasma arc welding in FEM analysis of the temperature profile. Weld. J.-New York **85**(12), 284 (2006)
3. Cabage, B.: New way to weld: ORNL team adds laser tech to expand the reach of a new welding technique. Oak Ridge National Laboratory Reporter (84), pp. 1–4 (2006)
4. Kou, S., Cao, G.: U.S. Patent No. 7,078,647. U.S. Patent and Trademark Office, Washington, DC (2006)
5. Deshpande, A.A., Short, A.B., Sun, W., McCartney, D.G., Xu, L., Hyde, T.H.: Finite element-based analysis of experimentally identified parametric envelopes for stable keyhole plasma arc welding of a titanium alloy. J Strain Anal. Eng. Des. **47**(5), 266–275 (2012)
6. Yaduwanshi, D.K., Bag, S., Pal, S.: Numerical modeling and experimental investigation on plasma-assisted hybrid friction stir welding of dissimilar materials. Mater. Des. **92**, 166–183 (2016)
7. Andrade, R.V.: Pipe Circularity Reformation Via Line Heating. Master's thesis, Massachusetts Institute of Technology (2001)
8. https://www.azom.com/article.aspx?ArticleID=6115
9. Teixeira, P.R.D.F., Araújo, D.B.D., Cunda, L.A.B.D.: Study of the gaussian distribution heat source model applied to numerical thermal simulations of TIG welding processes. Sci. Eng. J. **23**(1), 115–122 (2014)

Part VII
Process Simulation and Optimization

Chapter 68
Reduced-Order Finite Element Analysis of Rough Surface Contact Using Reduced Integration Elements

Ahad Noor, Saipraneeth Gouravaraju and Sachin Singh Gautam

Abstract This paper presents a reduced-order finite element analysis of rough surface contact. Frictionless, elastic contact between a rough deformable block with Gaussian roughness and a rigid flat plate is considered. An equivalent one-dimensional rough surface of a two-dimensional rough surface is generated using MATLAB. Generated rough surface is superimposed on ten layers of generated mesh of the block and analysis is done in commercial finite element software Abaqus. Four-noded quadrilateral elements with reduced-order integration are used to decrease the computational time. The rigid plate is placed close to the largest asperity of the rough surface and is then displaced in small steps. Contact force and real area of the contact are calculated. Results are compared with the results of the full-scale finite element analysis in the literature and are found to be in good agreement.

Keywords Finite element analysis · Rough surface · Gaussian roughness · Elastic contact · Abaqus

68.1 Introduction

Almost all the machines have contact interfaces which work under different types of loading conditions. It is a well-known fact that whatever be the manufacturing process, one cannot eliminate surface roughness. So roughness is important in determining the behaviour of the contacting surfaces. Roughness directly or indirectly affects friction, stress fluctuation, wear, leakage of liquid, and energy transfer

A. Noor · S. Gouravaraju · S. S. Gautam (✉)
Indian Institute of Technology Guwahati, Guwahati 781039, India
e-mail: ssg@iitg.ac.in

A. Noor
e-mail: ahad.noor@iitg.ac.in

S. Gouravaraju
e-mail: saipraneeth@iitg.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_68

between the surfaces. The study of rough surface contact plays an important role in many industries like bearing, lens, household items, electronics, brake and clutch, medical equipment, and medical implants manufacturing industry. For better tribological design and performance, analysis of the contact between rough surfaces is very important. The high points on a rough surface are known as asperities/peak/summit/hill, and the low points are known as valleys. When two rough surfaces are brought into contact, the actual contact occurs at the peaks/asperities. Because of surface roughness, contact occurs only at these discrete contact locations. Then, the real area of contact can be calculated by summing all the discrete areas of contact at each asperity. Moreover, in many cases the actual area of contact is limited to only 20–25% of the nominal contact area [1]. As the contact area of these asperities is very small compared to the total or nominal contact area, very high stresses develop in the asperities causing the material to yield. In some applications where two rough surfaces are brought into contact repeatedly, the increase in interfacial stress leads to damage of the materials resulting in wear. So, roughness and complexity of the contact surface have emerged as an important factor in analysis of strength, fatigue, wear, damage, friction, and energy dissipation. Hence, analysing the contact behaviour of the rough surfaces is an interesting domain for researchers.

Hertz [2] has performed the first contact analysis to investigate the optical properties of lenses in contact. Hertz considered two elastic smooth spheres in contact. He considered the contact zone to be a circle and solved considering minimum deformation energy. He concluded that the contact area is proportional to the $(2/3)$rd power of the applied load. In analytical as well as numerical analysis, to create rough surface a distribution of height of asperities must be defined. Many researchers have considered Gaussian distribution. Moreover, Gaussian distribution of asperities is experimentally verified by Celedón et al. [3]. They measured surface roughness of a gold film and concluded that the height of asperities follows Gaussian distribution. Greenwood and Williamson [1] in their pioneering work on nominally flat rough surface contact developed one of the first elastic contact models. This model represents a rough surface by a number of hemispheres with the same curvature radius R, but the height of asperities is varied according to the Gaussian distribution. Each asperity was assumed to deform elastically according to the Hertz contact theory. The flat plate is considered to be at some distance from mean of rough surface asperities and using the probability theory calculated number of asperities in contact at that distance. Then, using the Hertz contact theory, the authors solved for the real area of contact and the contact pressure. They concluded that actual area of contact varies linearly with the applied load. Lu et al. [4] studied the initial contact between a rough surface and a flat plate using Greenwood-Williamson model (GW model) for low loads over a small area of contact. They considered an exponential distribution of rough surface. They proposed a modification for the initial position of the flat plate. They set it so that the flat plate is just in contact with the highest asperity, while the GW model considers the initial position of the flat plate somewhere in the middle. However, they considered initial contact so they limited their analysis to low loads only. Yastrebov et al. [5, 6] have done the analysis of self-affine fractal rough surface in elastic,

frictionless, non-adhesive contact. They have created Gaussian rough surfaces numerically. They have used Fourier forward technique. They have applied pressure loading in small increments and have observed the evolution of real contact area under increasing pressure from zero up to full contact between rough surfaces. Later, they have done the analysis of contact between elasto-plastic rough surface and a rigid plane with two approaches: full-scale finite element analysis and to save computation time a reduced model with displacement boundary condition. In their reduced model, they considered single asperity, and to incorporate interaction between asperities, they have proposed a displacement boundary condition for the asperity outside the contact zone.

From the literature survey, it is clear that many researchers have extensively studied the contact between rough surface. Experimental calculation of actual area of contact is very difficult. So, researches mainly focused on analytical and numerical analyses. Due to the assumptions made in analytical studies, their accuracy is limited. Although the full-scale three-dimensional numerical simulations as in [5] give good results, the computational times are very high. So, in order to reduce the computational time without compromising on the accuracy in estimating the contact behaviour, in the present work, the method proposed by Scaraggi et al. [7] is used to generate an equivalent one-dimensional rough surface from a two-dimensional rough surface. A Gaussian, one-dimensional equivalent rough surface is created in MATLAB. This one-dimensional equivalent rough surface has all the information of two-dimensional rough surface. Finite element analysis is performed in commercial finite element software Abaqus. The rough surface is superimposed on the top surface of a two-dimensional block. To save the computational time, elements with reduced integration are considered. It is observed that the results obtained in the finite element analysis are consistent with the results of full-scale finite element analysis at low loads.

The remainder of the paper is structured as follows: Sect. 68.2 describes the representation and generation of rough surface. Section 68.3 presents the generation of the finite element mesh of the deformable block. Section 68.4 details the boundary conditions used and the methodology of the numerical simulations using Abaqus. Section 68.5 presents the obtained numerical results. Conclusions drawn from the study are presented in Sect. 68.6.

68.2 Description of the Rough surface Geometry

Any random rough surface can be represented by its power spectral density (PSD) denoted by $\Phi(k_x, k_y)$ of a profile in any direction [8] where k_x and k_y are the wavenumbers in x- and y-directions, respectively. In the present work, spectroscopic wavenumber which is the inverse of wavelength is used. It is normalized with the domain length L, i.e., $k = L/\lambda$. PSD contains all the statistical information

of the surface topography and is unbiased by the particular scan size and pixel resolution used. Mathematically, PSD is the Fourier transform of the signal's autocorrelation function and is a function of wave vector. In the present work, isotropic self-affine fractal surfaces are used whose power spectral density is defined as [5]

$$\Phi\left(k_x, k_y\right) \sim \left(\sqrt{k_x^2 + k_y^2}\right)^{2(H+1)} \tag{68.1}$$

where H is Hurst exponent.

In the present work, the following form of PSD is considered [5]:

$$\Phi_{2D}(|\mathbf{k}|) = \begin{cases} C_0 & \text{if } k_l < |\mathbf{k}| < k_r, \\ C_0 \frac{|\mathbf{k}|}{k_r} & \text{if } k_r \leq |\mathbf{k}| \leq k_s, \\ 0 & \text{otherwise} \end{cases} \tag{68.2}$$

where C_o is the constant that determines the roughness amplitude and $k = |\mathbf{k}|$ is the magnitude of the wave vector. In all the simulations, there is no plateau, so $k_l = k_r$.

Then, an equivalent one-dimensional rough surface profile is generated using the corrected equivalence condition between the mean square slopes of one-dimensional (1D) and two-dimensional (2D) rough surfaces. This is given by [7]

$$\Phi_{1D}^{eqv}(k) = \left(k_{1D}/k_{2D}\right)^2 \pi \Phi_{2D}(k) \tag{68.3}$$

where $\left(k_{1D}/k_{2D}\right)^2 \approx 1.56$ is the correction factor. This equivalence increases the root mean square height of one-dimensional rough surface

$$\sqrt{(h^2)_{1D}} = \left(k_{1D}/k_{2D}\right)\sqrt{(h^2)_{2D}} \tag{68.4}$$

For given PSD and surface area A, the surface heights can be calculated as

$$h(\mathbf{x}) = \sum_{|\mathbf{k}} \tilde{h}(|\mathbf{k}) \mathrm{e}^{i(|\mathbf{k}|.\mathbf{x} + \phi(|\mathbf{k}))} \tag{68.5}$$

where $\left|\tilde{h}(|\mathbf{k})\right| = \frac{2\pi}{\sqrt{A}}\sqrt{\Phi(|\mathbf{k})}$.

Then, the two-dimensional and the equivalent one-dimensional rough surfaces are generated using the inbuilt fast Fourier transforms (FFTs) and inverse Fourier transforms (IFFTs) functions in MATLAB. These are shown in Figs. 68.1 and 68.2, respectively.

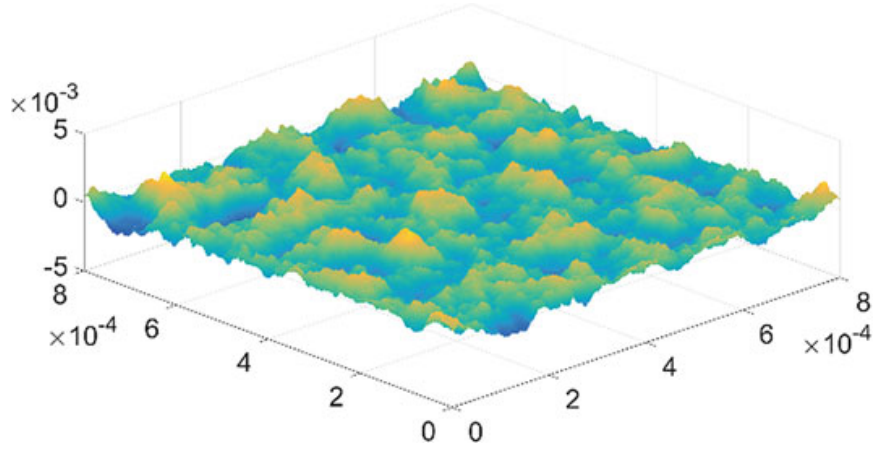


Fig. 68.1 Generated two-dimensional rough surface

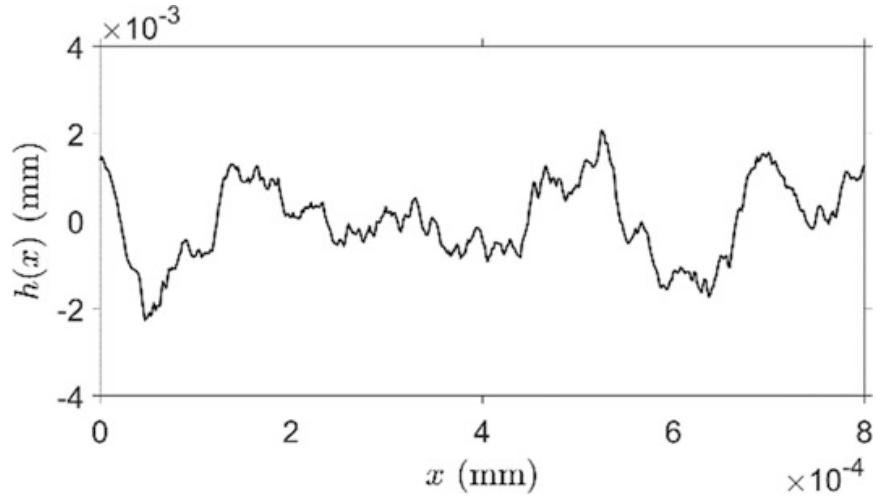


Fig. 68.2 Generated equivalent one-dimensional rough surface

68.3 Mesh Generation

To perform numerical analysis first, a mesh of the block having the dimensions 0.8 × 0.15 mm is considered. The finite element mesh is generated in Abaqus. It consists of 511 and 200 elements in the x- and y-directions, respectively. Four-noded plain stress quadrilateral element CPS4R with reduced integration is used for meshing. Since Abaqus does not have any option to create rough surface, it is generated in MATLAB. The height of the highest asperity is 0.0021 mm (2.1 μm). The generated rough surface heights $h(x)$ (see Fig. 68.2) are supposed to be superimposed onto the top layer of the initial mesh to create rough surface. But, this may produce badly distorted elements. In the worst case, it can produce negative Jacobians. Jacobian is the ratio of the area/volume of the deformed element to the undeformed element in a finite element mesh. So, if the Jacobian is negative, the

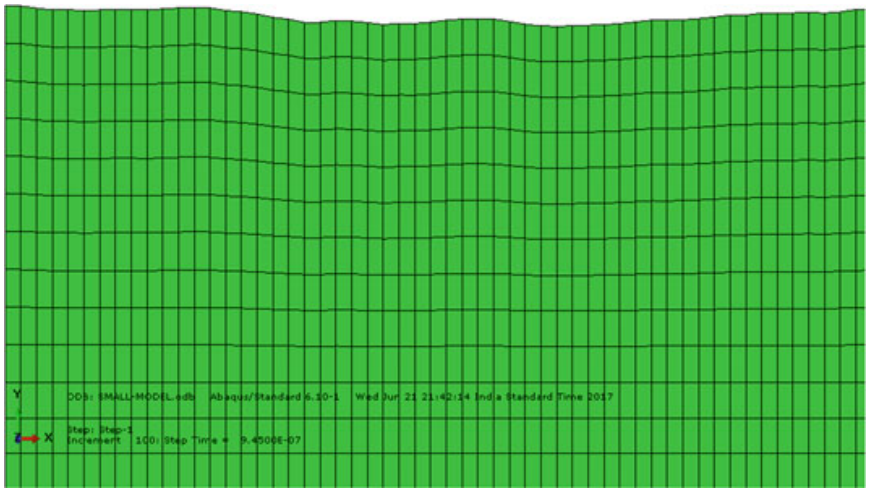


Fig. 68.3 Magnified view of generated roughness superimposed on ten layers of the block to avoid negative Jacobians

Table 68.1 Mechanical properties of steel

E (GPa)	v	ρ (Kg/m^3)
200	0.3	7800

elements have undergone a non-physical movement. To avoid this, all the nodes in top 10 layers are moved by a fraction of the heights of the top surface and change in height is gradually brought to zero in tenth layer (see Fig. 68.3). This is done in following steps: Coordinates of calculated height in MATLAB is added on the nodes of top layer. Then, nodes of second layer are moved by adding 0.9 times height of rough surface; similarly, third layer is moved by 0.8 times, and this is done up to tenth layer nodes. The rigid surface is modelled as a straight horizontal line with its initial position at the 0.0001 mm from the highest point on the top surface. The resulting finite element mesh (zoomed view near the top surface) is shown in Fig. 68.3. The material for the rough surface is considered to be steel. Material properties of steel are given in Table 68.1. In the present work, only elastic contact is considered.

68.4 Boundary Conditions

In this simulation, fixed boundary conditions are applied on the bottom surface of the block in both the x- and y-directions. The vertical surfaces of the block are fixed in the x-direction. The top surface is in contact with a rigid flat plate which is placed at a distance of 0.0022 mm from the smooth surface (without roughness) and

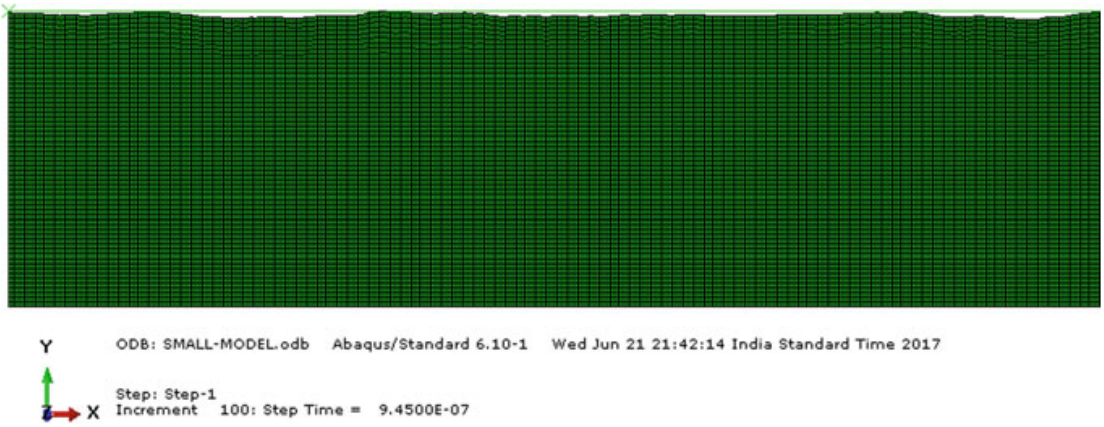


Fig. 68.4 Assembly of rigid plate and rough surface

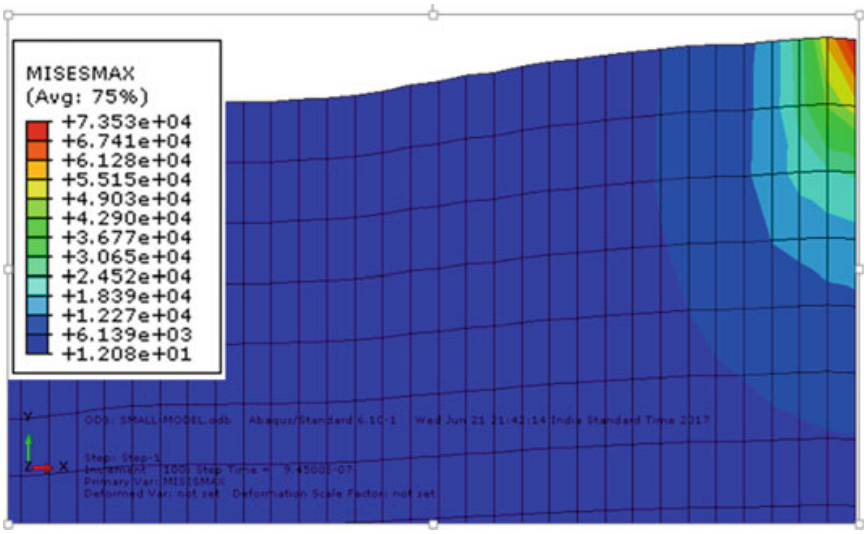


Fig. 68.5 Deformed surface with superimposed von Mises stress

0.0001 mm from the highest asperity as shown in Fig. 68.4. The rigid plate is given a total displacement of 0.002 mm in hundred time steps. However, it is worth mentioning that the Abaqus can reduce the time step to account for change in contact status. It is observed that initially the time step is nearly 10^{-8} s but as more and more asperities start to come in contact, Abaqus made the time step in the range of 9×10^{-9} s.

In Abaqus static analysis, the movement of upper rigid plate is guided by displacement boundary condition rather than velocity. Figure 68.5 shows the deformed surface at the end of simulation with superimposed von Mises stress. The problem is simulated on a Dell T620 workstation (32 GB RAM and 16 cores). The whole simulation is completed in approximately 4 min only.

68.5 Results and Discussion

The actual or real area of contact A_r and the corresponding contact force F are calculated at each load step. Total contact area is sum total of contact area of each asperity. Analytically, contact area is calculated by calculating the number of nodes in contact. Then, ratio of nodes in contact and total number of nodes on the top layer is calculated. This ratio is multiplied with nominal area to give actual area of contact. The same algorithm is used in numerical methods to calculate actual area of contact. Actual area is normalized with nominal area A_o to give dimensionless area of contact $A^* = A_r/A_o$. Contact force is normalized with the product of elastic modulus E and nominal area A_o to give dimensionless pressure $P^* = F/EA_o$. Figure 68.6 shows the evolution of normalized area of contact A^* with the normalized pressure P^*. The results obtained by Yastrebov et al. [6] for the contact of a two-dimensional rough surface with a rigid block are also included. It can be seen that the results from the present work match well with the numerical results of Yastrebov et al. [6] for low deformation regime while there is some discrepancy at higher deformation. It is important to note that the results of Yastrebov et al. [6] are obtained from a full-scale finite element simulation of two-dimensional rough surface which took 220 h, whereas in the current work the results are obtained from the analysis of an equivalent one-dimensional surface and at a much reduced computational time of 4 min. Although plot between actual area of contact and contact pressure is not linear in the present analysis, variation from linear behaviour as predicted by the full-scale FE model of Yastrebov et al. [6] is not very high for low loads, i.e., $P^* < 0.1$. However for loads $P^* > 0.1$, the current analysis predicts a nonlinear behaviour.

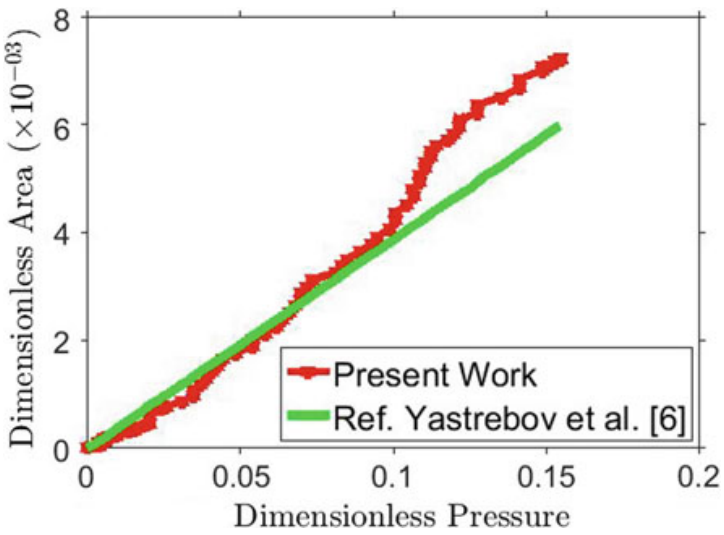


Fig. 68.6 Variation of dimensionless contact and dimensionless contact pressure

68.6 Conclusion

Numerical analysis of the contact between a linear elastic block and a self-affine rough surface is carried out using the finite element analysis software Abaqus. To save computational time, one-dimensional equivalent rough surface with reduced integration elements is used for analysis. Actual area of contact and contact force are calculated. It is observed that the variation of real area of contact with the applied contact pressure is close to linear in low load regime, and at higher loads, the variation is observed to be nonlinear. This behaviour can be explained in many ways. First, to decrease computational time, finite element approximation elements with reduced integration are considered which contain single Gauss point. So accuracy of result is compromised at the cost of calculation time. Full-scale analysis took 220 h [6] while this analysis of one-dimensional equivalent rough surface is completed in four minutes only. Second, equivalent one-dimensional rough surface used in this analysis can capture the essential features of its two-dimensional counterpart, and it is not exact. Although equivalent rough surface with reduced integration results showed some discrepancy at higher loads, computational time is reduced from 220 h to 4 min. Third, fixed boundary conditions are used at the vertical sides of the rough surface. This boundary condition is easy to apply. This can be applied by selecting "Encastre" option in Abaqus, but it can be used for approximation only because in real condition elements at the edges will not be fixed and show some deformation. For example, as rigid plate approaches the block, it will crush the rough surface and elements at edges will try to move outside. For exact results, more complex boundary conditions are needed which will be considered in the future work.

Acknowledgements The authors are grateful to the SERB, DST for supporting this research under project SR/FTP/ETA-0008/2014.

References

1. Greenwood, J.A., Williamson. J.B.P.: Contact of nominally flat rough surfaces. Proc. Royal Soc. London A Math. Phys. Eng. Sci. **295**(1442), 300–319 (1966)
2. Johnson, K.L.: Contact Mechanics. Cambridge University Press, Cambridge, UK (1985)
3. Celedon, C., Flores, M., Haberle, P., Valdes, J.: Surface roughness of thin gold films and its effects on the proton energy loss straggling. Braz. J. Phys. **36**(3B), 956–959 (2006)
4. Lu, P., O'Shea, S.: Mechanical contact between rough surfaces at low load. J Phys D Appl. Phys. **45**(47), 475303-1–475303-15 (2012)
5. Yastrebov, V.A., Anciaux, G., Molinari. J.-F.: From infinitesimal to full contact between rough surface: evolution of the contact area. Int. J. Solids Struct. **52**, 83–102 (2015)
6. Yastrebov, V.A., Durand, J., Proudhon, J., Cailletaud, G.: Rough surface contact analysis by means of finite element method and new reduced model. C.R. Mec. **339**(7–8), 473–490 (2011)
7. Scaraggi, M., Putignano, C., Carbone, G.: Elastic contact of rough surfaces: a simple criterion to make 2D isotropic roughness to equivalent 1D one. Wear **297**(1–2), 811–817 (2013)
8. Nayak, P.R.: Random process model of rough surfaces. ASME J. Lubr. Technol. **93**(3), 398–407 (1971)

Chapter 69
FEM-Based Modeling of Drilling-Induced Delamination in Laminated FRP Composites

Vishwas Divse, Deepak Marla and Suhas S. Joshi

Abstract In fiber-reinforced plastics (FRP) laminate, different laminae are adhesively bonded to each other. The interfaces between plies are weak because these are resin rich; therefore, there is a low-resistance path for crack propagation. As a result of this, delamination is a frequent mode of failure in the FRP laminate. The drilling of laminated FRP is always associated with delamination damage; therefore, delamination-free drilling is a major concern for aerospace industries. Experimental investigations of drilling-induced delamination have been widely studied in the literature, whereas very few studies are available related to its modeling. In this paper, the effect of the chisel-edge length of drill-bit on push-out delamination was studied. Finite element modeling (FEM) approach was used to model interfaces of laminae, and it was further extended to model delamination in drilling of FRP laminate. Cohesive zone method (CZM) based on traction–separation law was used in FEM to model the interfaces of laminae. It was observed that load-carrying capacity of beam in double cantilever beam (DCB) test reduced by 50% for '1 mm' initial crack length. Also, push-out delamination increased by around 10% with an increase in the chisel-edge length of drill bit from '0 mm' to '0.35 mm.' The FE models developed in this work were able to simulating delamination phenomenon with enough accuracy.

Keywords Delamination · Fiber-reinforced plastics · Drilling · Finite element modeling · Cohesive zone method

V. Divse · D. Marla · S. S. Joshi (✉)
Department of Mechanical Engineering, Indian Institute of Technology Bombay, Mumbai 400076, India
e-mail: ssjoshi@iitb.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_69

69.1 Introduction

Composite materials are increasingly used as a structural material in aerospace and other engineering applications. Among different composites, FRP composites have wide applications in aerospace industries due to their high specific strength and stiffness. While assembling the FRP components, holes are required for fasteners like bolts or rivets. Mechanical drilling is widely used to create holes in these components [1]. However, this operation is always associated with damages such as fiber pullout or breakage, matrix cracking or thermal degradation, delamination due to its poor machinability [2]. Consequently, this could result in strength and stiffness degradation and may lead to failure of a component in service. Additionally, they also cause poor assembly tolerance and reduction in structural integrity. Delamination is a failure of an interface between plies. This is a highly undesirable problem because it affects the mechanical properties of FRP composites [3]. A number of methodologies have been adopted for the purpose of minimization of delamination. The quantification of delamination is done by the delamination factor (F_d). Several researchers developed empirical models to predict delamination factor in drilling [1, 4, 5]. They identified that drilling speed and feed have the highest effect on delamination. These models are applicable only for particular drilling conditions and materials. Artificial neural networks (ANNs) can predict delamination factor at different values of spindle speed, feed rate, drill diameter, drill point geometry, etc. [6–8]. It requires huge experimental data to train an ANN, and therefore, it is a tedious process to apply for different materials. Experimental techniques used to characterize delamination factor are challenging and costly. Hence, modeling of delamination is seen as an effective tool to predict delamination in drilling. FEM is a potential technique to model the delamination phenomenon in FRP composites [9]. FE models developed by some researchers can predict drilling-induced delamination [10–12]. These models are very complex in nature and have not considered the effect of drill-bit geometry. Most of the models presented in the literature to predict delamination are empirical and focused only on drilling parameters and neglected the geometry of the drill bit.

In this paper, a simple 2D FE model of drilling is presented to predict drilling-induced delamination. The main objective is to study the effect of chisel edge of drill bit on the push-out delamination. In Sect. 69.2, the modeling techniques of FRP laminate and their interfaces are discussed. Further, in Sect. 69.3, FE models of the double cantilever beam (DCB) and drilling of FRP laminate are presented. The results and conclusions based on simulations performed in this work are summarized in the last section.

69.2 Modeling Approach

FRP laminates are composed of a number of plies with different orientations. Each FRP ply is assumed to be homogeneous with orthotropic elastic properties. These plies are glued together with a resin-rich interface. A two-phase (FRP ply and interface) model (Fig. 69.1) is required to simulate the delamination process. The associated material and damage models of FRP ply and interface adhesive are described in the following subsections.

69.2.1 Fiber-Reinforced Plastics (FRP)

While modeling of FRP laminate, an individual ply in a laminate is assumed to be homogeneous, orthotropic, and in a state of plane stress [13]. Hashin's failure criterion [14] has been used to model the damage initiation in FRP material. The criterion considers the failure of both fibers and matrix under both tension and compression. Failure mechanisms of fibers are governed by the longitudinal stress with reference to the fiber orientation. Meanwhile, failure mechanisms of the matrix are governed by the transverse and tangential stresses to the fiber. Failure is said to occur if the failure criterion of that failure mode is equal to or greater than one.

69.2.2 Cohesive Zone Method (CZM)

The interface between the plies is resin rich; hence, it is weaker compared to the plies. Therefore, the delamination occurs only at the interface. Special elements such as cohesive elements are used in the FE model to simulate delamination. The

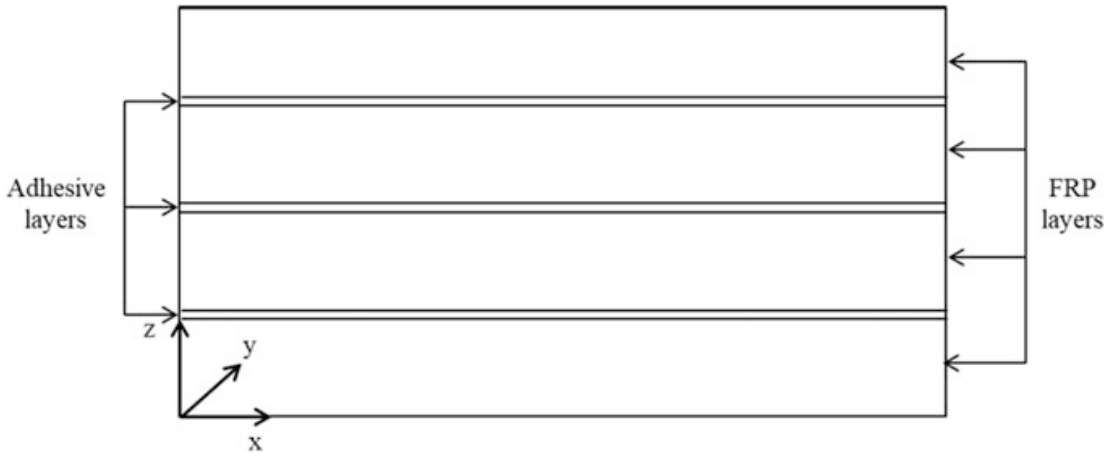


Fig. 69.1 Schematic representation of an FRP laminate

cohesive zone method (CZM) is based on traction–separation law [15]. The law has the capability to simulate the behavior of an interface from an initial linear elasticity to the final failure. Quadratic nominal stress criterion is used as a failure criterion [16]. The criterion assumes damage is initiated when a quadratic interaction function involving all nominal stress component ratios reaches a value of one. Damage evolution law based on fracture energy is incorporated in the model to simulate fracture at the interface. When the material is fully damaged, a crack will grow according to the quadratic fracture energy criterion.

69.3 Finite Element Modeling

2D FE models are developed to study the phenomenon of delamination based on the cohesive zone method. A commercial FE code Abaqus© is used for the simulations of the double cantilever beam test and the drilling of FRP laminate. DCB test is performed to understand delamination at interfaces of FRP laminate. Then, this approach is further extended to simulate push-out delamination in drilling process.

69.3.1 Double Cantilever Beam (DCB) Test

DCB test [17] is used to calculate the fracture energies of different modes of interface failure. Figure 69.2 shows an FE model of a DCB specimen. Two FRP layers are adhesively bonded together with an adhesive (here epoxy) layer. Each FRP layer is of 10 mm in length and 0.5 mm in thickness. Their elastic behavior is modeled with orthotropic elasticity, and damage initiation is modeled with Hashin's damage model [14]. An interface of 0.01 mm thickness between two

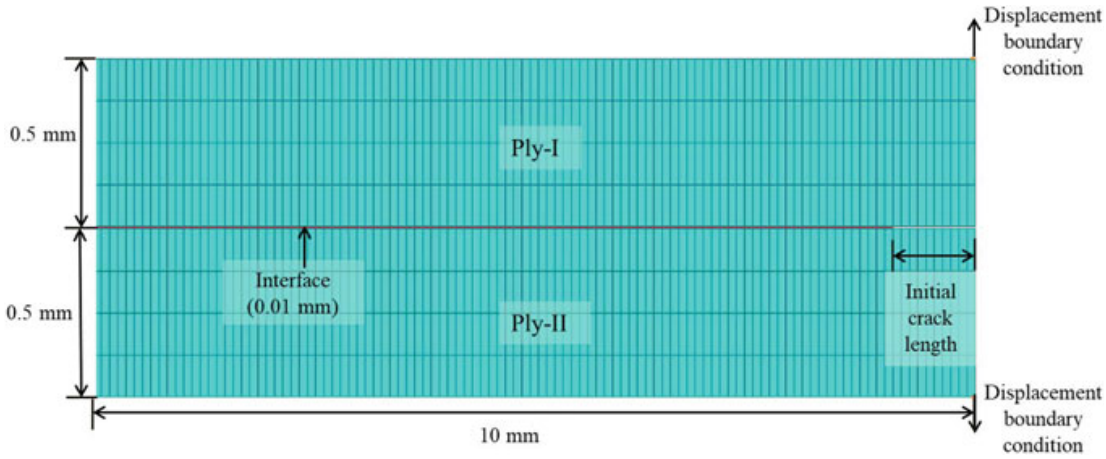


Fig. 69.2 FE model of double cantilever beam test

FRP layers is modeled based on CZM (see Sect. 69.2.2). Plain stress elements, 'CPS4R,' and cohesive elements 'COH2D4' are used to mesh FRP layers and interface layer, respectively.

The number of nodes on the adhesive layer and FRP layers is the same to achieve nodal connectivity at the interface. A displacement boundary condition is applied on the top and bottom corner of the nodes of FRP layers. Displacement of the nodes induces stresses in the elements which causes failure of the weakest section, i.e., interface. Quadratic damage criterion with an element deletion technique is used to simulate crack propagation along the interface. The load-carrying capacity of the interface in the mode-I type of delamination is predicted for different initial crack lengths.

69.3.2 Drilling of FRP Laminate

CZM can be effectively used to predict delamination of FRP laminate in a complex loading. Here, an attempt is made to predict delamination in drilling of FRP laminate. Drilling is a complex machining operation; hence, it is simplified by assuming no rotation of drill bit and a simple 2D FE model is developed. The main objective of the FE model is to predict the effects of thrust forces on the push-out delamination. Due to the simplification of the drilling process, this model does not take drilling torque into consideration. This model can predict delamination results with enough accuracy because push-out delamination is largely affected by thrust forces in drilling. The FE model of drilling (Fig. 69.3) shows a drill bit and four FRP plies that are adhesively bonded together with adhesive layers (epoxy). Drill bit is assumed to be rigid as it is 3–4 times harder than the FRP laminate. To minimize the computational time, whole geometry is scaled down. Each FRP layer is of 10 mm in length and 0.5 mm in thickness, and the diameter of a drill bit is taken as 1 mm. Their elastic behavior is modeled with orthotropic elasticity, and damage initiation is modeled with Hashin's damage model. Interfaces of 0.01 mm in thickness between the FRP layers are modeled based on CZM. Plain stress elements, 'CPS4R,' and cohesive elements 'COH2D4' are used to mesh the FRP layers and the interface layers, respectively. Element deletion technique is applied to delete failed elements based on damage criterion.

The interaction between the drill bit and workpiece is modeled using surface–surface contact available in Abaqus/Explicit. The contact is defined between the drill-bit surface and the FRP laminate nodes using sliding friction model with a coefficient of friction equal to 0.3. End nodes are fixed, and linear velocity in the vertical direction is applied to the drill bit. The effect of three different chisel-edge lengths on push-out delamination is predicted. The results obtained are discussed in the next section.

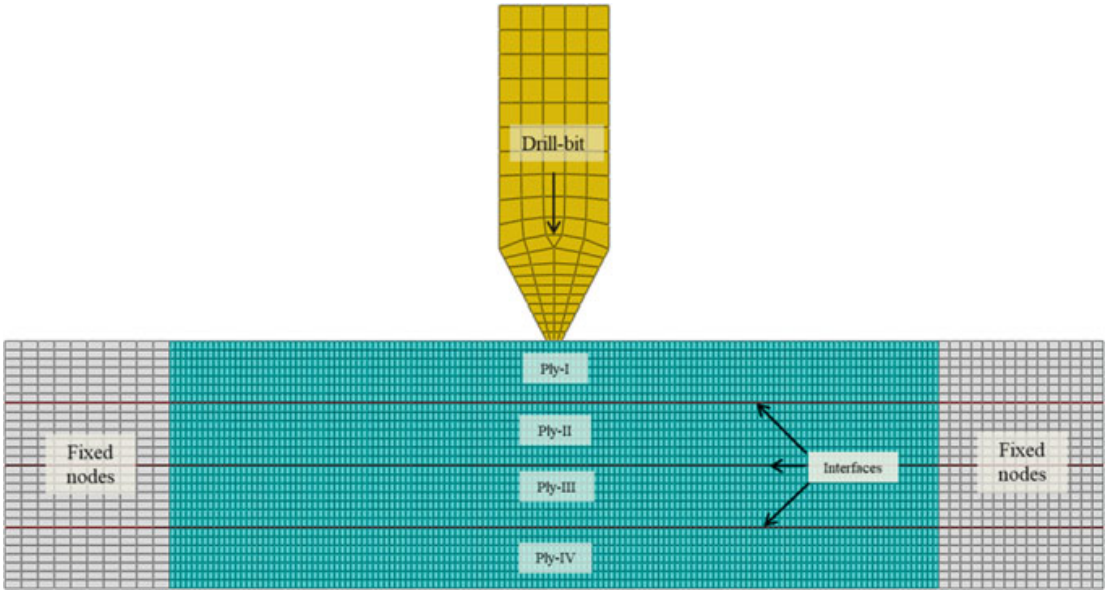


Fig. 69.3 FE model of drilling of FRP laminate

69.4 Results and Discussion

The results obtained in the simulation of double cantilever beam test and drilling of FRP laminate are discussed in the section.

69.4.1 Double Cantilever Beam (DCB) Test

These simulations were performed to predict the resistance offered by an interface for various initial crack lengths. Figure 69.4 shows a separation achieved at the interface by deleting the failed elements. Stresses in the *y*-direction (S22 in MPa) are the highest in the next critical elements at the interface. In this way, the crack propagates along the interface and delamination occurs. The simulation of DCB is run for five different initial crack lengths (0, 0.05, 0.25, 0.5, 1 mm). The initial crack helps in delamination; hence, the resistance offered by the interface layer reduces. This results in a reduction in load-carrying capacity with the size of initial crack length as shown in Fig. 69.5.

The shape of the load–displacement curve is triangular like a typical traction–separation curve that governs the behavior of an interface. The displacement indicates the separation between the nodes at the interface at the right edge of the beam (Fig. 69.4). The initial linear rise in the curves indicates the elastic behavior of the interface. The peak point is the maximum load-carrying capacity of the interface where the damage is initiated. Once the damage is initiated, the interface becomes weaker, its load-carrying capability goes down, and then it follows a linear softening law (Fig. 69.5). The load-carrying capacity is a strong function of an

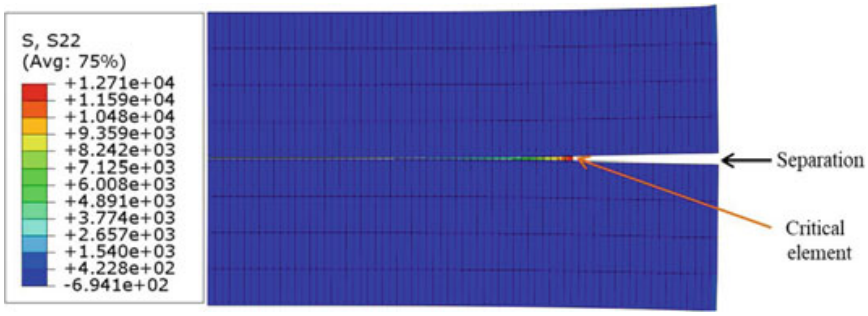


Fig. 69.4 Crack propagation along the interface during delamination

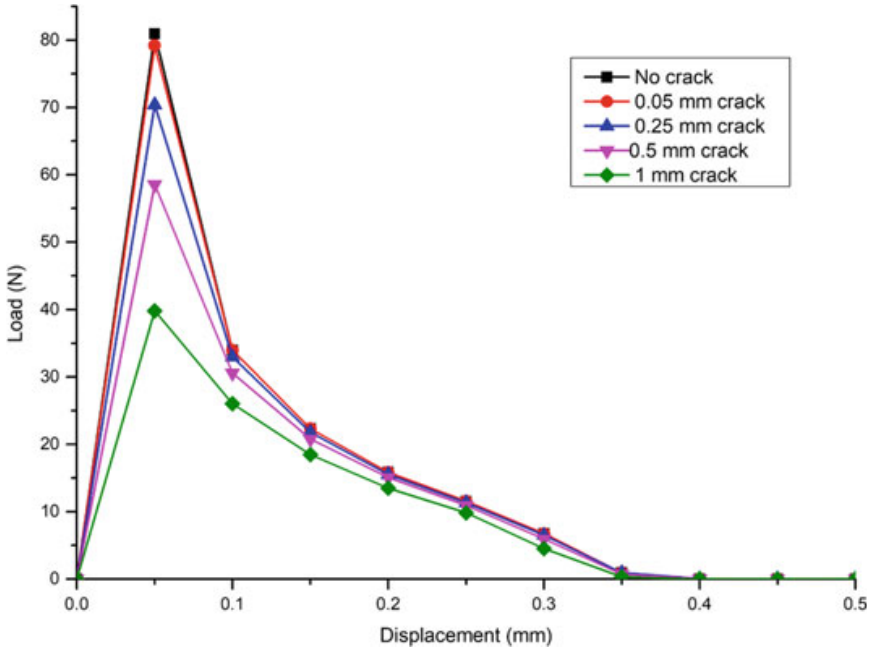


Fig. 69.5 DCB test results of load versus displacement for different initial crack lengths

initial crack length. The maximum resistance offered is 80 N with no initial crack, whereas 40 N was observed with a 1 mm initial crack.

69.4.2 Drilling of FRP Laminate

Simulation of drilling is performed for different chisel-edge lengths. For a given chisel edge, the drilling is achieved by deleting the failed elements (Fig. 69.6). As the drill bit moves in a vertically downward direction, the elements of the work-piece in contact with the drill-bit deform excessively. As the end nodes of the

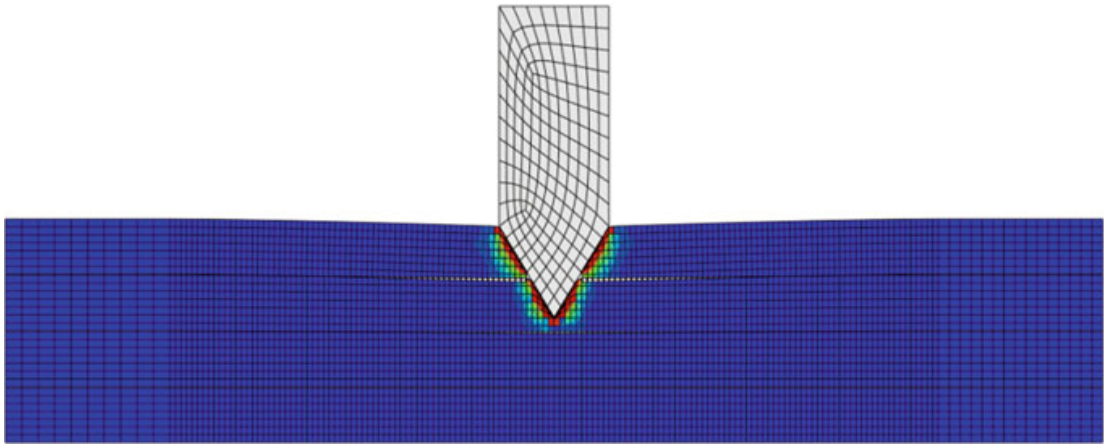


Fig. 69.6 Elements on the onset of failure in contact with drill bit

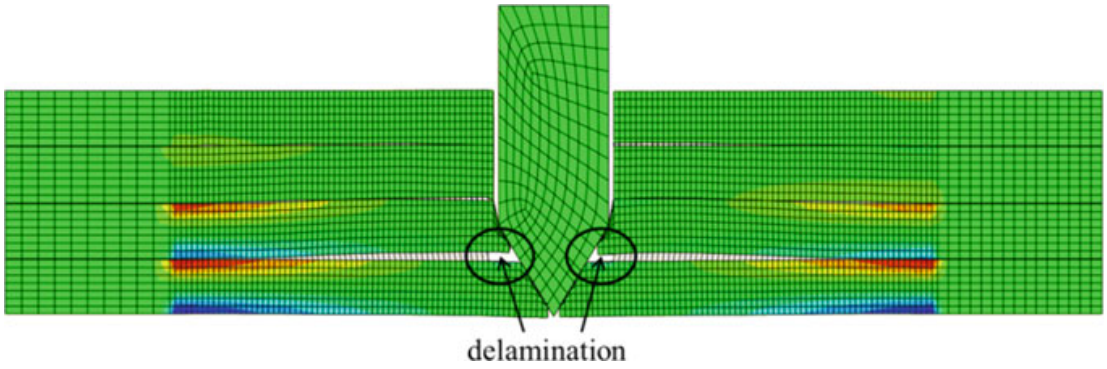


Fig. 69.7 Push-out delamination at the last interface in the drilling of FRP laminate

laminate are fixed, compressive and tensile stresses are induced in the laminate. When these stresses are enough to fail the FRP material, the damage is initiated based on Hashin's damage criterion.

Deletion of interface elements is based on quadratic damage criterion. In this way, simulation of a drilling process is accomplished in the 2D FE model. Drilling simulations are performed for three different chisel-edge lengths: 0, 0.15, and 0.35 mm, respectively. Figure 69.7 shows the delamination occurred at the last ply with a drill bit of '0 mm' chisel edge. The quantification of delamination is done using delamination factor $(F_d = D_{\max}/D)$, where $D_{\max}$ indicates the maximum diameter of the damaged zone and D is the actual hole diameter. As the length of the chisel-edge increases, there is an increase in the thrust force in drilling. Therefore, it is observed that delamination also increases with an increase in a chisel-edge length of a drill bit (Fig. 69.8).

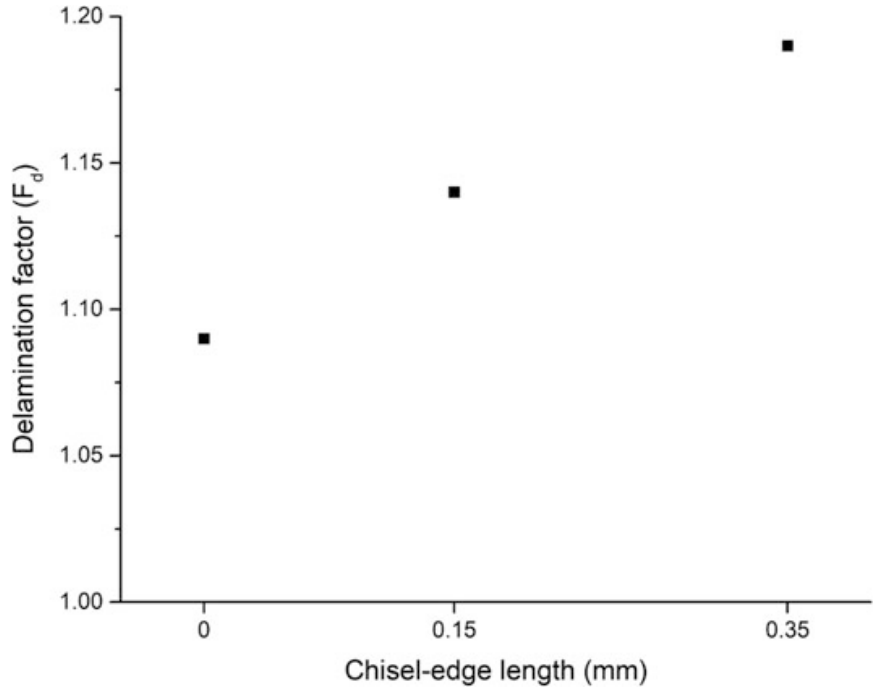


Fig. 69.8 Effect of the chisel-edge length on delamination

69.5 Conclusions

The finite element method can be effectively used to predict delamination in DCB test as well as in drilling of laminated composites. It was observed that load-carrying capacity of beam in double cantilever beam (DCB) test reduced by 50% for '1 mm' initial crack length. The available models to predict delamination in drilling are either analytical or the complex finite element models. These models have neglected the effect of chisel-edge length on the delamination. The present FE model is simple, takes less computational time, and produces results with enough accuracy. It takes a chisel-edge length effect into account. It was observed that push-out delamination increased by around 10% with an increase in the chisel-edge length from '0 mm' to '0.35 mm.' In addition, these models can be easily extended to predict the delamination for different drilling conditions and different FRP laminates.

References

1. Liu, D., Tang, Y., Cong, W.: A review of mechanical drilling for composite laminates. Compos. Struct. **94**(4), 1265–1279 (2012)
2. Davim, J.: Machinability of Fibre-Reinforced Plastics, vol. 4, pp. 1–70. DE Gruyter (2015)
3. Tao, C., Qiu, J., Yao, W., Ji, H.: The effect of drilling-induced delamination on tensile strength and prediction of residual strength of carbon fiber-reinforced polymer laminate. J. Compos. Mater. **50**(24), 3373–3384 (2016)

4. Gaitonde, V., Karnik, S., Rubio, J., Correia, A., Abrao, A., Davim, J.: Analysis of parametric influence on delamination in high-speed drilling of carbon fiber reinforced plastic composites. J. Mater. Process. Technol. **203**(1–3), 431–438 (2008)
5. Tsao, C., Hocheng, H.: Taguchi analysis of delamination associated with various drill bits in drilling of composite material. Int. J. Mach. Tools Manuf. **44**(10), 1085–1090 (2004)
6. Karnik, S., Gaitonde, V., Rubio, J., Correia, A., Abrão, A., Davim, J.: Delamination analysis in high speed drilling of carbon fiber reinforced plastics (CFRP) using artificial neural network model. Mater. Des. **29**(9), 1768–1776 (2008)
7. Mishra, R., Malik, J., Singh, I.: Prediction of drilling-induced damage in unidirectional glass-fiber-reinforced plastic laminates using an artificial neural network. Proc. Inst. Mech. Eng. Part B J. Eng. Manuf. **224**(5), 733–738 (2010)
8. Dhawan, V., Debnath, K., Singh, I., Singh, S.: A novel intelligent software-based approach to predict forces and delamination during drilling of fiber-reinforced plastics. Proc. Inst. Mech. Eng. Part L J. Mater. Des. Appl. **230**(2), 603–614 (2016)
9. Barbero, E.: Finite Element Analysis of Composite Materials using Abaqus, First edn. CRC press (2013)
10. Feito, N., Lopez-Puente, J., Santiuste, C., Miguélez, M.: Numerical prediction of delamination in CFRP drilling. Compos. Struct. **108**(1), 677–683 (2014)
11. Durao, L., Moura, M., Marques, A.: Numerical prediction of delamination onset in carbon/epoxy composites drilling. Eng. Fract. Mech. **75**(9), 2767–2778 (2008)
12. Phadnis, V., Makhdum, F., Roy, A., Silberschmidt, V.: Composites: Part A drilling in carbon/epoxy composites: experimental investigations and finite element implementation. Compos. Part A **47**, 41–51 (2013)
13. Mallick, P.: Fiber Reinforced Composites, Third edn. Taylor and Francis group (2007)
14. Hashin, Z.: Failure criteria for unidirectional fibre composites. J. Appl. Mech. **47**(June), 329–334 (1980)
15. Dassault Systèmes Simulia: Analysis User's Manual Volume 4. Abaqus 6.12, vol. IV, p. 29.3.8–10 (2012)
16. Abaqus Inc.: Analysis User's Guide, vol. III, p. 294 (2014)
17. Astm committee.: D5528-01: Standard test method for mode I interlaminar fracture toughness of unidirectional fiber-reinforced polymer matrix composites 1. Am. Stand. Test. Methods **3,** 1–12 (2014). (Reapproved 2007)

Chapter 70
Fuzzy Inference System-Based Neuro-Fuzzy Modeling of Electron-Beam Welding

Sanjib Jaypuria and Dilip Kumar Pratihar

Abstract Electron-beam welding, one of the high-energy density welding processes, is preferred for welding of nuclear, aerospace, and automotive materials. However, the input–output modeling of the welding is a challenging problem faced by many of the welding experts because of addition of new alloys to the industries. Soft computing techniques have been evolved in recent days to address this issue efficiently. This paper discusses two potential approaches of fuzzy inference system (FIS)-based neuro-fuzzy system for modeling of electron-beam welding response parameter. The following adaptive neuro-fuzzy inference systems (ANFIS): grid partition-based fuzzy interference system (G-ANFIS) and subtractive clustering-based fuzzy inference system (S-ANFIS) have been incorporated for the determination of depth of penetration in electron-beam welded copper plate in bead-on-plate configuration. In this study, the performances of above networks are compared on the basis of accuracy and computation time. Coefficient of correlation for S-ANFIS and G-ANFIS are found to be 0.98 in both approaches. Prediction of S-ANFIS is found to be 4 and 6% more accurate than results of G-ANFIS and regression model, respectively. Grid ANFIS consumes around 42 min to complete the computation, while it takes only 9.819 s for S-ANFIS. Based on the comparison, it is observed that the S-ANFIS model has the better prediction capability and less computational complexity than G-ANFIS and nonlinear regression model.

Keywords Electron-beam welding · ANFIS · Grid partitioning · Subtractive cluster · Fuzzy inference system

S. Jaypuria (✉) · D. K. Pratihar
Soft Computing Laboratory, Department of Mechanical Engineering,
Indian Institute of Technology Kharagpur, 721302 Kharagpur, India
e-mail: sanjibjaypuria@iitkgp.ac.in

D. K. Pratihar
e-mail: dkpra@mech.iitkgp.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_70

70.1 Introduction

Electron-beam welding (EBW) is an advanced fusion welding process, in which a high-speed electron beam is injected to the joining area of materials. The materials melt as the kinetic energy of the electrons transforms into heat upon impact. As a high-energy density process, EBW has inherent advantage of low-heat-affected zone and high aspect ratio, which yields relatively deep and narrow weld profile [1]. Electron-beam technologies are particularly useful in the industries like automotive, aviation, nuclear, medical, and railway. The primary input parameters, which affect the mechanical and metallurgical properties, are acceleration voltage, beam current, scanning velocity, focusing distance, and beam oscillation parameters [2]. Precipitated hardened Cu–Cr–Zr alloys have been selected as heat dissipation materials for the components of international experimental thermonuclear reactor (ITER) with high heat flux (HHF), such as the diverter, the limiter, and the first wall of the vacuum reactor [3].

Design of experiments, soft computing, and evolutionary computational networks have developed promptly and used for input–output modeling and optimization of welding [4]. Dutta and Pratihar [5] had modeled the dependence between input–output of gas tungsten welding (TIG) with three approaches, namely conventional regression, multilayer backpropagation neural network (BPNN), and genetic algorithm-tuned neural network (GANN). From the results, it was concluded that neural network is more robust in nature as compared to conventional regression. Moreover, GANN was found to outperform BPNN, as the former has fair chance of reaching the globally optimal solution. Jaypuria et al. [6] investigated the influence of EB welding input variables on mechanical properties and weld geometry of CuCrZr butt joint. A linear regression model had been developed to study the input and output relationship. Kim et al. [7] used a GA and a RSM simultaneously to predict the optimal welding conditions in the metal inert gas arc (MIG) welding process. Koleva [8] used statistical multiple regression techniques to estimate the depth and width of the welding schedule of EB welded stainless steels. Kanigalpula et al. [9] carried out electron-beam welding of copper plate and these bead-on-plate experiments were performed according to the central composite design. A statistical regression was performed to establish the input–output relationship of the EB joint. An optimization problem without constrains had been formulated to minimize the problem of spiking phenomena with maintaining the maximum average bead penetration. Results of genetic algorithm approach showed high accuracy with the real experiment results. Kanigalpula et al. [10] used genetic algorithm with penalty function approach to solve the complex constrained optimization problem associated with the EB welding of Cu–Cr–Zr alloy plates. They used four welding input parameters, namely voltage, beam current, scanning speed, and focusing distance to get suitable values of bead penetration, bead width, hardness of weld, and heat-affected zone.

Although many researches modeled input–output relation of EB welding for automation with respect to the accuracy of the model, there is no comparative study

about the computational complexity. A few studies have been reported on fuzzy logic (FL)-based analysis related to welding process. Most of the articles concentrated their studies regarding fuzzy inference system (FIS) parameter optimization through genetic algorithm (GA) and particle swarm optimization (PSO) algorithm. However, no literature is available for different kinds of FIS-based ANFIS model for EB welding. Here, the modeling of EBW is done through two proposed neuro-fuzzy approaches to predict the response. The following systems: (i) grid partition-based fuzzy inference system, known as G-ANFIS and (ii) subtractive clustering-based fuzzy inference system, known as S-ANFIS, have implemented for establishing the nonlinear dependence between the EBW input and output parameters. Accelerating voltage, beam current, welding speed, and focusing distance are used as inputs to the FIS model to predict the bead penetration. The performance and computational complexity are compared for all the developed approaches.

70.2 Experimentation and Data Collection

Bead-on-plate (BOP) welding of Cu–Cr–Zr alloy plates of the dimensions 100 mm × 65 mm × 10 mm is taken in the present consideration. The input process variables which influence the bead penetration (BP) of EB welding are accelerating voltage (V), beam current (I), welding speed (S), and focusing distance (D). These experiments were conducted by using response surface central composite design (CCD) matrix with three center points. Hence, a total of $2^4 + 2\times4 + 3 = 27$ combinations of experimentation were conducted. Since the factors were expected to have quadratic effects, three central points were taken and three replicates were considered for each combination of input variables. The CCD matrix having 81 experimental data is utilized for this study. The weld geometry of the finished and etched samples was studied under optical microscope and macroscope. To test the performances of the developed input–output models, 12 random cases out of the 81 combinations were selected after maintaining their respective ranges unviolated. Training and test data set for verification of developed model was taken from Kanigalpula et al. [10]. Table 70.1 shows the test data matrix with different combinations of process parameters.

70.3 Adaptive Neuro-Fuzzy Inference System (ANFIS)—Based Modeling

Adaptive Neuro-Fuzzy Inference System (ANFIS) is a combination of artificial neural network (ANN) and fuzzy Takagi–Sugeno system [11]. It is a powerful tool, for carrying out various supervised learning tasks, such as regression and classification. The membership function parameters of initial FIS are fine-tuned by the use

Table 70.1 Test data matrix for testing proposed methods

Accelerating voltage (kV)	Beam current (mA)	Welding speed (mm/min)	Focusing distance (mm)	Experimental value (mm)
55	90	600	370	6.144
55	90	800	370	5.879
50	70	600	375	3.096
55	110	800	370	6.349
60	90	800	370	5.551
60	90	800	370	5.759
60	70	600	365	4.661
60	70	1000	375	3.797
55	90	800	370	5.423
60	90	800	370	5.181
55	90	800	370	5.121
50	110	1000	365	4.74

of backpropagation algorithm or in combination with the least square method. Then, the network learns and trains the FIS-based network on the data provided. This popular fuzzy system consists of following components: (1) a rule base, (2) a database, and (3) a decision-making mechanism [12]. The rule base consists of all possible rules from the number and levels of input variables. The membership function is decided form the database and the inference mechanism helps to establish the relation between input variables and rules. In ANFIS architecture, the initial FIS clusters the entire region and describes the behavior of the rules in the fuzzy embedded region.

70.3.1 Grid Partition-Based Fuzzy Inference System (G-ANFIS)

G-ANFIS is first reported by Abonyi et al. [13] in 1999. G-ANFIS works on the principle of generating single-output Sugeno-type fuzzy inference system using a grid partition on the data. The initial FIS generates the membership function parameters randomly just like any ANFIS network. G-ANFIS uses the grid partition (axis-paralleled partition) of fuzzy on the basis of predefined number of membership functions and their dimension. ANFIS starts from zero output during the training; it gradually learns and refines the fuzzy rules and functions parameters. The number of fuzzy rules increases exponentially, when the number of internal input variables increases, so with the increase of number of rules, it becomes computationally expensive and impractical for industries where a large number of variables are associated with the process. The grid partition might be useful for the

system with small number of input variables mostly less than six and a large size of training data [14]. Here, four input variables are considered in this study. The proposed approaches have been developed with MATLAB 2015.

70.3.2 Subtractive Clustering-Based Fuzzy Inference System (S-ANFIS)

S-ANFIS [15] uses Sugeno-type FIS structure with subtractive grouping. S-ANFIS network is comparatively faster because of the faster estimation of cluster and cluster center in a given dataset. Here, each data point is assumed as cluster center and then defines the suitable cluster center of the dataset by calculating potential density of surrounding data points. In this manner, the first cluster center is chosen among the given data points and subsequently, the potential of data points near the cluster center is destroyed and second cluster center will be chosen based on potential density of surrounding data points. So, it becomes crucial to determine influence radius to construct one specific cluster and subsequent number of clusters with given data points. S-ANFIS returns a Sugeno-type FIS structure that models the data behavior and contains a set of fuzzy rules to cover the feature space [16].

70.3.3 ANFIS Models for Prediction of Depth of Penetration

In this work, both methods (G-ANFIS and S-ANFIS) are used for training of ANFIS and an attempt has been made to model and predict the response of EBW. Figure 70.1 shows the schematic view of developed ANFIS architecture for this study. A graphical user interface (GUI) of the updated rule base for a specific combination of input parameters is given in Fig. 70.2.

In the present study, out of the 81 input data pairs, 69 combinations are considered for training the FIS and 12 data pairs are used to test the accuracy and computation time. Takagi–Sugeno-type FIS structure of zero-order is used in both developed approaches. Gaussian and linear membership functions are considered in the current work for input and output layer, respectively. The reason behind choosing Gaussian function in input layer is the capability of the same to map the nonlinear process efficiently. The initial parameters associated with the membership function are randomly generated by ANFIS and tuned after subsequent training of FIS. A hybrid optimization algorithm, which is a mixture of backpropagation and least square method, is used to minimize the error (with error goal = 0) and a maximum iteration of 200. The number of membership functions is kept fixed at

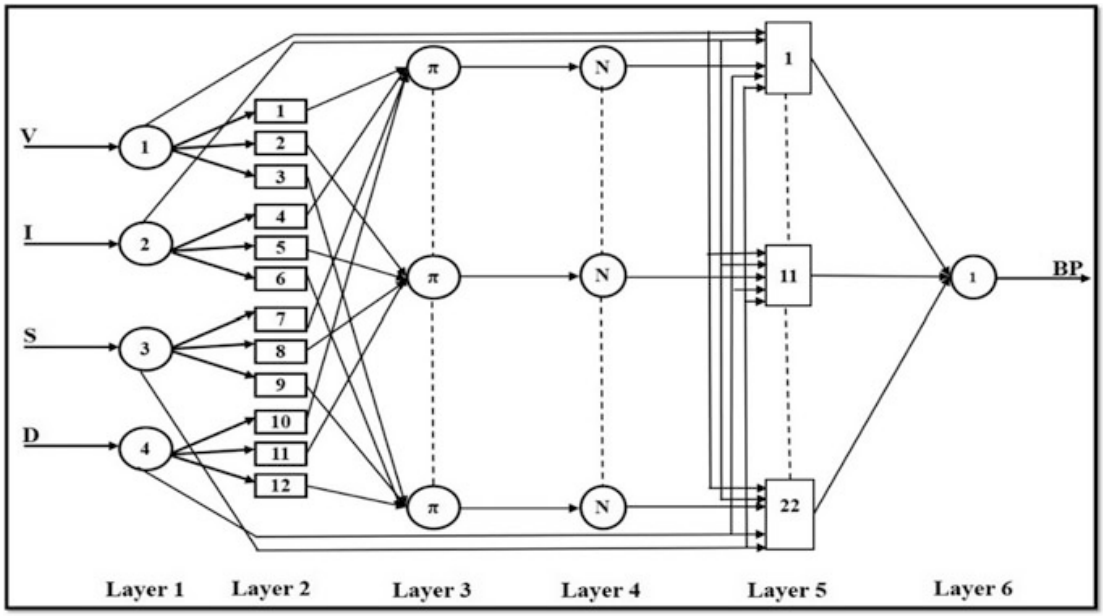


Fig. 70.1 ANFIS architecture

Fig. 70.2 Updated rule base for a specific input data combination [55, 90, 800, 370]

five, representing the linguistic variables like very low, low, medium, high, and very high. Table 70.2 represents the internal parameters associated with G-ANFIS and S-ANFIS.

Table 70.2 Internal parameters for proposed methods

Internal parameters	G-ANFIS	S-ANFIS
Number of nodes	1297	227
Number of linear parameters	3125	110
Number of nonlinear parameters	40	176
Total number of parameters	3165	286
Number of training data pairs	69	69
Number of checking data pairs	12	12
Number of fuzzy rules	625	22

70.4 Results and Discussion

70.4.1 Grid Partition-Based Fuzzy Inference System (G-ANFIS)

G-ANFIS model has been established to predict bead penetration as a function of welding input variables. The root-mean-square error (RMSE) values of G-ANFIS model are 0.15797 and 0.24419 for training and test data sets, respectively. The regression plots of the training data and test data sets are shown in Fig. 70.3 with the goodness of fit (R^2) values as 0.99635 and 0.98907, respectively. The total CPU time for execution of the program is 2510.9725 s. The computational complexity can be easily understood from 625 number of fuzzy (if-then) rules and 3165 number of internal parameters, which leads to curse of dimensionality [12, 14].

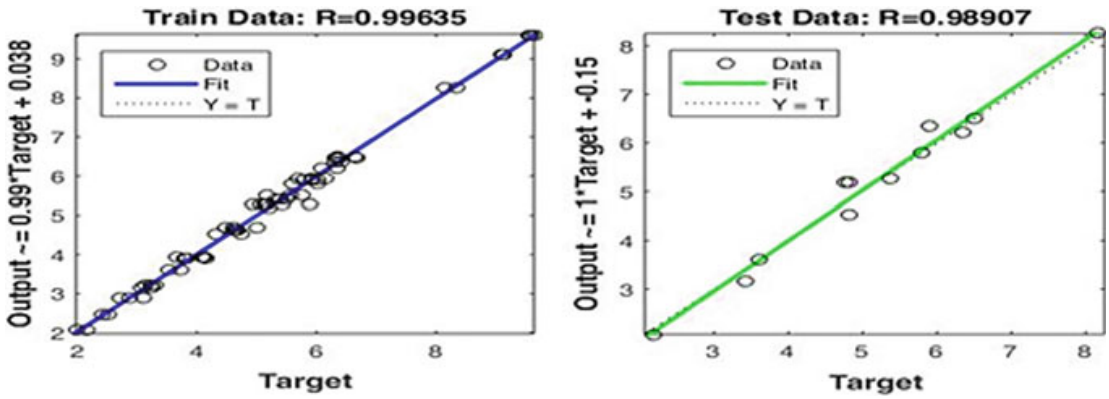


Fig. 70.3 Regression plots for training data and test data in G-ANFIS

70.4.2 Subtractive Clustering-Based Fuzzy Inference System (S-ANFIS)

The root-mean-square error (RMSE) values calculated for S-ANFIS architecture are 0.15507 and 0.23323 for the training and test data sets, respectively. The regression plots of the training data and test data sets are displayed in Fig. 70.4 with the (R^2) values as 0.99646 and 0.98915, respectively. The total CPU time for execution of the program is 9.819 s. Here, computation time is found to be quite less as compared to G-ANFIS because of elimination of unnecessary rules and dealing with only a few rules and internal parameters [15, 16].

70.4.3 Nonlinear Regression

A nonlinear statistical regression analysis is considered from available literature [10] to establish input–output relationships of the process and compare the results with other discussed ANFIS approaches. The generalized nonlinear regression expression is shown in below Eq. (70.1).

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i=1}^{k} \sum_{i<j}^{k} \beta_{ij} x_i x_j + \varepsilon, \tag{70.1}$$

where y is the response parameter; $\beta_0, \beta_i, \beta_{ii}$ …. are the coefficients to be computed using a least-squares method. Here, ε stands for the error in fitting. Nonlinear regression equation for bead penetration is evaluated following Eq. (70.2), and given below.

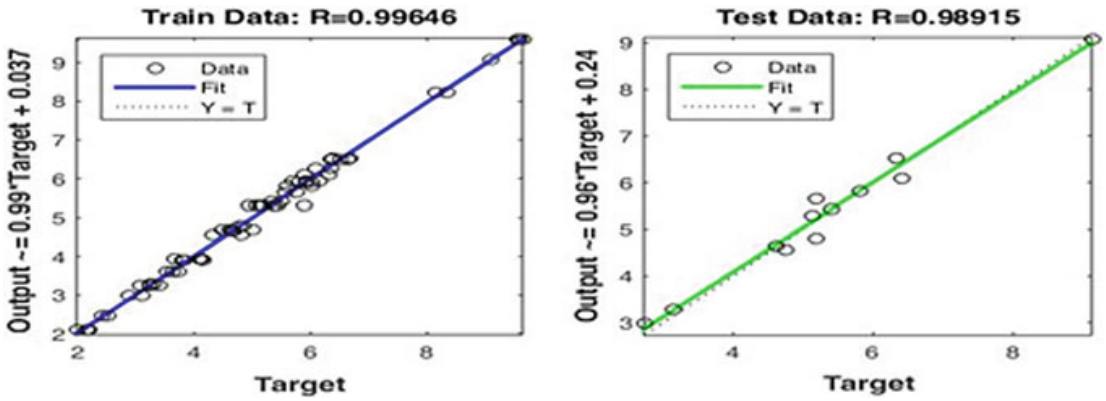


Fig. 70.4 Regression plots for training data and test data in S-ANFIS

$$\begin{aligned} BP_{uncoded} = {} & 5857.98 + 8.02798\,V + 0.0920434\,I + 0.0623438\,S - 33.0775\,D \\ & - 0.0313765\,V^2 - 0.00146853\,I^2 + 0.00000207134\,S^2 + 0.0456195\,D^2 + 0.0000113958\,VI \\ & - 0.0002530645\,VS - 0.0113766\,VD - 0.000145435\,IS + 0.00100352\,ID - 0.000114410\,SD \end{aligned} \tag{70.2}$$

70.4.4 Discussion

The RMS values for both training and test dataset in various developed ANFIS approaches are represented in Table 70.3. The least RMS value of 0.15507 and 0.23323 for training and test cases of S-ANFIS signifies a better convergence rate because of efficient tuning of internal parameters. Here, the model uses subtractive clustering technique with Gaussian and linear membership function for input and output layers, respectively. It is very fast enough to extract the set of rules that models the input–output relationship of a given system by effectively calculating the number of clusters and cluster center. Figure 70.5 shows distinctly that there is a satisfying similarity between the experimental value and the forecasted assessment of depth of penetration. The percentage of deviation of predicted data with respect to the test cases is also seen to be minimum in S-ANFIS-based approach, which is shown in Fig. 70.6. The computation time involved in S-ANFIS approach is least of all approaches. It is due to the selected number of rules chosen for the training of a specific response, so it deals with less number of fuzzy rules and internal parameters.

Grid portioning or G-ANFIS is also showing comparable RMS error values for both the training and test cases. This may be due to the inclusion of all possible combinations of fuzzy rules in the FIS, which ensure all relevant rules are accounted in the training process. Although this approach is precise, the computational complexity involved in this FIS is complex [17]. So, here, both the performance criteria, i.e., RMSE and computation time of the FIS are opposing in nature. The observed trend is certainly due to the involvement of all possible rules and accumulation of internal parameters during training. In addition to this, the RMSE value of train and test cases of nonlinear regression model was found to be larger than train and test results of S-ANFIS and G-ANFIS, respectively. The prediction accuracy of nonlinear regression and G-ANFIS models has shown 6 and 4% inferior results than that of S-ANFIS model. The computation complexity of S-ANFIS is also found to be significant lower than that of G-ANFIS.

Table 70.3 Performance comparison of the proposed approaches

Performance criteria	G-ANFIS	S-ANFIS	Nonlinear regression
RMSE (training cases)	0.15797	0.15507	0.21936
RMSE (test cases)	0.24419	0.23323	0.24644
Computation time (s)	2510.972	9.819	

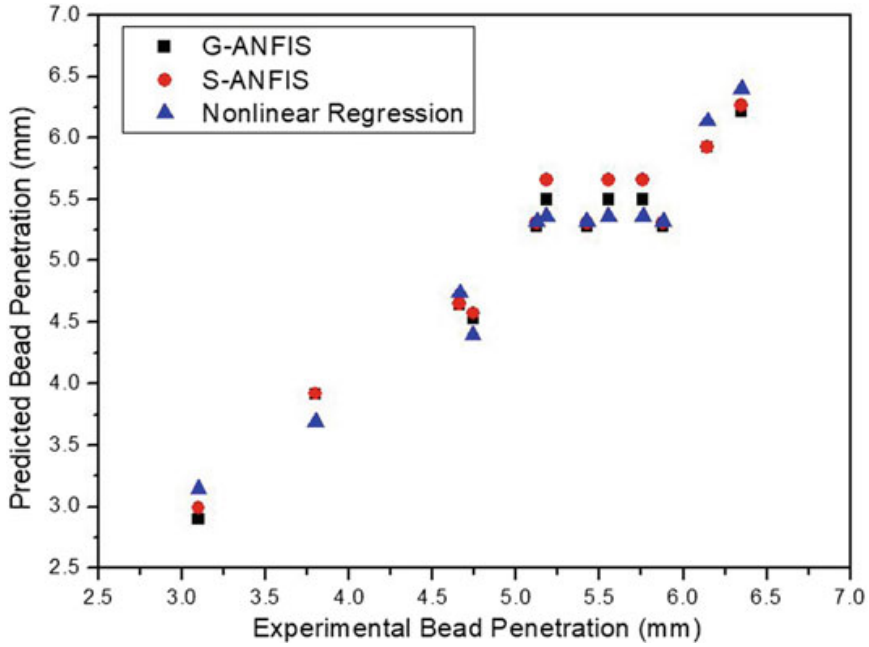


Fig. 70.5 Comparison of proposed approaches in terms of % deviation in prediction of bead penetration in test cases

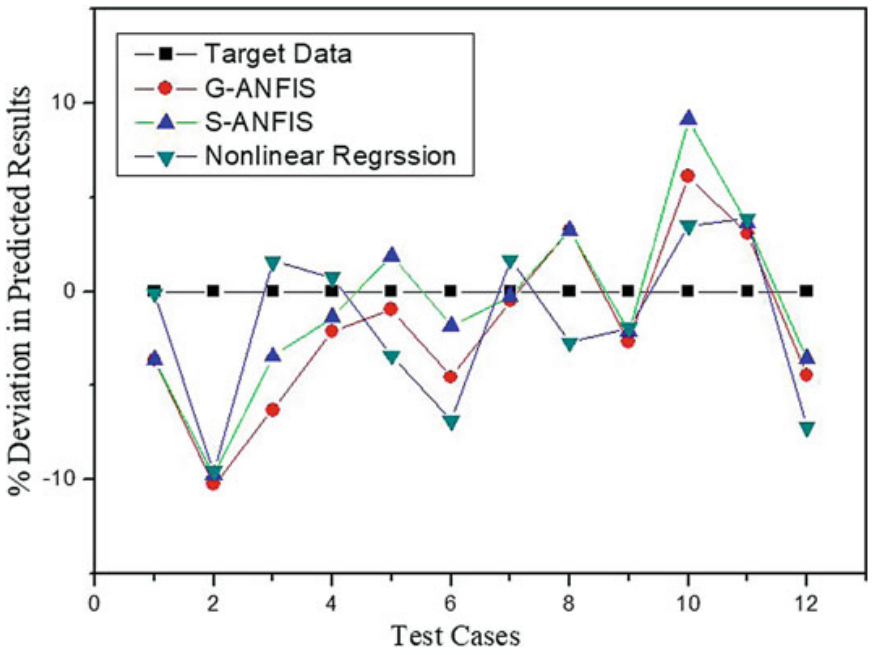


Fig. 70.6 Comparison between predicted and actual depth of penetration for the proposed approaches

Percentage deviation (Refer Fig. 70.6) result of S-ANFIS also signifies the good agreement with target data. Although the coefficient of correlation of both discussed approaches is nearly 0.98, the computational complexity of G-ANFIS creates difficulty for G-ANFIS model for real-time prediction. Therefore, it can be concluded

that S-ANFIS architecture is the more suitable one to make predictions without hampering the accuracy and computational complexity. It can also be used for real-time prediction due to their less consumption of computational resources and relatively higher accuracy in prediction.

70.5 Conclusion

The prediction of bead penetration using input variables of EBW is approximated by the FIS architecture based on grid partition and subtractive clustering. The developed model is verified with the test data. From the results, the following conclusions have been drawn:

1. Two different kinds of FIS generated models, namely G-ANFIS and S-ANFIS, are used to train the ANFIS network. The developed models have a good agreement with the experimental results. It can be validated with high value of coefficient of correlation value, which is 0.98 in both cases. Root-mean-squared error of the test cases is found to be equal to 0.24419 and 0.23323 for G-ANFIS and S-ANFIS, respectively. The computation time for G-ANFIS and S-ANFIS is found to be equal to 2510 s and 10 s, respectively.
2. The performance of above-discussed ANFIS models is also compared with nonlinear regression model, which is taken from the literature. S-ANFIS prediction accuracy in terms of RMSE is found to be 6% more accurate than the nonlinear regression results.
3. S-ANFIS model is found to be suitable among discussed approaches in terms of prediction accuracy and computational complexity. S-ANFIS is also more preferred for real-time prediction due to less consumption of computational resources.

In future, the performances of proposed approaches can be compared with that of fuzzy C-means clustering (FCM) approaches and evolutionary-tuned ANFIS approaches to check the suitability of the algorithm for automation and online inspection.

These tools are very useful for automation and controller design of any manufacturing process, where it involves costly trial and error experiments to find the suitable parameters for a desired response. The application of this tool is not limited to manufacturing industries, it can also be widely used for all domains of industries, where real-time or online modeling is required.

References

1. Schultz, H.: Electron Beam Welding. Abington Publishing, England (2004)
2. Fabritsiev, S.A., Zinkle, S.J., Singh, B.N.: Evaluation of copper alloys for fusion reactor divertor and first wall components. J. Nucl. Mater. **233**, 127–137 (1996)
3. Lacki, P., Adamus, K., Wojsyk, K., Zawadzki, M., Nitkiewicz, Z.: Modeling of heat source based on parameters of electron beam welding process. Arch. Metall. Mater. **56**(2), 455–462 (2011)
4. Benyounis, K.Y., Olabi, A.G.: Optimization of different welding processes using statistical and numerical approaches—a reference guide. Adv. Eng. Softw. **39**(6), 483–496 (2008)
5. Dutta, P., Pratihar, D.K.: Modeling of TIG welding process using conventional regression analysis and neural network-based approaches. J. Mater. Process. Technol. **184**(1–3), 56–68 (2007)
6. Jaypuria, S., Doshi, N., Pratihar D.K.: Effects of welding parameters on mechanical properties in electron beam welded CuCrZr alloy plates. In: IOP Conference Series: Materials Science and Engineering, vol. 338, issue No. (1), pp. 012013 (2018)
7. Kim, D., Rhee, S., Park, H.: Modelling and optimization of a GMA welding process by genetic algorithm and response surface methodology. Int. J. Prod. Res. **40**(7), 1699–1711 (2002)
8. Koleva, E.: Statistical modelling and computer programs for optimisation of the electron beam welding of stainless steel. Vacuum **62**(2–3), 151–157 (2001)
9. Kanigalpula, P.K.C., Jaypuria, S., Pratihar, D.K., Jha, M.N.: Experimental investigations, input-output modeling, and optimization of spiking phenomenon in electron beam welding of ETP copper plates. Measurement **129**, 302–318 (2018)
10. Kanigalpula, P.K.C., Pratihar, D.K., Jha, M.N., Derose, J., Bapat, A.V., Pal, A.R.: Experimental investigations, input-output modeling and optimization for electron beam welding of Cu–Cr–Zr alloy plates. Int. J. Adv. Manuf. Technol. **85**(1–4), 711–726 (2016)
11. Jang, J.S.: ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man. Cybern. **23**(3), 665–685 (1993)
12. Gowtham, K.N., Vasudevan, M., Maduraimuthu, V., Jayakumar, T.: Intelligent modeling combining adaptive neuro fuzzy inference system and genetic algorithm for optimizing welding process parameters. Metall. Mater. Trans. B **42**(2), 385–392 (2011)
13. Abonyi, J., Andersen, H., Nagy, L., Szeifert, F.: Inverse fuzzy-process-model based direct adaptive control. Math. Comput. Simul. **51**(1–2), 119–132 (1999)
14. Jang, J.S.: Input selection for ANFIS learning. In: Proceedings of IEEE 5th International Fuzzy Systems, vol. 2, pp. 1493–1499. IEEE (1996)
15. Cobaner, M.: Evapotranspiration estimation by two different neuro-fuzzy inference systems. J. Hydrol. **398**(3–4), 292–302 (2011)
16. Admuthe, L.S., Apte, S.: Adaptive neuro-fuzzy inference system with subtractive clustering: a model to predict fiber and yarn relationship. Text. Res. J. **80**(9), 841–846 (2010)
17. Mathworks.: Fuzzy Reasoning Toolbox: User's Guide (r2015b) (2016)

Chapter 71
Numerical Analysis of Erosion of Rough Surface

Ahad Noor, Saipraneeth Gouravaraju and Sachin Singh Gautam

Abstract This paper presents a numerical analysis of solid particle erosion of a rough surface. A total of ten particles are modelled to impact the target surface. For validation, first a smooth surface is impacted with rigid particles and the erosion results are compared with experimental results in the literature. After validation, erosion of rough surface is simulated using the finite element software ABAQUS. Rough surfaces with different average surface roughness are generated using inbuilt fast Fourier transform and random distribution functions in MATLAB. The coordinates of the generated rough surfaces are then imported to and superimposed on the mesh of the target surface generated in ABAQUS and finite element analysis is performed. It is observed that surface roughness affects the erosion rate significantly. For impact angles of 30° and 40° erosion is found to be less than that of the smooth surface.

Keywords Finite element analysis · Rough surface · Gaussian roughness · Johnson-Cook model · ABAQUS

71.1 Introduction

Precise prediction of lifespan of machine elements and replacement (wherever necessary) can avoid industrial hazards. There are several reasons due to which premature failure of engineering machine elements take place. Solid particles moving at high velocity is considered to be one of the main reasons for premature failure of machine elements. Study of erosion is very important in engineering

A. Noor · S. Gouravaraju · S. S. Gautam (✉)
Indian Institute of Technology Guwahati, Guwahati 781039, India
e-mail: ssg@iitg.ac.in

A. Noor
e-mail: ahad.noor@iitg.ac.in

S. Gouravaraju
e-mail: saipraneeth@iitg.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_71

applications where fluid travels at very high speed like oil and gas pipeline, rocket nozzles, compressor blades, drilling platforms, etc. When solid particles hit the surface repeatedly it causes material removal from the surface, this phenomenon is known as solid particle erosion (SPE). So, it would be beneficial to industries to predict life of machine elements for which the phenomena of erosion must be understood extensively. Understanding of erosion phenomena includes understanding loading condition, response of materials, materials properties, failure mechanism, and environmental factors like temperature and pressure. Erosion due to the impact of solid particle is one of the most extensively researched areas since 1960.

Finnie [1] performed analytical study of erosion. He considered a single non-spherical particle which has two sides, viz. flat side and cutting edge. He made an assumption that when particle strikes with flat side, erosion does not take place, and when it strikes with cutting edge, erosion takes place. He proposed a simple mathematical expression for prediction of erosion rate at different angles. When compared with experimental results, his equation was able to predict correct results at lower impact angles but not at normal impact. Later Bitter [2, 3] proposed that solid particle erosion cannot be explained with one model. He proposed that solid particle erosion takes place due to cutting action and plastic deformation. He solved energy equation for both the cases and provided mathematical expression to predict erosion rate at different impact angles. He developed concept of threshold velocity, according to which if a particle hits the surface with a velocity more than threshold velocity then only erosion takes place. His formula was simple but needed a constant, which was supposed to be determined experimentally.

By the era of Finnie and Bitter, it was well established that erosion is highly influenced by the impact angle. Although mechanism proposed by researchers mentioned above can predict erosion of ductile material, a generalized agreement could not be obtained even for simpler cases. So, to understand the erosion phenomena, some researchers concentrated on microscopic analysis by experiment. Hutching and Winter [4] attempted to explain the mechanism of erosion by scanning electron micrograph of impact zone. They used a 3 mm steel ball to impact surface of different samples made of aluminium, copper, and steel. They concluded that erosion of materials takes place due to shearing of material ultimately leading to lip formation.

Yerramareddy and Bahadur [5] performed experiments on Ti–6A–4V alloy substrate using silicon carbide particles at different impact angles. They studied the effect of heat treatment which includes effect of precipitation hardening microstructure. They reported that the temperature of heat treatment changes microstructure and so affects erosion rate. To understand mechanism of erosion, they studied scanning electron microscope image of impact zone. They reported that erosion rate depends on the particle's shape, i.e., blunt or sharp-edged. Depending on the shape, erosion takes place by different mechanism like cutting, ploughing, lip information, etc.

Oka et al. [6] performed an experimental analysis of solid particle erosion of 304 steel at different impact angles. They considered the effect of surface hardness also.

Based on the results of their experiment they proposed a predictive equation to estimate the erosion rate.

$$E_{90} = K(H_v)^{k_1}(v)^{k_2}(D)^{k_3} \tag{71.1}$$

where E_{90} is erosion in normal impact. H_v, v, and D are hardness, velocity, and diameter of particle, respectively. K accounts for the particle property factor, which includes angularity, hardness, etc., of the particle. k_1, k_2, and k_3 are exponents.

Erosion at a particular angle of impact $E(\alpha)$ is considered to be related to the erosion in normal impact E_{90} as

$$E(\alpha) = g(\alpha)E_{90} \tag{71.2}$$

where $g(\alpha)$ denotes the dependency of erosion on impact angle and is given as

$$g(\alpha) = (\sin\alpha)^{n_1}(1 + H_v(1 - \sin\alpha))^{n_2} \tag{71.3}$$

Although these equations as proposed by Oka et al. [6] predict the erosion rate correctly, they did not consider the effect of roughness.

Phenomena of solid particles erosion involve very high-velocity impact and very low interaction time, which makes it very difficult to study experimentally. So researchers were attracted by numerical methods to study erosion phenomena. Shimizu et al. [7] performed FE analysis of impact of a single particle at different angles. They concluded that for impact of spherical particle there is significant plastic deformation before failure. Before the development of numerical methods, it was very difficult to study the behaviour of erosion resistance coating and composite materials. Their thickness was determined with a trial and error basis. But numerical methods emerged as an important tool for analysis of coating and composite materials. Many researchers like Bielawski and Beres [8] performed a numerical analysis of composite materials.

Woytowitz and Richman [9] studied the effect of gap between impacting particles. Eltobgy et al. [10] performed finite element analysis of Ti–6A–4V composite and studied effect of impact angle and speed. Giriffin et al. [11] performed finite element analysis of steel coated with Al_2O_3. He used two different failure models for Al_2O_3 steel. Lui et al. [12] performed numerical analysis to investigate the effect of shape of the impacting particle.

Nelson et al. [13] performed erosion experiment at 30° impact angle with different rough sample made of mild steel. Yildierim and Muftu [14] have done analysis of particle impact on rough surface but they have not considered damage as they have performed their analysis in context of shot peening process. Anaraki et al. [15] has performed the analysis of erosion on rough surface and investigated effect of velocity, density, and diameter of particle but they have taken surface roughness value much high rather than in practical range and also they have not considered damage they assumed elements with 0.2% plastic strain as eroded volume.

From the literature survey, it was observed that extensive work has been done to analyse the effect of angel impact, size of particle, velocity of impact, etc. So it was decided to investigate the effect of surface roughness on erosion. We first validated our model which is described in the next section with experimental results of Oka et al. [6] Then different roughness were included in the model. Results were found to be consistent with the experimental results of Nelson et al. [13].

71.2 Model Description

In the present work, impact of ten particles striking at two closely spaced locations on the target material is simulated in ABAQUS/Explicit as shown in Fig. 71.1. The material considered here is 4340 steel. A stainless steel block of size 2 mm × 1.6 mm × 0.6 mm is considered as the target material. The bottom surface and all the vertical surfaces of the surface are assumed to be fixed (called the Encastre boundary condition in ABAQUS). The impacting particles are considered rigid. The diameter of particles is taken as 326 μm and mass of each particle is taken as 4.717×10^{-8} kg, which resembles with mass of silicon carbide particle of similar size. Rotary inertia of the particles is ignored. Particles are arranged such that each particle strikes the surface just after the impact of previous striking particle is complete, i.e., a rebound has occurred. Particles are placed at desired angles using rotation option in assembly module of ABAQUS. The particles are given an initial velocity of 104 m/s using predefined field option in ABAQUS. Particles are given velocity in *X*-and *Y*-direction depending on the angle of impact such that their resultant velocity comes to 104 m/s. Contact interaction between the particles and the rough surface is defined using surface to surface node contact algorithm, where the particle is considered as the master surface and all nodes of target are considered as slave surface. This contact algorithm provides an interaction between particle and target only, so no particle

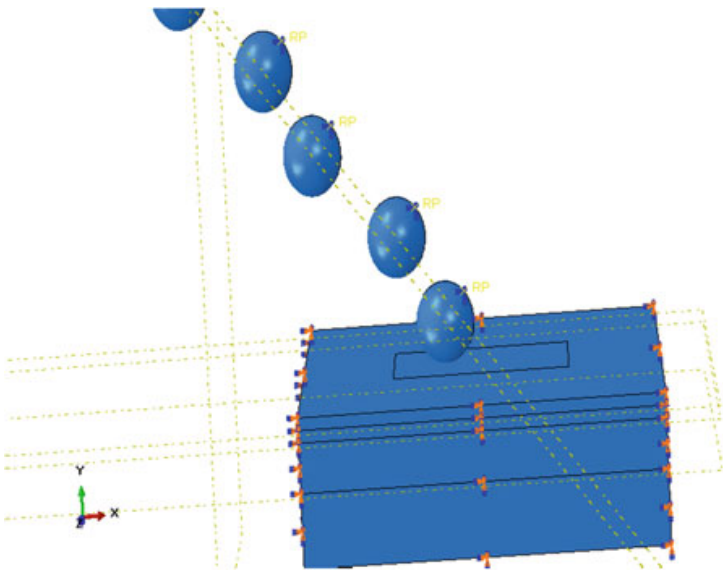


Fig. 71.1 Assembly of rigid particles and target material at an impact angle of 45° (only four particles are shown)

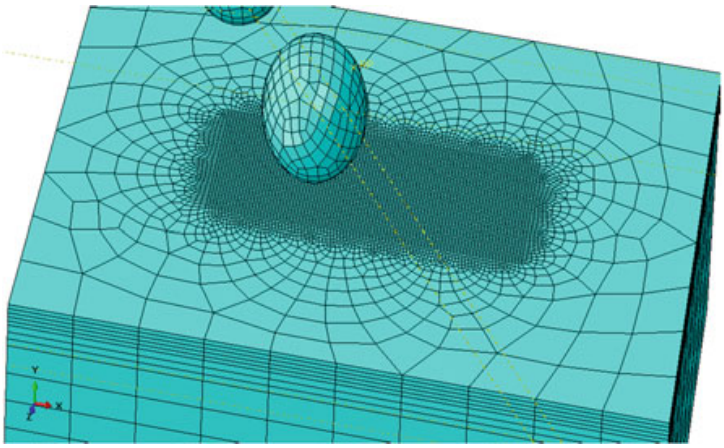


Fig. 71.2 Fine mesh near impact zone of 0.01 mm size while coarse mesh at other locations

collides during rebound of the previous particle. This helps us in placing particles closely and shortening the total simulation time. Friction between the surfaces is also considered with coefficient of friction value of 0.2, which was implemented using penalty, using interaction property option in ABAQUS. Heat generation due to friction and impact of particles were also considered. Eight nodded brick element (C3D8RT) with displacement degree of freedom is used for the analysis. Fine mesh of size 0.01 mm was used near impact zone while relatively coarse mesh was used in rest of the region as shown in Fig. 71.2.

The material data for 4340 steel is given in Table 71.1 [16].

Plastic properties are implemented using Johnson-Cook material model, which relates flow stress with equivalent plastic strain, temperature, and strain rate as shown in equation. Parameters of material model are given in Tables 71.2 and 71.3.

Table 71.1 Mechanical properties of 4340 steel

E (GPa)	v	ρ (kg/m^3)
200	0.3	7800

Table 71.2 Parameters for Johnson-Cook plasticity of 4340 steel

A (MPa)	B (MPa)	n	c	m	D_1	D_2	D_3	D_4	D_5
792	510	0.26	0.014	1.03	0.05	3.44	−2.12	0.002	0.61

Table 71.3 Composition of 304 and 4340 steel (%)

	C	Mn	Si	Ni	Cr	S	P
304 Steel	0.08	2.0	0.75	8–12	18–20	0.08	0.045
4340 Steel	0.42	0.70	0.28	1.83	0.79	0.0012	0.0009

$$\sigma_Y = \left(A + B\epsilon_p^n\right)(1 + C\ln\epsilon^*)(1 - T^*) \tag{71.4}$$

where ϵ_p is equivalent plastic strain, $\dot{\epsilon}^* = \frac{\dot{\epsilon}}{\dot{\epsilon}_0}$ is dimensionless plastic strain rate for $\dot{\epsilon}_0 = 1\ \mathrm{s}^{-1}$ and T^* is homologous temperature defined as

$$T^* = \left(\frac{T - T_{\text{room}}}{T_m - T_{\text{room}}}\right) \tag{71.5}$$

The Johnson-Cook fracture model is considered in this work to simulate damage. The damage parameter D is defined as

$$D = \Sigma\frac{\Delta\epsilon}{\epsilon_f} \tag{71.6}$$

where $\Delta\epsilon$ is increment in equivalent plastic strain and ϵ_f is the equivalent strain to fracture. Fracture is assumed to occur when D reaches value of 1. The expression for equivalent strain to fracture ϵ_f is given by

$$\epsilon_f = (D_1 + D_2\exp(D_3\sigma^*))(1 + D_4\ln\epsilon^*)(1 + D_5T^*) \tag{71.7}$$

where $D_i(i = 1, 2, 3, 4, 5)$ are the five constant. Their values are given in Table 71.2 [16].

In their erosion experiments, Oka et al. [6] used 304 stainless steel. Reliable Johnson-Cook damage parameters for 304 steel are not available in the literature [17]. So, in the current work, the results of Oka et al. [6] are compared with the results in the current work obtained with 4340 steel in a qualitative manner. Liu et al. [12] also considered the properties of ARMCO Iron for their FE simulations while comparing with the results of Oka et al. [6]. Although both the steels are of different grade, trend of failure follows the same pattern. A comparison of composition of both materials is shown in Table 71.3.

71.3 Results and Discussion

In this section, first the erosion from a smooth surface is discussed for impact velocity of 104 m/s and at six different impact angles, viz. 10°, 20°, 30°, 40°, 60°, and 90°, respectively. Results are validated with the experimental results of Oka et al. [6]. Then the erosion from the rough surface is simulated.

Initially, up to the impact of four to five particles no erosion occurred as the plastic flow and plastic deformation of the material is observed. With further impact, damage to the material starts and erosion takes place. Eroded volume is calculated and normalized with eroded volume in case of normal impact, i.e., for 90°. For calculation of eroded volume, number of elements deleted after the impact

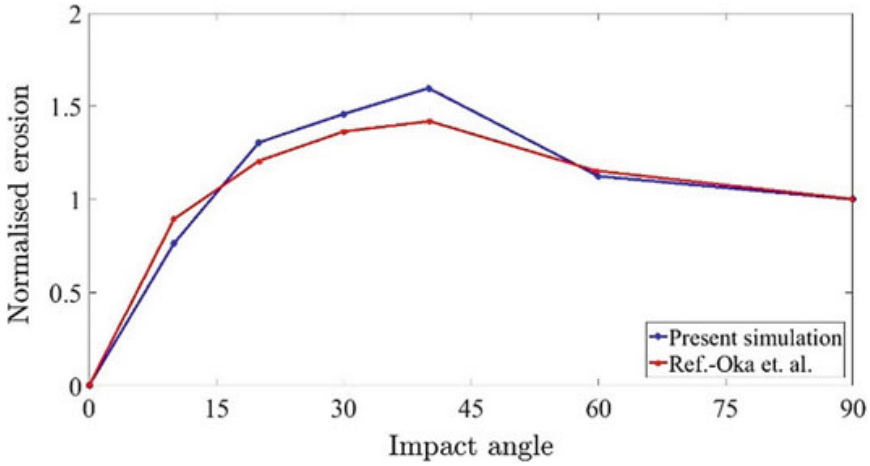


Fig. 71.3 Result for smooth surface and comparison to reference result

is counted. Then summing the volume of all the deleted elements gives us the total eroded volume. To convert eroded volume to dimensionless volume it is divided with eroded volume in case of normal impact. Results from the current finite element simulation are compared with results obtained by Oka et al. [6]. A comparison of present work and results obtained by Oka et al. [6] in their experiment is shown in Fig. 71.3. It can be seen that they are in good agreement. It is also observed that maximum depth occurs at normal impact but maximum volume occurs at 40° impact.

After validation, the effect of surfaces roughness on the erosion is studied. Four different rough surfaces with average surface roughness of 0.5 μm, 1 μm, 1.5 μm, and 2 μm are considered. Rough surface with autocorrelation length of 0.01 mm and mentioned average roughness values are generated using inbuilt Fourier transform and inbuilt random distribution function in MATLAB. Generated rough

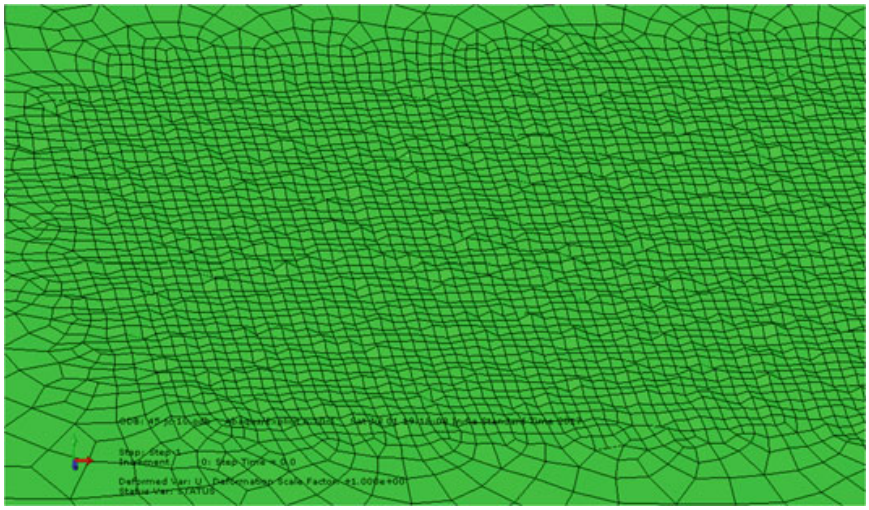


Fig. 71.4 Sample mesh of the target surface of the block after superimposing the coordinates of the generated rough surface

surface is superimposed on the top of the smooth surface discussed in the previous section (see Fig. 71.4). Then particle with a velocity of 104 m/s at different angles is impacted on the rough surface. Similar to smooth surface, eroded volume is calculated and normalized with volume of normal impact. Interestingly it is observed that at low impact angles, erosion decreased, and high angle is increased. Nelson et al. [13] also reported a decrease in erosion with increase in roughness for most of the cases but in one case they observed opposite trend also. At 30° impact angle, similar trend was observed in the current finite element simulation.

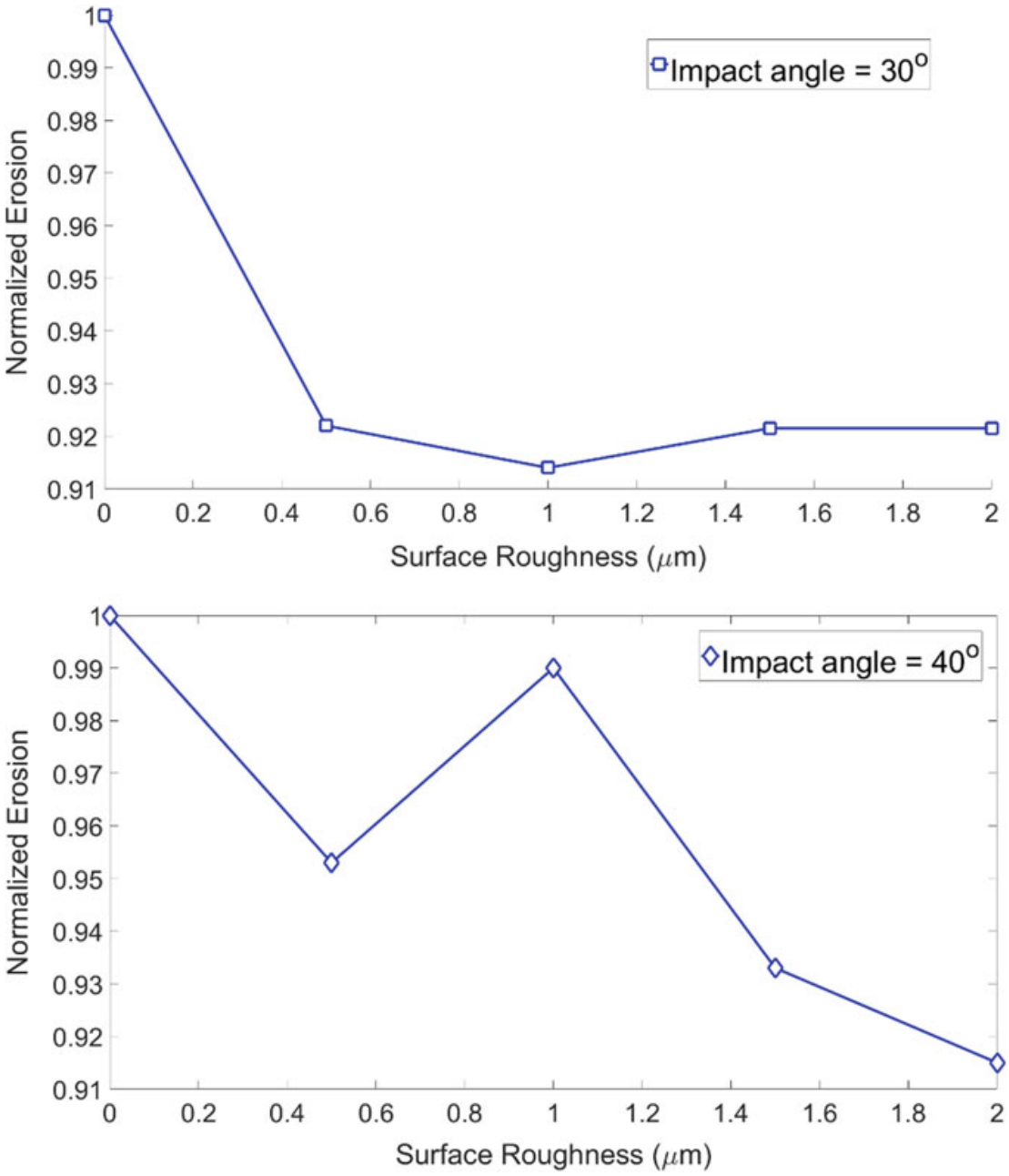


Fig. 71.5 Evolution of normalized erosion with surface roughness at impact angles of 30° and 40°

To quantify the effect of roughness, the eroded volume is normalized with the eroded volume of the smooth surface at each angle. Figure 71.4 shows the variation of normalized volume of erosion for rough surface with surface roughness for different impact angles. It can be seen that at lower impact angles, as the surface roughness increases, the erosion decreases. But at higher impact angles, erosion increases with increase in roughness. This observed behaviour is consistent with the observations of Nelson et al. [13] at an impact angle of 30°. (Figure 71.5)

71.4 Conclusion

A finite element analysis is carried out to study the erosion characteristics of a rough surface due to particle impact at different angles of impact. It is observed that surface roughness affects the erosion rate significantly. In the present analysis, two angles of impact are considered for the rough surface impact, viz. 30° and 40°. In both the cases, erosion rate is observed to be lesser than the erosion of smooth surface. This can be explained by Finnie's theory [1]. Finnie [1] argued that the surface roughness changes actual angle of contact for 30° impact. Results of the present numerical analysis follow the similar trend which was observed by Nelson et al. [13] experimentally.

Acknowledgements The authors are grateful to the SERB, DST for supporting this research under project SR/FTP/ETA-0008/2014.

References

1. Finnie, I.: Erosion of surface by solid particles. Wear **3**, 87–103 (1960)
2. Bitter, J.G.A.: A study of erosion phenomena part I. Wear **6**(1), 5–21 (1962)
3. Bitter, J.G.A.: A study of erosion phenomena part II. Wear **6**(3), 169–190 (1962)
4. Hutchings, I.M., Winter, R.E.: Particle erosion of ductile metals: a mechanism of material removal. Wear **27**(1), 121–128 (1974)
5. Yerramareddy, S., Bahadur, S.: Effect of operational variables, micro structure and mechanical properties on the erosion of Ti–6Al–4V. Wear **142**(2), 253–263 (1991)
6. Oka, Y.I., Okamura, K., Yoshida, T.: Practical estimation of erosion damage caused by solid particle impact: Part 1: effects of impact parameters on a predictive equation. Wear **259**(1–6), 95–101 (2005)
7. Shimizu, K., Noguchi, T., Seitoh, H., Muranaka, E.: FEM analysis of dependency of impact angle during erosive wear. Wear **233–235**, 157–159 (1999)
8. Bielawski, M., Beres, W.: FE modelling of surface stresses in erosion-resistant coatings under single particle impact. Wear **262**(1–2), 167–175 (2007)
9. Woytowitz, P., Richman, R.: Modeling of damage from multiple impacts by spherical particles. Wear **233–235**, 120–133 (1999)
10. ElTobgy, M.S., Ng, E., Elbestawi, M.A.: Finite element modeling of erosive wear. Int. J. Mach. Tools Manuf **45**(11), 1337–1346 (2005)

11. Griffin, D., Daadbin, A., Datta, S.: The development of a three dimensional finite element model for solid particle erosion on an alumina scale/MA956 substrate. Wear **256**(9–10), 900–906 (2004)
12. Liu, Z.G., Wan, S., Nguyen, V.B., Zhang, Y.W.: A numerical study on the effect of particle shape on the erosion of ductile materials. Wear **313**(1–2), 135–142 (2014)
13. Nelson, S.A., Baker, M.J., Deans, W.F.: The effect of surface roughness on the erosion of mild steel. In: International Pipeline Conference. American Society of Mechanical Engineers, pp. 803–810 (2004)
14. Yıldırım, B., Muftu, S.: Simulation and analysis of the impact of micron-scale particles onto a rough surface. Int. J. Solids Struct. **49**, 1375–1386 (2012)
15. Anaraki, A. P., Kadkhodapour, J., Taherkhani, B.: Simulation of erosion by particle impact on a rough surface. J. Fail. Anal. Prev. 784–789 (2014)
16. Johnson, G.R., Cook, W.H.: Fracture characteristics of three metals subjected to various strains, strain rates, temperatures and pressures. Eng. Fract. Mech. 31–48 (1985)
17. Bosetti, P., Maximiliano Giorgio Bort, C., Bruschi, S.: Identification of Johnson–Cook and Tresca's parameters for numerical modeling of AISI-304 machining processes. ASME. J. Manuf. Sci. Eng. **135**(5), 051021–051021-8 (2013)

Chapter 72
Numerical Simulation and Experimental Validation of Nanoindentation of Silicon Using Finite Element Method

Borad M. Barkachary and Shrikrishna N. Joshi

Abstract This paper reports finite element method-based numerical simulations of nanoindentation of Silicon to extract the mechanical properties, viz. Young's modulus and hardness. The simulations were carried out by using the Drucker–Prager constitutive model assuming a rigid spherical indenter of radius 300 nm to mimic the Hertzian contact which is typically observed underneath the Berkovich indenter. The dispalcement controlled boundary conditions comprising the depths of indentation as 45, 75, 95, 115, 130, and 140 nm were applied. Oliver–Pharr analytical method was used to obtain load-displacement plots. By using these plots, the values of Young's modulus and hardness were estimated. The results from the developed finite element simulation have also been validated by using the published experimental results. The estimated results from our simulations were found to be fairly matching with the experimental results. It was noticed that the FEM-based simulations provide an economical and fast option to the physical experiments in the estimation of mechanical properties of a material.

Keywords Nanoindentation · Silicon · Finite element method · Mechanical properties · Drucker-Prager material model

B. M. Barkachary
Department of Mechanical Engineering, Jorhat Institute of Science & Technology, Jorhat 781010, Assam, India
e-mail: borad.barkachary@gmail.com

S. N. Joshi (✉)
Department of Mechanical Engineering, Indian Institute of Technology Guwahati, Guwahati 781039, Assam, India
e-mail: snj@iitg.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_72

72.1 Introduction

Nanoindentation testing is useful to obtain the unknown mechanical properties such as hardness and Young's modulus. This method comprises of pressing the substrate material (specimen) against a harder material (usually a diamond indenter), and the material properties are estimated based on the force–displacement plots obtained during the loading and unloading sequence. Over the decades, various techniques such as tensile/compressive test, shear test, triaxial, and scratch test, nanoindentation have been developed to measure the mechanical properties of materials, viz. Young's modulus, Poisson's ratio, yield strength, strain-hardening, shear strength, hardness, etc. Out of these techniques, the tensile/compressive are commonly used for characterization of ductile materials. However, these are not suitable for brittle materials such as Silicon, because the brittle materials have higher yield strength than the fracture strength. Thus, during these tests, brittle materials fail before yielding. The applied stress reaches the fracture strength first, thereby fracturing the specimen before any appreciable strain can form. Nanoindentation test is a simple method, which determines the mechanical properties of materials such as hardness and Young's modulus. It is also used to obtain strain-hardening exponent, creep, fracture toughness, viscoelastic properties, and stress-strain data.

While some experimental work on Silicon exists, little theoretical work, especially in the arena of numerical simulations, has been reported. Our motivation is to simulate the Nanoindentation of Silicon using the currently available material models and evaluate how well the results from such models compare to those obtained in experiments. In sum, the objectives of this study were (i) to explore nanomechanical response of Silicon to quasi-static nanoindentation, (ii) to demonstrate the feasibility of modeling a typical hardness test using finite element analysis, (iii) to show how basic mechanical properties can be extracted from such simulated results, and (iv) to compare and contrast the simulated results with the experimental results.

72.2 Theoretical Background

Nanoindentation is a robust technique used to investigate the nanomechanical properties of materials, such as hardness, Young's modulus, fracture toughness, and incipient plasticity, all of which are vital for designing structural/functional elements of MEMS/NEMS devices. Nanoindentation entails a simpler setup and specimen preparation is compared to other mechanical experimental characterization methods. Typical nanoindentation involves pressing a hard indenter into a substrate material whose properties are to be determined.

During a nanoindentation process, the indenter is pressed into the material as shown in Fig. 72.1a, due to which both elastic and plastic deformation processes

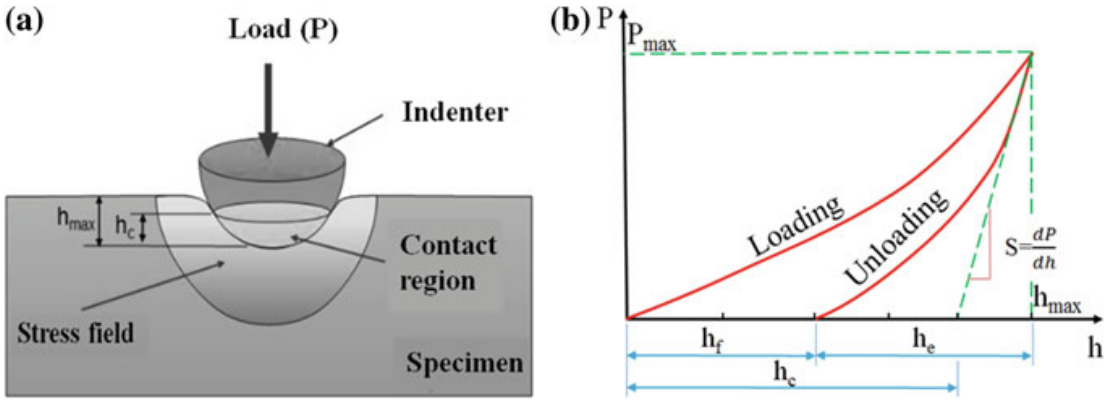


Fig. 72.1 Schematic of **a** nanoindentation process and **b** load–displacement plot

occur. This process produces an impression with a projected area A_p and surface area A_s of contact. This area depends on the geometry of the indenter and the contact depth, h_c. Basically, the nanoindentation process includes a loading and unloading cycle with or without dwell cycle. Figure 72.1b shows the indentation load (P) plotted against the displacement of the indenter tip (h) with respect to the top of the surface (before deformation) for one complete indentation cycle. The important parameters that have to be noted are the maximum depth (h_{max}) of penetration at peak load, the peak load (P_{max}) and the final residual depth after unloading (h_f). The slope S, also known as contact stiffness is determined by linearly fitting the upper 1/3th portion of the unloading curve. The contact stiffness, hardness, and Young's modulus are determined by using the Oliver–Pharr method [1, 2] as described in following Sects. 72.3.1 and 72.3.2 which is also mentioned in ISO 14577 and ASTM E2546.

72.3 FEM Modeling and Simulation of Nanoindentation of Silicon

With the advent of the computational techniques, it is now possible to use the FEA for analyzing complex engineering problems in a variety of engineering areas. Some of these problems include the study of heat transfer, contact mechanics, fluid dynamics, and electromagnetism [3]. In the present study, a commercial FE package ABAQUS™ was used to model the nanoindentation process of Silicon with a spherical rigid indenter of a known radius (300 nm). The use of a spherical indenter introduces a wider yet small stress field resulting in elasto-plastic deformation at shallow loads. Details of the model development are presented below.

72.3.1 Assumptions

As with any other engineering problem, there were several assumptions made while undertaking this study. First, it was assumed that no slippage occurs at the interface. Second, the contact between the indenter and the surface of the work material is assumed to be frictionless [4–6]. Workpiece and diamond indenter materials are assumed to be homogeneous and isotropic. The indentation process was carried out in quasi-static condition, so, temperature effect was neglected in the current simulation.

72.3.2 Geometric Modeling

The axisymmetric FEM model with its geometric dimensions is shown in Fig. 72.2. The dimensions of the work material were chosen in such a way that the deformation-induced stress waves reflected back from the fixed boundary can be avoided. Accordingly, the work material in this model was set to be an axisymmetric rectangular block of 2 μm × 2 μm. These dimensions are large enough to simulate a semi-infinite half-space for indentation with a maximum depth of 140 nm in this model. The rigid indenter was modeled with a radius of 0.3 μm.

72.3.3 Material Model

In this present simulation, a nanoindentation problem was considered to be based on continuum mechanics, where quasi-static loading of the indenter is applied. Since the von Mises yield criterion imply the dependence of material yielding solely on second deviatoric stress tensor and is independent of the first stress invariant, the yielding sensitivity to hydrostatic stress tensor is not incorporated for

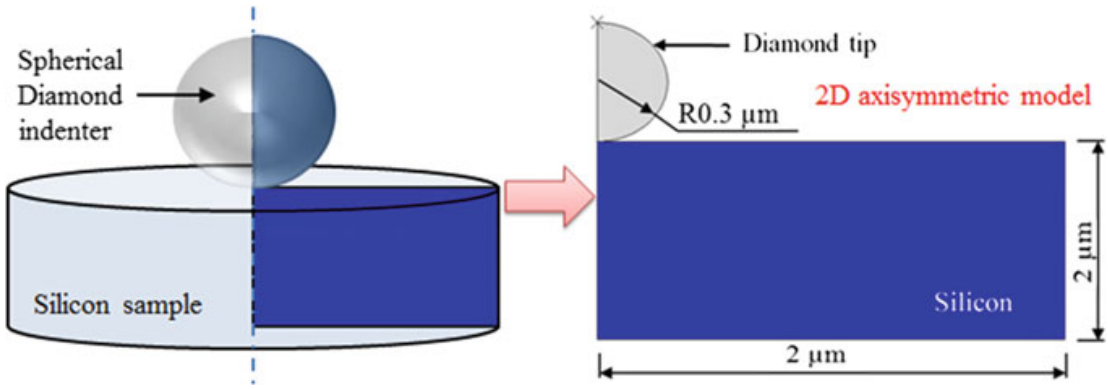


Fig. 72.2 Schematic of 2D axisymmetric indentation simulation model

pressure-sensitive materials. Drucker–Prager material model is used to model the material behavior of Silicon along with the elastic material properties. The Drucker–Prager yield criterion given by ref. [7] is a pressure-dependent model for determining whether the material has failed or undergone plastic yielding. The Drucker–Prager yield criterion can be described as a simple modification of the von Mises criterion, in which the hydrostatic stress component is also included to introduce pressure-sensitivity. The Drucker–Prager yield criterion is given by the following:

$$\sqrt{3 \cdot J_2} + I_1 \cdot \alpha - \kappa = 0 \tag{72.1}$$

where I_1 is the first invariant of the stress tensor, J_2 is the second invariant of the deviatoric stress tensor, α is the pressure-sensitivity coefficient, and κ is a material constant.

The quantity J_2 is given by the following:

$$J_2 = \frac{1}{6}\left[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2\right] \tag{72.2}$$

where σ_i are the principal stresses. The quantity κ is given by the following, which is equal to the yield stress in the case, when $\sigma_c = \sigma_t$, that is, no pressure dependency:

$$\kappa = \frac{2\sigma_c\sigma_t}{\sigma_c + \sigma_t} \tag{72.3}$$

where σ_c and σ_t are the yield stress in tension and compression, respectively.

The elastic materials properties for diamond and silicon are listed in Table 72.1 and Drucker–Prager material constants for silicon are listed in Table 72.2.

Table 72.1 Material properties of Diamond [7, 8] and Silicon [9]

Parameters	Unit	Diamond	Silicon
Density (ρ)	(kg/m^3)	3350	2328
Young's modulus (E)	(GPa)	1141	130
Shear modulus (G)	(GPa)	512	79.9
Poison's ratio (υ)		0.07	0.27
Fracture toughness (K_c)	(MPa m$^{1/2}$)	5.3 – 8	0.9
Thermal conductivity (k)	(W/m K)	700	150
Specific heat	(J/kg K)	520	700

Table 72.2 Constants for Drucker–Prager constitutive model for Silicon [10]

Parameters	Angle of friction (°)	Flow stress ratio	Dilation angle (°)	Yield stress (Nm^{-2})
Si	26	0.82	−20	15,867

72.3.4 Element and Meshing

Figure 72.3 shows the meshed models of the diamond tip and the workpiece. In the present work, two-dimensional, four-node quadrilateral element CAX4R was employed to discretize the model geometries. It is four-node quadrilateral element and provides reduced integration with hourglass control.

This element is chosen here because it is relatively inexpensive in terms of computation cost and time for problems involving non-linear constitutive behavior. For this type of element, the material calculations are only done at one point in each element by integrating all four nodes. Moreover, quadrangle element is one order higher than triangular element due to which it gives more accurate results. Also, by using this type of element, the symmetrical three-dimensional results can be viewed. The solution accuracy is dependent upon the type of mesh and its

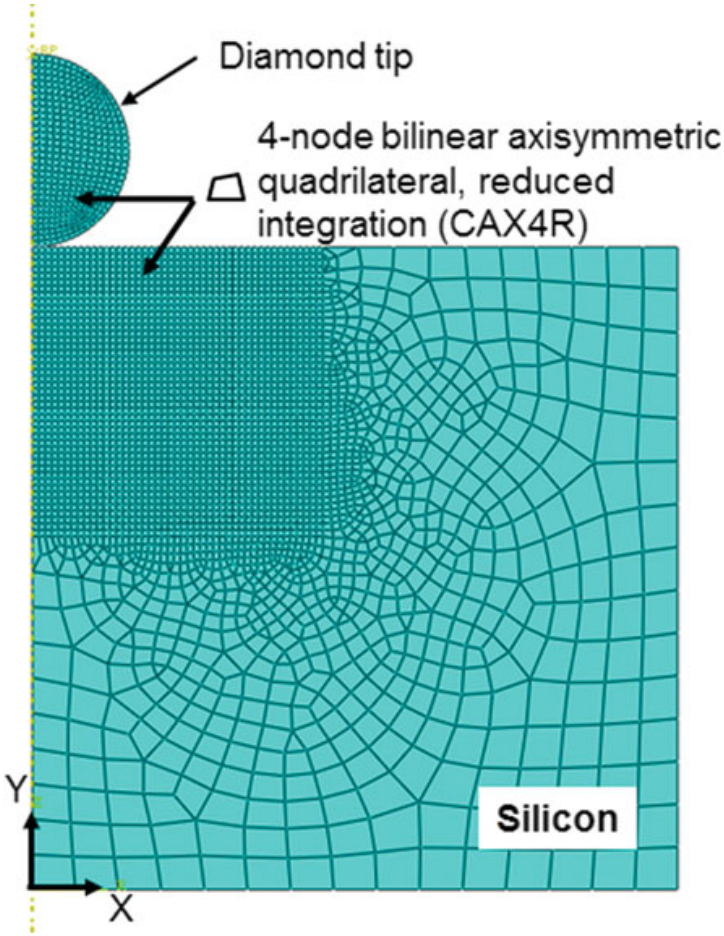


Fig. 72.3 Meshed geometries of the diamond tip and the workpiece

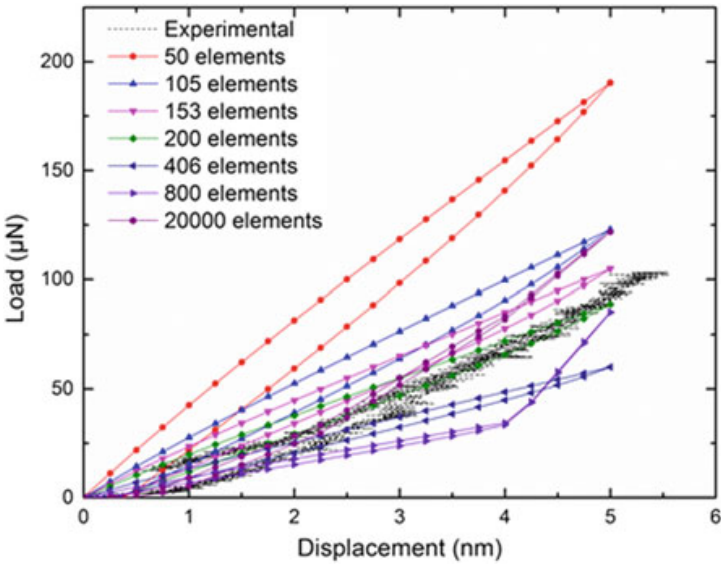


Fig. 72.4 Comparison of load–displacement plot with no of elements in indentation zone

refinement. Modification of mesh, known as *mesh refinement* is important to obtain a more accurate solution in any FEM simulation. The silicon block is discretized using fine mesh near the indentation zone, and coarse meshing is done away from the indenter to reduce the computational time without compromising the accuracy. The mesh sensitivity analysis has been carried out for a typical process condition: indenter tip radius—300 nm, max loading depth—45 nm, loading speed—0.1 mm/s. These process conditions are the same as that of the experimental conditions mentioned in the published literature by Rao et al. [11]. From the analysis, it was seen that the load–displacement plots at 20,000 elements are closer to the experimental results [11] in comparison with those obtained for the remaining cases as shown in Fig. 72.4. Therefore, the element size of 0.00001 mm, which generates 20,000 elements to discritze the geometries, was chosen for the discretization of the indentation zone.

72.3.5 Modeling of Surface Contact

The bottom surface of the indenter and the top surface of work materials formed a contact pair in the present numerical model. Surface-to-surface contact has been assigned between the contact pairs. Based on the suggestions from the literature ref. nos. [5, 6, 12], frictionless contact property is enabled between the two contact surfaces. It is because only a small portion of the indenter is considered to be in contact with the surface of the workpiece.

72.3.6 Boundary Conditions

Figure 72.5 shows the boundary conditions applied to the tool and workpiece. A fixed boundary condition is applied to the base of the work material. All the nodes on the base cannot move in any directions. The X-symmetry boundary condition is applied to the vertical surface of the workpiece as shown in Fig. 72.5.

It means that the workpiece geometry and diamond tip are symmetric along with the symmetry line and both form a revolving 3D geometry around the line. Rigid body constraint is applied to the diamond indenter at the reference point. First, it is restrained in the *X*-direction along with all the rotational degrees of freedoms. Since the process is very slow, i.e., quasi-static loading condition is assumed. Therefore, the load applied to the indenter is the displacement/velocity load of 0.1 mm/s in the current simulation. It is applied to the reference point of the diamond tip. As the contact area is very small and the loading speed is very slow, the effects of variation in temperature are not considered in the current simulation.

72.3.7 Solution Methodology

The numerical simulation of nanoindentation has been carried out in two steps: loading and unloading. The nanoindentation process includes complex interaction between the diamond tip and the workpiece along with large elastic-plastic

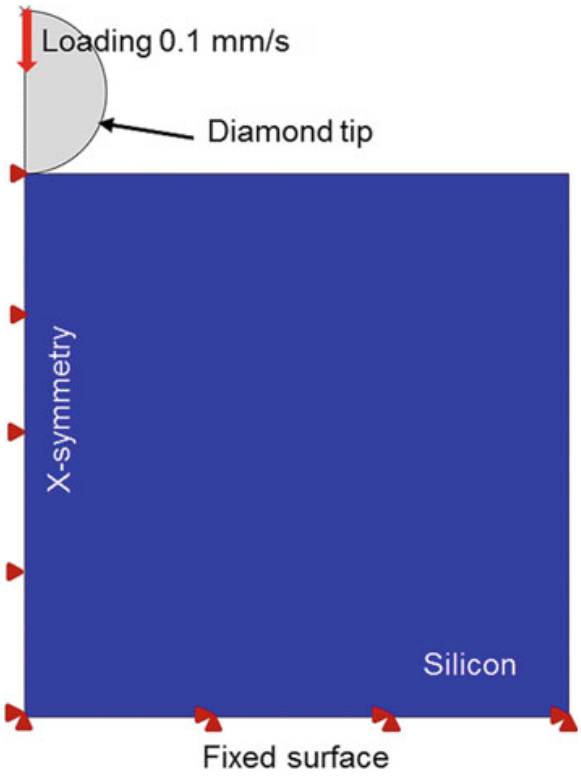


Fig. 72.5 Schematic of the boundary conditions applied

deformation of the workpiece during the nanoindentation process. Due to this complexity, both the steps have been solved by following the non-linear dynamic explicit solution step method with automatic (self-adaptive) time-step algorithm. In the dynamic explicit automatic time-step algorithm, the time increment is adjusted automatically to meet the corresponding convergence rate. The loading behavior of nanoindentation is simulated by increasing the indenter displacement in small increments. During the loading steps, the indenter is moved in the negative *Y*-direction, i.e., downward direction with a velocity of 0.1 mm/s and indents the work material to the maximum indentation depth. During the unloading step, the indenter returns to its original position incrementally. For each loading and unloading steps, 20 frames have been generated to reduce the data size.

72.4 Results and Discussion

The most common method for measuring hardness and Young's modulus using nanoindentation methods is the Oliver–Pharr analytical method. The underlying assumptions for the analytical model are:

(i) Deformation upon unloading is purely elastic
(ii) The compliance of the sample and the indenter tip is considered as springs in series
(iii) The process continuum is considered to be a homogeneous isotropic elastic half-space.

72.4.1 Measurement of Young's Modulus

From the extracted load–displacement plot, maximum load, corresponding maximum displacement at maximum load and the final displacement after the unloading was recorded. In the present case, the maximum load is 1275 μN, and maximum displacement is 45 nm. The displacement after unloading is 38.18 nm. It has been noted that there is no fixed rule for considering the unloading data points. The unloading data points are needed to be chosen by looking at the graphical data before deciding on the fitting parameters for analysis. Usually between 25 and 100% of the unload curve data are commonly used for the fitting function depending on the quality of the data [2, 13]. In the present case, we have found that 1/3th of the data points give good results while calculating Young's modulus and hardness. From the unloading curve, 1/3rd data points of the total unloading data points are linearly fitted, and the corresponding slope values were obtained as shown in Fig. 72.6.

Corresponding hardness and Young's modulus values were calculated for each depth of indentation following the Oliver–Pharr method. The stiffness of the contact

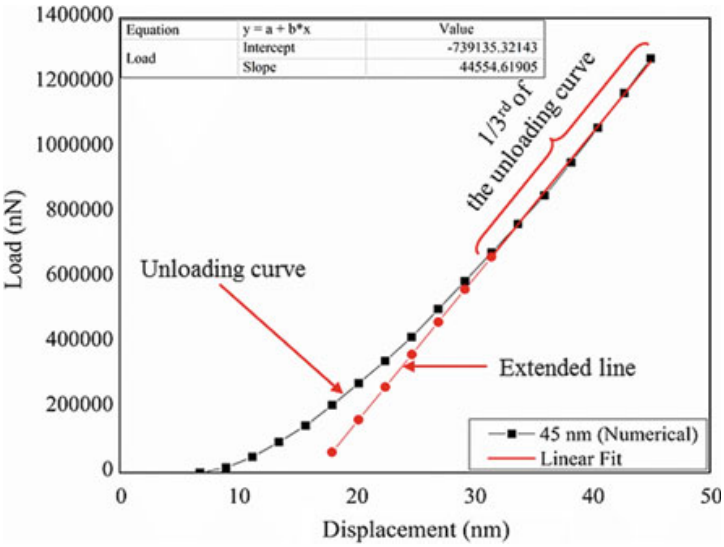


Fig. 72.6 Fitting the unloading curve and finding the slope

between the indenter and the material being tested is required to determine the mechanical properties of interest. The stiffness, S, is determined from the initial slope of the unloading curve.

$$\text{Stiffness}\,(S) = \frac{\mathrm{d}P}{\mathrm{d}h} \tag{72.4}$$

From the linearly fitted line, the slope is calculated using the linear line equation as shown in Fig. 72.6, which is nothing but the stiffness (S) = 44555 N/m^2. Once the stiffness is measured using (72.4), the reduced modulus E_r can be determined as:

$$E_r = S\frac{\sqrt{\pi}}{2\beta\sqrt{A}} \tag{72.5}$$

where S is the contact stiffness, A is the contact area, and E_r is the reduced modulus which accounts for the measured elastic displacement contributing both from the sample and the indenter tip. Depending on the indenter tip used, contact radius is calculated. In the present case, for spherical indenter, the contact radius is computed as

$$\text{Contact radius}, a = \sqrt{h(D-h)} = \sqrt{45(600-45)} = 158\,\text{nm} \tag{72.6}$$

where D is the diameter of the spherical indenter, and h is the maximum depth at maximum load. The contact area between the indenter tip and workpiece at maximum displacement is calculated as

$$\text{Contact area}, A = \pi a^2 = 7.85 \times 10^{-14}\ \text{m}^2 \tag{72.7}$$

After finding out the stiffness (S), contact area (A), and taking indenter geometry correction factor β = 1, the reduced modulus (E_r) is calculated as 1.41×10^{11} N/m^2

(by using the Eq. (72.5)). Next, the reduced modulus can be used to calculate the actual modulus of the sample, using the following relationship:

$$\frac{1}{E_r} = \frac{1 - \upsilon_i^2}{E_i} + \frac{1 - \upsilon_s^2}{E_s} \tag{72.8}$$

where $E_i = 1141$ GPa and $\upsilon_i = 0.07$ are the known values of the Young's modulus and Poisson's ratio for diamond indenter, respectively. Thus, by using Eq. (72.8), the actual modulus E_s for silicon sample is found to be $E_s = 153$ GPa.

72.4.2 Measurement of Hardness

Hardness is the resistance to the permanent plastic deformation-induced in the material and is expressed as the ratio of the applied force ($P_{\max}$) to the corresponding projected contact area (A) as shown in Eq. (72.9):

$$\text{Hardness}, H = \frac{P_{\max}}{\text{Contact area}} \tag{72.9}$$

By using Eq. (72.9), the hardness value of silicon were calculated as $H = 16.25$ GPa for $P_{\max} = 1275$ μN and $A = 7.85 \times 10^{-14}$ m^2. After determining the required output parameters, the results were validated with the published experimental results. These are presented in the following section.

72.5 Experimental Validation of Nanoindentation Simulation

Nanoindentation simulations were carried out for silicon workpiece to validate the developed finite element-based nanoindentation model with experimental results [11]. All the process conditions in the simulations were set similar to the experimental process conditions mentioned in the Rao et al. [11] except the indenter geometry as the simulation was carried out in 2D. In numerical simulation spherical indenter tip with a diameter of 600 nm was considered, whereas, in the experiment, indentations were performed with a Berkovich indenter. The approximation of Berkovich indenter with a spherical indenter with nanometric radius is still valid as the tip of the Berkovich indenter is not completely sharp. Plots of experimental and numerical simulation results in terms of the load–displacement graph for indentation depths 45, 75, 95, 115, 130, and 140 nm are shown in Fig. 72.7. From the figure, it was observed that numerical results are in good agreement with the experimental results; however, the numerical results over predict the experimental results. Using Oliver–Pharr analytical procedure explained in previous Sects. 72.4.1

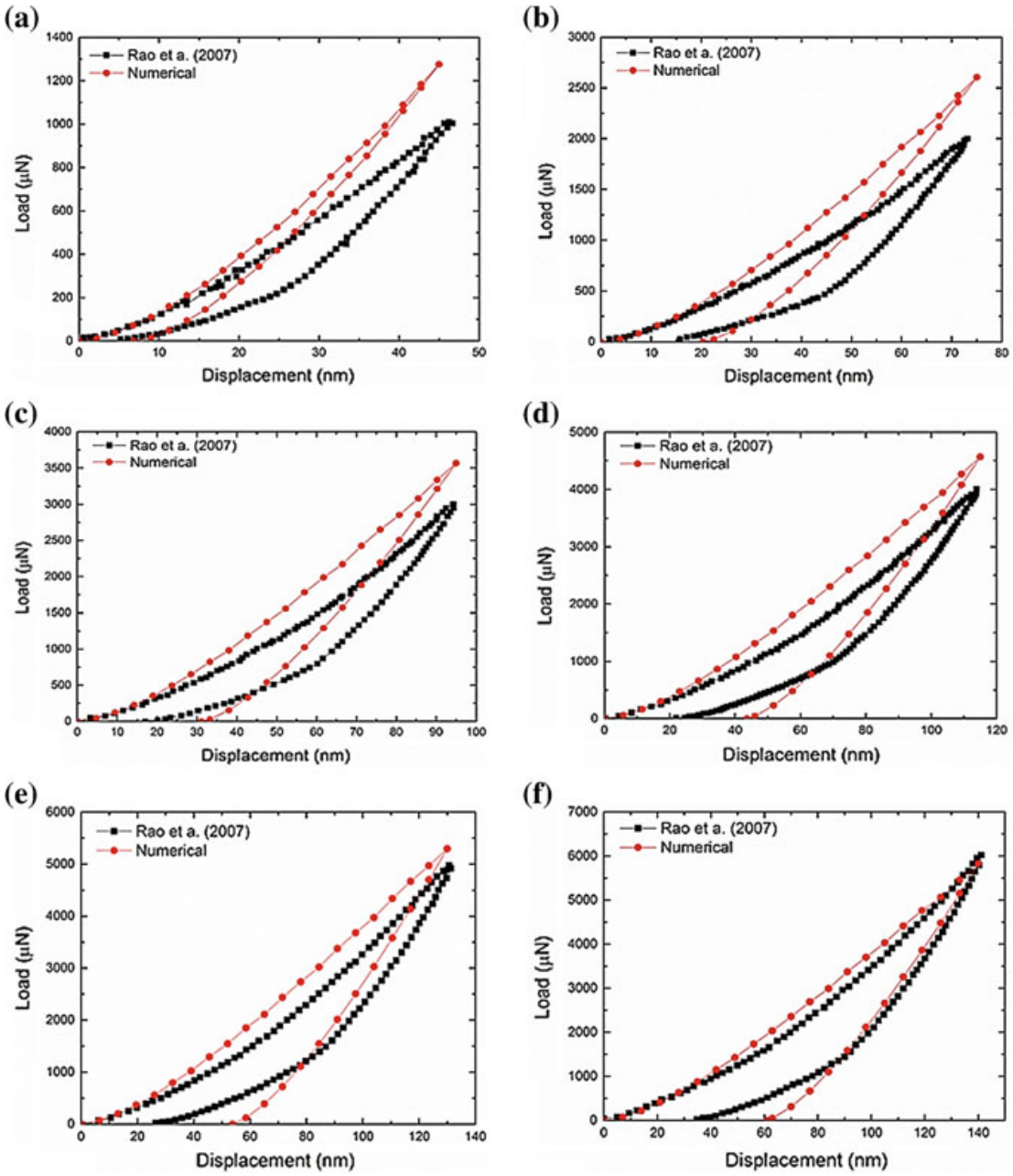


Fig. 72.7 Experimental [11] versus numerical comparison of P-h graph

and 72.4.2, Young's modulus and hardness values were calculated and shown in Table 72.3.

Table 72.3 also depicts the comparison of experimental results with numerical results, such as peak loads, projected area, calculated Young's modulus, and hardness values obtained for various levels of indentation depths. From the table, it is observed that peak load for simulations are higher than the experimental peak load. However, the projected area, Young's modulus, and hardness are found to be similar with minimal prediction error. The simulations are found to be predicting well with an overall prediction error of 10.9%. The projected area is found to be

Table 72.3 Comparison of Young's modulus and hardness at different indentation depth

Peak displacement	Peak load (μN)			Projected area (nm^2)			Young's modulus (GPa)			Hardness (GPa)		
	Experiment	Simulation	Prediction error (%)	Experiment	Simulation	Prediction error (%)	Experiment	Simulation	Prediction error (%)	Experiment	Simulation	Prediction error (%)
45 nm	1004.2	1275.2	27	81,157.7	78,492.9	−3	139	153	10	12	16	31
75 nm	2000.4	2604.4	30	121,056.3	123,750.0	2	176	167	−5	17	21	27
95 nm	3000.0	3566.8	19	149,824.4	150,778.6	1	196	176	−10	20	24	18
115 nm	4001.7	4568.5	14	173,974.9	175,292.9	1	193	185	−4	23	26	13
130 nm	4975.3	5296.1	6.4	192,637.2	192,028.6	−0	211	191	−9	26	28	7
140 nm	6026.0	5826.6	−3	203,447.6	202,400.0	−1	243	196	−19	30	29	−3
Average prediction error			16.6			1.3			9.2			16.5

similar to the experimental value with 1.3% prediction error. The average prediction error of Young's modulus and hardness were found to be 9.2 and 16.5%, respectively with the modified method.

From Fig. 72.6 and Table 72.3, it is observed that the calculated Young's modulus, hardness, and projected area obtained from the experiment and numerically simulated load–displacement graph are in good agreement. Furthermore, it is to be noted that the established values of Young's modulus and hardness values for silicon are 190 GPa and 12 GPa, respectively [14]. The slight variations in the results between the experiments and numerical simulations may be due to the consideration of rigid spherical indenter, isotropic and homogeneous material properties in the numerical model. Moreover, the presence of residual stresses strongly affects Young's modulus and hardness which may cause both to decrease [15–17]. Because, in the nanoindentation experiment, samples are prepared by mechanical means (such as machining), due to which residual stresses are left over the workpiece surface. Also, during the mechanical processing, brittle material undergoes amorphization, which generally has less Young's modulus and hardness as compared to pristine sample [18]. Further, Young's modulus calculation is largely dependent on an empirical factor presented by the Oliver–Pharr. However, despite all these limitations, the FEA fairly estimates the Young's modulus and hardness without conducting physical experiments, which saves a considerable time and cost.

72.6 Conclusions

In the present work, FEA-based simulations of displacement controlled quasi-static nanoindentation were performed on a silicon specimen using a spherical (300 nm tip radius) indenter. The simulations were primarily carried out as a testbed study to evaluate the mechanical properties, viz. Young's modulus and hardness of silicon. The computed values were found in good agreement with the experimental results. Calculated Young's modulus and hardness from the load–displacement plots were 153–196 GPa and 16–29 GPa, while the experimental values were noted to be 139–243 GPa and 12–30 GPa, respectively. The difference between the numerical and published experimental results may be due to (a) residual stresses (tensile or compressive) present in the experimental specimen, (b) experimental process may not follow the continuum assumption due to the presence of voids and dislocations, and (c) diamond indenter may not be that stiff as it is considered in the numerical model. Overall, it is found that finite element simulation provides an economical and quicker alternative to the costly and time-consuming physical experiments in obtaining the mechanical properties of the materials.

References

1. Oliver, W.C., Pharr, G.M.: An improved technique for determining hardness and elastic modulus using load and displacement sensing indentation experiments. J. Mater. Res. **7**(6), 1564–1583 (1992). https://doi.org/10.1557/JMR.1992.1564
2. Oliver, W.C., Pharr, G.M.: Measurement of hardness and elastic modulus by instrumented indentation: advances in understanding and refinements to methodology. J. Mater. Res. **19**(1), 3–20 (2004). https://doi.org/10.1557/jmr.2004.19.1.3
3. Zienkiewicz, O.C., Taylor, R.L.: The finite element method, vol. 36. McGraw-hill, London (1977)
4. Yu, N., Polycarpou, A.A., Conry, T.F.: Tip-radius effect in finite element modeling of sub-50 nm shallow nanoindentation. Thin Solid Films **450**(2), 295–303 (2004). https://doi.org/10.1016/j.tsf.2003.10.033
5. Bhattacharya, A.K., Nix, W.D.: Analysis of elastic and plastic deformation associated with indentation testing of thin films on substrates. Int. J. Solids Struct. **24**(12), 1287–1298 (1988). https://doi.org/10.1016/0020-7683(88),90091-1
6. Lu, C.J., Bogy, D.B.: The effect of tip radius on nano-indentation hardness tests. Int. J. Solids Struct. **32**(12), 1759–1770 (1995). https://doi.org/10.1016/0020-7683(94),00194-2
7. Mariayyah R (2007) Experimental and numerical studies on ductile regime machining of silicon carbide and silicon nitride. Doctoral dissertation, The University of North Carolina at Charlotte
8. Goel, S., Kovalchenko, A., Stukowski, A., Cross, G.: Influence of microstructure on the cutting behaviour of silicon. Acta Mater. **105**, 464–478 (2016). https://doi.org/10.1016/j.actamat.2015.11.046
9. Venkatachalam S (2007) Predictive modeling for ductile machining of brittle materials. Doctoral dissertation, Georgia Institute of Technology
10. Mir, A., Luo, X., Sun, J.: The investigation of influence of tool wear on ductile to brittle transition in single point diamond turning of silicon. Wear **364**, 233–243 (2016). https://doi.org/10.1016/j.wear.2016.08.003
11. Rao, R., Bradby, J.E., Ruffell, S., Williams, J.S.: Nanoindentation-induced phase transformation in crystalline silicon and relaxed amorphous silicon. Microelectron. J. **38**(6–7), 722–726 (2007). https://doi.org/10.1016/j.mejo.2007.04.006
12. Panich, N., Kraivichien, V., Yong, S.: Finite element simulation of nanoindentation of bulk materials. J. Sci. Res. Chulalongkorn Univ. **29**, 145–153 (2004)
13. Shuman, D.J., Costa, A.L., Andrade, M.S.: Calculating the elastic modulus from nanoindentation and microindentation reload curves. Mater. Charact. **58**(4), 380–389 (2007). https://doi.org/10.1016/j.matchar.2006.06.005
14. Reddy JD (2008) Mechanical properties of Silicon Carbide (SiC) thin films. Doctoral dissertation, University of South Florida
15. Tsui TY, Oliver WC, Pharr GM (1996) Influences of stress on the measurement of mechanical properties using nanoindentation: Part I. Experimental studies in an aluminum alloy. J. Mater. Res. **11**(3), 752–759. https://doi.org/10.1557/jmr.1996.0091
16. Kese, K.O., Li, Z.C., Bergman, B.: Influence of residual stress on elastic modulus and hardness of soda-lime glass measured by nanoindentation. J. Mater. Res. **19**(10), 3109–3119 (2004). https://doi.org/10.1557/JMR.2004.0404
17. Lee, W.S., Chang, S.L.: Nanoindentation behaviour and microstructural evolution of annealed single-crystal silicon. World academy of science, engineering and technology. Int. J. Chem. Mol. Nucl. Mater. Metall. Eng. **9**(7), 799–803 (2015). https://doi.org/10.1999/1307-6892/10001683
18. Yan J, Takahashi H, Tamaki JI, Gai X, Harada H, Patten J (2005) Nanoindentation tests on diamond-machined silicon wafers. Appl. Phys. Lett. **86**(18), 181913 (1–3). https://doi.org/10.1063/1.1924895

Chapter 73
Effect of Microchannel Pattern on Fluid Flow and Heat Transfer Characteristics

Dungali Sreehari and Apurbba Kumar Sharma

Abstract Three different patterns (rectangular-serpentine, U-serpentine and V-serpentine) of microchannels were modelled to investigate the fluid flow and heat transfer characteristics using the commercially available finite element analysis software—CFD Fluent. The microchannel models were based on 0.525-mm-thick silicon wafer substrate. The hydraulic diameter (0.4 mm) and channel length (114 mm) were kept constant in all the three different patterned microchannels to compare the performance of the microchannels. Three-dimensional (3D) simulations were performed to simulate near practical condition using the conjugate method in which heat transfer was considered in both solid and fluid zones. Reynolds number was varied in the range of 100–400, and uniform heat flux (30 kW/m^2) was applied at the bottom wall of the microchannels to investigate the effect of microchannel pattern on pressure drop, temperature, velocity profile and heat transfer coefficients. Results confirm that the higher heat transfer coefficient was obtained in U-serpentine microchannel, indicating better performance among the three.

Keywords Microchannel · Channel pattern · Heat transfer · Simulation

73.1 Introduction

With increasing the importance of microchannels in various fields of applications, especially in the electronics industry for cooling of microelectronic components, the research on the design, fabrication and performance evaluation of microchannels has increased. Due to the dense packing of the several microelectronic components in an electronic device, heat dissipation is a major concern. The microchannel is one

D. Sreehari (✉)
National Institute of Technology, Uttarakhand Srinagar (Garhwal) 246174, India
e-mail: sreehari@nituk.ac.in

D. Sreehari · A. K. Sharma
Indian Institute of Technology Roorkee, Roorkee, Haridwar 247667, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_73

of the best solutions to bring down these rising temperatures. Large surface-area-to-volume ratio and low weight and volume attribute the microchannel as an effective heat sink. The concept of the microchannel was introduced by Tuckerman and Pease [1] in the early 1980s. They fabricated microchannels on silicon by an orientation-dependent etching technique and concluded a maximum heat carried out was about 790 W/cm^2 from 1 cm^2 area. After their pioneering study, research on the microchannels application was performed by several researchers across the globe. Peiyi and Little [2] fabricated very fine microchannels on a silicon wafer by the photolithography technique and revealed that the width, depth and surface roughness of the microchannel affect the values of friction factor in the application of microminiature refrigerator. Weisberg et al. [3] numerically analysed the silicon microchannels using conjugate heat transfer method and presented some design formulations for the selection of channel dimensions. Experimental and numerical investigations were carried out by Qu et al. [4] on trapezoidal cross-sectional silicon microchannels. They predicted surface roughness of the microchannel could be the reason to obtain a significant difference between the experimental and numerically determined Nusselt number. Heat transfer characteristics such as temperature, Nusselt number and effect of Reynolds number were numerically analysed in three-dimensional (3D) rectangular silicon microchannels by Qu and Mudawar [5] using conjugate heat transfer method. Lee et al. [6] investigated rectangular microchannels of hydraulic diameter ranging from 318 to 903 μm by experimental and numerical analysis considering both thin wall model and 3D conjugate method. To investigate the geometrical effect of the microchannels on flow and heat transfer characteristics, Gunnasegaran et al. [7] numerically analysed three different cross sections (rectangular, trapezoidal and triangular) and concluded that rectangular cross-sectional microchannels have better heat transfer coefficients.

In general, the microchannel heat sinks are straight channels in which the streamlines of the coolant are nearly straight. Due to poor mixing of fluid in the straight microchannels, the heat transfer is ineffective. Therefore, an alternative wavy microchannel was proposed by Sui et al. [8]. They numerically investigated the performance of straight and wavy microchannels with the same rectangular cross sections using CFD simulations. From their results, it was observed that the performance of wavy microchannel is better than the straight microchannels. Lin et al. [9] further investigated the performance of wavy microchannels varying the wavelength and amplitude of the microchannels in the flow direction. Flow and heat transfer analysis was done in serpentine channels by Karale et al. [10]. They have compared the performance of general serpentine, sinusoidal serpentine and sawtooth serpentine microchannels. Different channel configurations such as nonlinear, wavy, spiral, pin-finned and fractal-shaped were explored to enhance the heat transfer rate [11–14] further.

Although several serpentine microchannels were analysed in previous work, still it is not well understood which serpentine is more appropriate for high heat transfer

with moderate pressure drop. Therefore, in the present work, three different patterns (rectangular-serpentine, U-serpentine and V-serpentine) of microchannels were modelled and analysed to compare the fluid flow and heat transfer characteristics.

73.2 Geometric Models of Microchannels

Silicon wafer (0.525-mm-thick) was selected as substrate material as it is commonly used to fabricate IC chips in the electronic devices. The schematic of three serpentine microchannels is shown in Fig. 73.1. The selection of dimension was based on the fabrication and substrate thickness constraints. Microchannels with a constant length of 114 mm were fabricated by micro-ultrasonic machining technique on the silicon wafer to a depth of 0.3 mm using 0.6-mm-diameter cylindrical tungsten carbide tool [15]. Hence, the same dimensions ($L \times W \times H$: 114 mm 0.3 mm × 0.6 mm) were used to model and analyse the fluid flow and heat transfer characteristics in all the three serpentine microchannels. However, it was observed, to maintain the constant length of the microchannel in all three different serpentines, the base areas of the microchannels vary significantly. The base areas of the rectangular-serpentine, U-serpentine and V-serpentine microchannels are 336, 244.5 and 384 mm^2, respectively. Therefore, a new dimensionless parameter

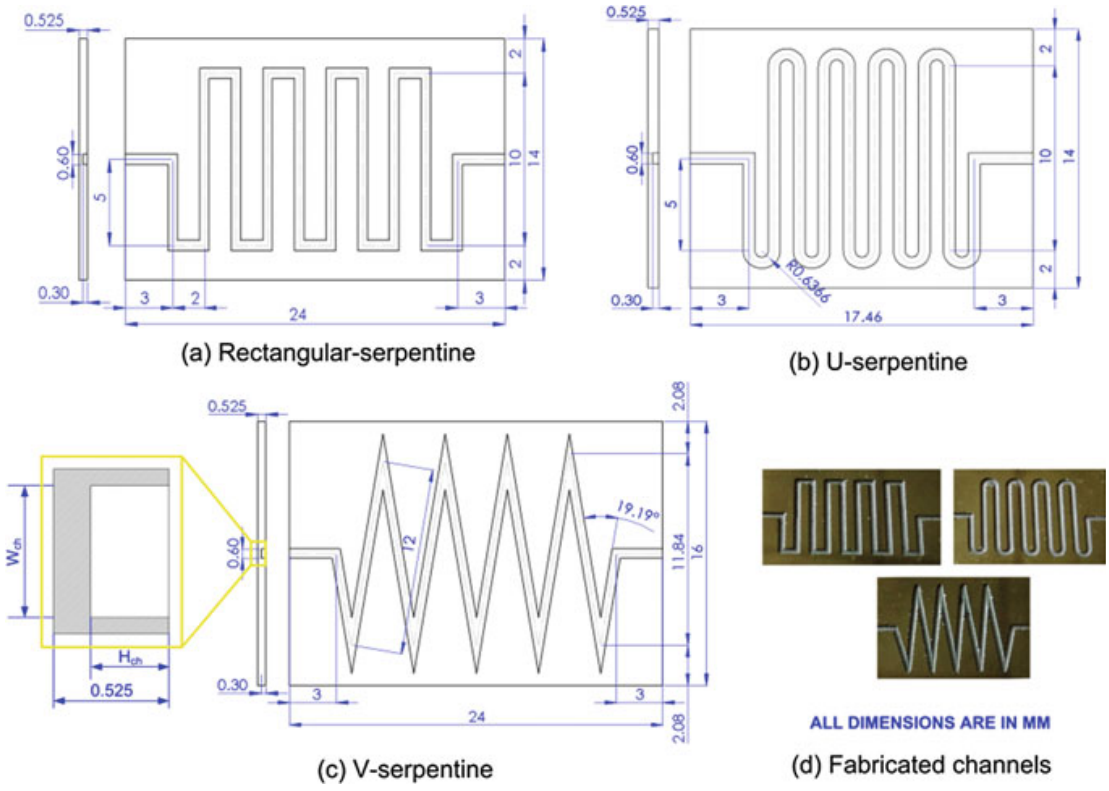


Fig. 73.1 Schematics of **a** rectangular-serpentine, **b** U-serpentine and **c** V-serpentine microchannels and **d** image of fabricated microchannels

Table 73.1 Sink area factors of microchannels

Pattern of microchannel	Total channel wall area (mm^2)	Base area (mm^2)	Sink area factor (%)
Rectangular-serpentine	136.8	336	40.71
U-serpentine	136.8	244.5	55.95
V-serpentine	136.8	384	35.63

—sink area factor (SAF) defined as the ratio of the total channel wall area to the total base area (Eq. 73.1)—was introduced to compare the performance of different serpentine microchannels with constant channel length and different base areas.

$$\text{Sink Area Factor (SAF)}(\%) = \frac{\text{Total channel wall area}}{\text{Base area}} \times 100\% \tag{73.1}$$

Table 73.1 gives the calculated values of sink area factors (SAFs) of the serpentine microchannels. The total channel wall area is the actual convective area responsible for heat transfer, i.e. the bottom wall area ($0.6 \times 114 = 68.4$ mm^2) and the two side wall areas ($2 \times 0.3 \times 114 = 68.4$ mm^2) of the microchannel.

73.3 Numerical Study

The fluid flow in the modelled microchannels was based on the continuity, momentum and energy equations. The flow was assumed to be steady, laminar and incompressible. Uniform heat flux (30 kW/m^2) was applied at the bottom wall of the microchannels. The inlet velocity of the fluid was based on the Reynolds number which was varied in the range of 100–400.

73.3.1 Boundary Conditions

The boundary conditions used in the simulation are as follows:

(i) Inlet fluid temperature was constant ($T_{in} = 300$ K).
(ii) A constant heat flux (30 kW/m^2) was given to the bottom wall, keeping the top and side walls as adiabatic.
(iii) Inlet velocity (V_{in}) was calculated using Eq. (73.2)

$$V_{in} = \frac{Re\mu}{\rho_f D_h} \tag{73.2}$$

$$D_h = \frac{4A}{P} = \frac{2W_{ch}H_{ch}}{W_{ch} + H_{ch}} \tag{73.3}$$

where Re—Reynolds number, μ—viscosity (Kg/m s), ρ_f—density of the fluid (kg/m^3), D_h—hydraulic diameter (μm), W_{ch} and H_{ch} are the width (μm) and depth (μm) of the microchannel.

(iv) Outlet pressure (P_{out}) was assigned a zero gauge pressure.

73.3.2 Grid Independence Test

The grid independence of the solution for rectangular serpentine was performed by considering the computational cells 2849658, 3289705 and 3951850. A variation of 1–2% was observed in the results of temperature and pressure using the computational cells of 3289705 and 3951850; while comparing the results obtained by the computational cells 2849658 and 3289705, a variation of 5–6% was observed. Therefore, 3951850 cells were considered for computation purpose in the numerical simulation study of rectangular-serpentine microchannels. Similar grid independence tests were performed for U-serpentine and V-serpentine microchannels. After obtaining the nearly identical solutions using the different number of cells, 5879476 and 6522096 cells were considered for computation purpose for U-serpentine and V-serpentine microchannels, respectively. It took nearly 60, 75 and 105 min to converge the solution in the rectangular-, U- and V-serpentine microchannels, respectively.

73.3.3 Model Validation

The model was validated with the work of Qu and Mudawar [5] to obtain the accurate results from the present numerical analysis. The boundary conditions, material—silicon substrate and fluid medium (liquid water) were considered to reproduce the results of Qu and Mudawar [5]. In order to compare the result, one of the operating conditions, i.e. wall heat flux of q = 50 W/cm^2 and Re = 140, was simulated. Figure 73.2a, b shows the comparative result of temperature distribution across a vertical plane at the middle of the substrate. It was observed that the present model showed similar results as of Qu and Mudawar [5].

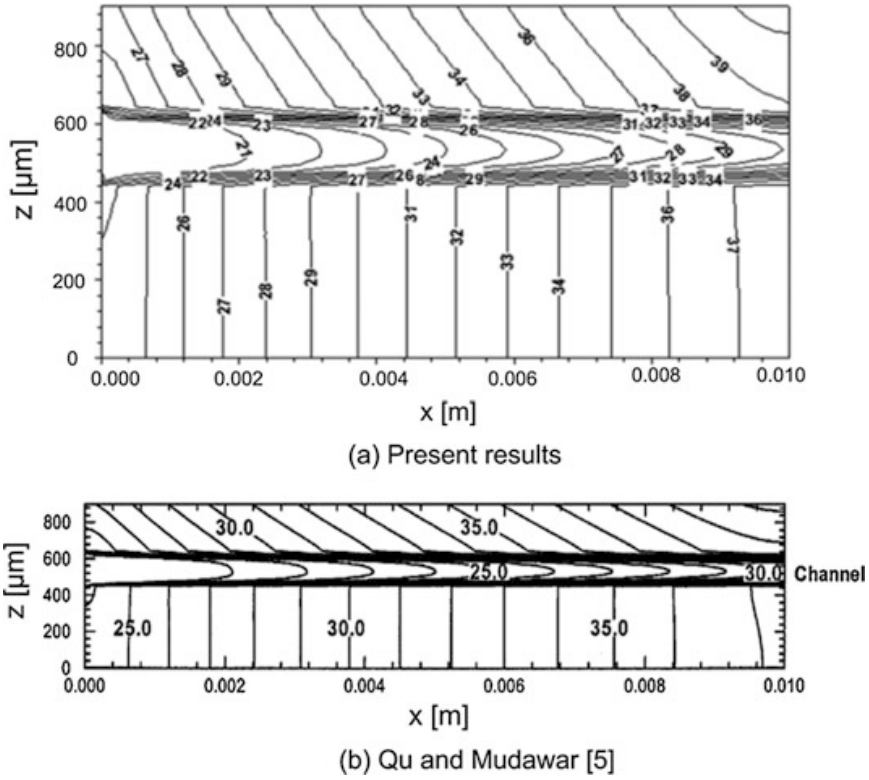


Fig. 73.2 Validation of the present model with Qu and Mudawar [5] results

73.4 Results and Discussion

The numerical simulations were carried out on the three patterns of microchannels considering a constant heat flux (30 kW/m^2) at the bottom wall and varying the Reynolds number from 100 to 400 in steps of 100. In the present section, comparative results in terms of pressure drop, temperature, velocity and heat transfer coefficient are presented.

73.4.1 Pressure Drop

Fig. 73.3 shows the variation of pressure drop (the difference between outlet pressure and inlet pressure) in different serpentine microchannels at different Reynolds numbers (*Re*). It was observed that the pressure drop increases with increase in the Reynolds number, irrespective of the serpentine. The reason is that as the Reynolds number increases, the velocity of the fluid flow increases which in turn increases the pressure drop as it is directly proportional to the square of the velocity. The results are in good agreement with the previous work of Gunnasegaran el al. [7]. Further, it was also observed that pressure drop in all the three serpentines is near similar at *Re*—100 and 200; highest in U—serpentine

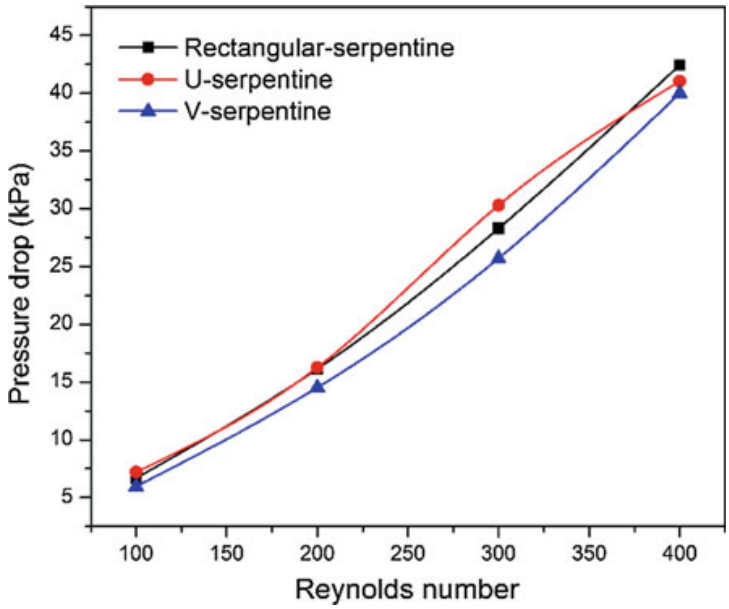


Fig. 73.3 Pressure drop variation with Reynolds number in different pattern microchannels

microchannels at *Re*—300 was observed compared to the other two serpentines. However, at higher Reynolds number (400), the pressure drop variation was marginal.

73.4.2 Temperature

The inlet temperature of the fluid was taken as one of the boundary conditions and was kept constant at 300 K in every simulation of the serpentines. The temperature at the outlet of the fluid was estimated from the simulation results. A difference of the inlet and outlet fluid was calculated in all the serpentine microchannels at different Reynolds numbers. A plot was drawn for the fluid temperature differences at different Reynolds numbers in rectangular-serpentine, U-serpentine and V-serpentine as shown in Fig. 73.4. It was observed that the fluid temperature difference decreases with increasing the Reynolds number, indicating that as the Reynolds number increases, the amount of heat carried away by the fluid was more. However, at higher Reynolds number (beyond *Re*—300), the rate of temperature difference was less. From Fig. 73.4, it was observed the V-serpentine microchannels have a larger fluid temperature difference compared to the remaining microchannels. The reason could be that due to the larger base area of V-serpentine, the amount of heat input was more and hence results in more outlet fluid temperature. Figure 73.5 indicates the contours of the temperature difference in the microchannel and substrate material.

The average base temperature of the substrate was considered from the simulation results, in order to investigate further. The same was plotted against the Reynolds number in Fig. 73.6. It was observed that the base temperature decreases with increase in the Reynolds number. The reason is that due to higher velocities at

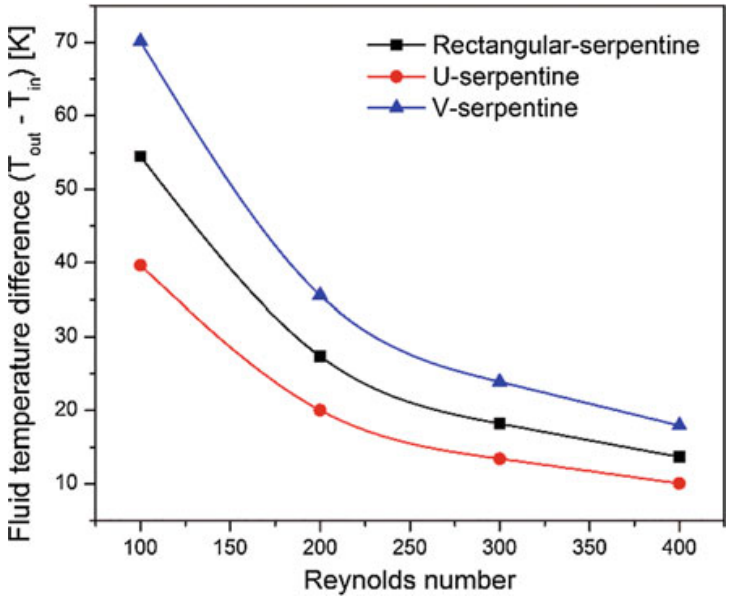


Fig. 73.4 Fluid temperature difference versus Reynolds number

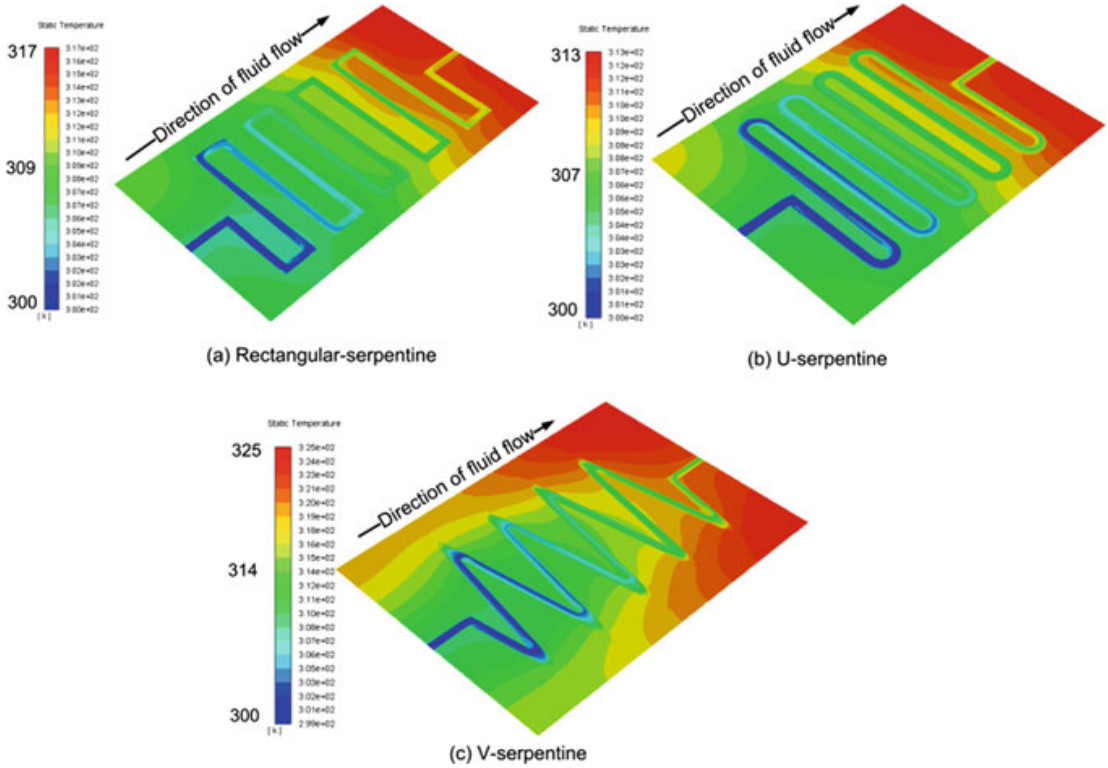


Fig. 73.5 Contours of temperature distribution

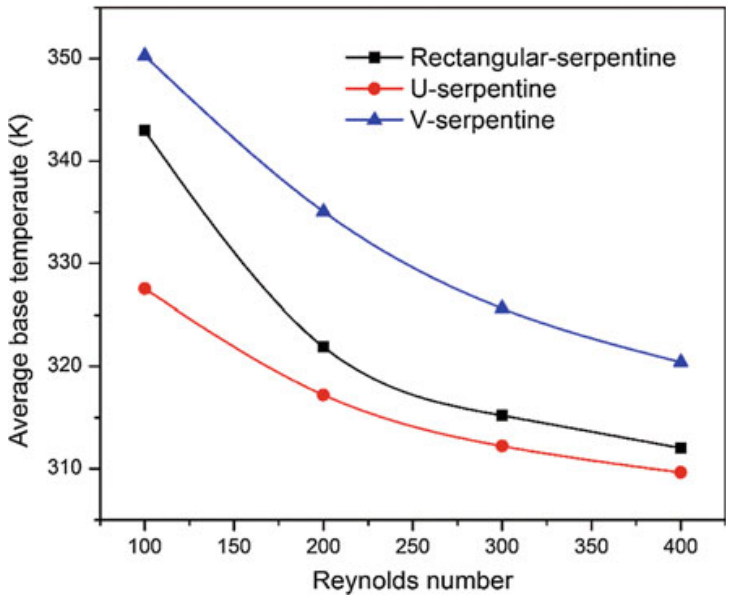


Fig. 73.6 Variation of the base temperature of the substrate with Reynolds number in different serpentine microchannels

higher Reynolds numbers, mixing of fluid was more to carry out more amount of heat from the substrate. However, at any particular Reynolds number, the base temperature of the V-serpentine microchannel was more. This attributes that the cooling effect on the V-serpentine was less compared to the other two serpentine channels. The reason could be that due to the presence of sharp corners in the V-serpentine, the possibility of fluid stagnation was more which results in higher base temperatures.

73.4.3 Velocity Profile

Figure 73.7 indicates the velocity profiles of all the three serpentine microchannel bends. From Fig. 73.7, it is clearly visible that the more amount of fluid was stagnated at the sharp corners of the V-serpentine followed by rectangular and least

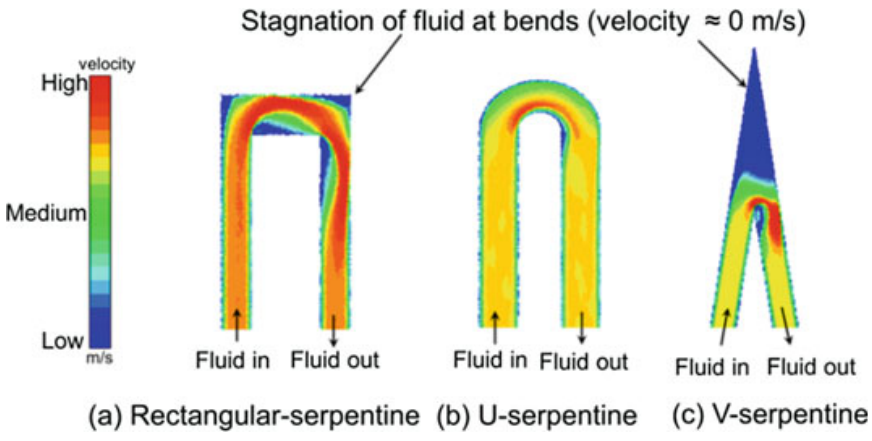


Fig. 73.7 Velocity profiles at bends of serpentine microchannels

in the U-serpentine microchannel. Therefore, the contribution of the fluid in these regions in heat transfer was very low. As the stagnant fluid region was high in V-serpentine, this leads to higher base temperatures. However, the stagnation region of fluid in U-serpentine was very less and hence lowers substrate base temperature (Fig. 73.6).

73.4.4 Heat Transfer Coefficient

The heat transfer coefficient of the fluid passing through the microchannels was calculated using Eq. (73.4).

$$h = \frac{q}{(T_{aw} - T_{af})} \tag{73.4}$$

where

T_{aw} Average bottom wall (base) temperature, K
T_{af} Average fluid temperature = $(T_{in} + T_{out})/2$, K.

In the calculation of heat transfer coefficient, both the substrate bottom wall (base) temperature and average fluid temperature were considered. Therefore, this parameter was used to compare the performance of the serpentine microchannels. Figure 73.8 shows the heat transfer coefficient of different serpentine microchannels at different Reynolds numbers. At higher Reynolds number, the velocity of the fluid is more which results in increased heat transfer coefficients as it maintains a

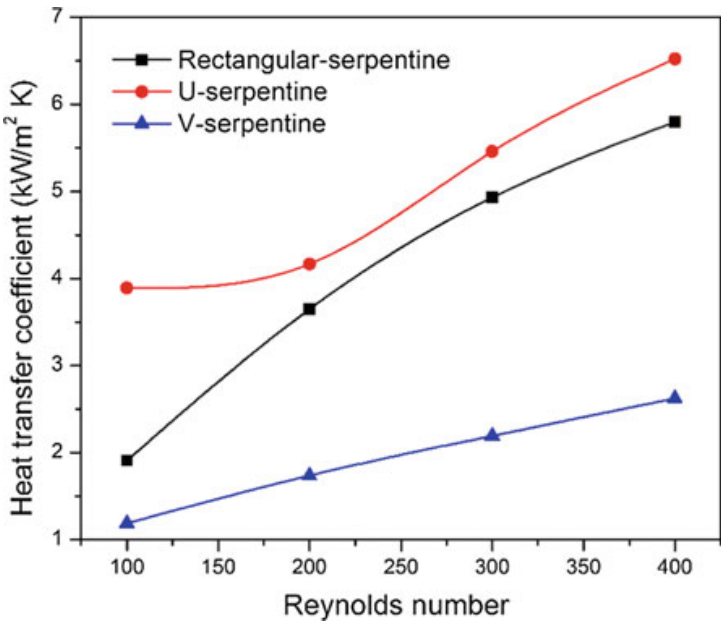


Fig. 73.8 Heat transfer coefficient variation with Reynolds number in different serpentine microchannels

high-temperature difference per unit time. At every *Re,* the value of heat transfer coefficient is more in U-serpentine microchannel, indicating better performing serpentine microchannel.

73.5 Conclusions

The characteristics of fluid flow and heat transfer such as pressure drop, temperature of fluid and substrate, velocity profile and heat transfer coefficient in three different serpentine microchannels at different Reynolds numbers were investigated numerically, and the results were discussed. The major conclusions drawn are as follows:

(a) Pressure drop was nearly similar in every microchannel. At *Re*—300, U-serpentine microchannel exhibits higher pressure drop compared to the remaining two microchannels.
(b) V-serpentine microchannel has the highest values of fluid temperature difference and average substrate temperatures.
(c) A larger amount of stagnation of fluid (velocity $\approx$ 0 m/s) was observed at the sharp bends of V-serpentine microchannels from the velocity profiles. However, smooth passage of the fluid was observed in U-serpentine microchannels.
(d) Among the three serpentine microchannels, the overall heat transfer coefficient in the U-serpentine microchannel has shown better performance compared to the other two.

References

1. Tuckerman, D.B., Pease, R.F.W.: High-performance heat sinking for VLSI. IEEE Electron Device Lett. **2**(5), 126–129 (1981)
2. Peiyi, W., Little, W.A.: Measurement of friction factors for the flow of gases in very fine channels used for microminiature Joule-Thomson refrigerators. Cryogenics **23**(5), 273–277 (1983)
3. Weisberg, A., Bau, H.H., Zemel, J.N.: Analysis of microchannels for integrated cooling. Int. J. Heat Mass Transf. **35**(10), 2465–2474 (1992)
4. Qu, W., Mala, G.M., Li, D.: Heat transfer for water flow in trapezoidal silicon microchannels. Int. J. Heat Mass Transf. **43**(21), 3925–3936 (2000)
5. Qu, W., Mudawar, I.: Analysis of three-dimensional heat transfer in micro-channel heat sinks. Int. J. Heat Mass Transf. **45**(19), 3973–3985 (2002)
6. Lee, P.S., Garimella, S.V., Liu, D.: Investigation of heat transfer in rectangular microchannels. Int. J. Heat Mass Transf. **48**(9), 1688–1704 (2005)
7. Gunnasegaran, P., Mohammed, H.A., Shuaib, N.H., Saidur, R.: The effect of geometrical parameters on heat transfer characteristics of microchannels heat sink with different shapes. Int. Commun. Heat Mass Transf. **37**(8), 1078–1086 (2010)

8. Sui, Y., Teo, C.J., Lee, P.S., Chew, Y.T., Shu, C.: Fluid flow and heat transfer in wavy microchannels. Int. J. Heat Mass Transf. **53**(13–14), 2760–2772 (2010)
9. Lin, L., Zhao, J., Lu, G., Wang, X.D., Yan, W.M.: Heat transfer enhancement in microchannel heat sink by wavy channel with changing wavelength/amplitude. Int. J. Therm. Sci. **118**, 423–434 (2017)
10. Karale, C.M., Bhagwat, S.S., Ranade, V.V.: Flow and heat transfer in serpentine channels. AIChE J. **59**(5), 1814–1827 (2013)
11. Zhang, J., Zhang, Y.F., Miao, M., Jin, Y.F., Bai, S.L., Chen, J.Q.: Simulation of fluid flow and heat transfer in microchannel cooling for LTCC electronic packages. In: 2009 International Conference on Electronic Packaging Technology and High Density Packaging, pp. 327–330 (2009)
12. Dehghandokht, M., Khan, M.G., Fartaj, A., Sanaye, S.: Numerical study of fluid flow and heat transfer in a multi-port serpentine meso-channel heat exchanger. Appl. Therm. Eng. **31**(10), 1588–1599 (2011)
13. Gong, L., Kota, K., Tao, W., Joshi, Y.: Parametric numerical study of flow and heat transfer in microchannels with wavy walls. J. Heat Transfer **133**(5), 051702 (2011)
14. Kandlikar, S.G.: History, advances, and challenges in liquid flow and flow boiling heat transfer in microchannels: a critical review. J. Heat Transfer **134**(3), 034001 (2012)
15. Sreehari, D., Sharma, A.K.: On form accuracy and surface roughness in micro-ultrasonic machining of silicon microchannels. Precis. Eng. **53**, 300–309 (2018)

Chapter 74
Comparative Analysis of GPU-Based Solver Libraries for a Sparse Linear System of Equations

Utpal Kiran, Subhajit Sanfui, Shashi Kant Ratnakar, Sachin Singh Gautam and Deepak Sharma

Abstract In this paper, a comparison of GPU-based linear solver libraries for the solution of sparse positive-definite matrices is presented. These large sparse matrices arise in a number of computational disciplines seeking a solution for partial differential equations. The solution of these matrices is often a time-consuming process that can be reduced by parallel computing. Since the development of GPU for general-purpose computing, a number of numerical solver libraries have evolved that can accelerate the solution procedure. The performance of three solver libraries has been evaluated in this paper for five different test matrices. These test matrices have been taken from different application domains with different sparsity patterns. Results demonstrate a higher speedup from the iterative solver over the direct solver on GPU and also over a multithreaded CPU implementation.

Keywords Finite element analysis · GPGPU · Sparse matrix · Linear solver

U. Kiran · S. Sanfui · S. K. Ratnakar · S. S. Gautam · D. Sharma (✉)
Department of Mechanical Engineering, Indian Institute of Technology Guwahati, Guwahati, Assam 781039, India
e-mail: dsharma@iitg.ac.in

U. Kiran
e-mail: ukiran@iitg.ac.in

S. Sanfui
e-mail: s.sanfui@iitg.ac.in

S. K. Ratnakar
e-mail: r.shashi@iitg.ac.in

S. S. Gautam
e-mail: ssg@iitg.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_74

74.1 Introduction

Most of the physical phenomena are governed by partial differential equations (PDEs). In various problems of engineering and science, it is not always possible to analytically solve the governing PDE to get the required solution. Most of the time simplifying approximations are made to reduce the PDE to a set of algebraic equations, using numerical methods [1].

Numerical simulations of manufacturing processes also use these numerical approximation techniques. Finite element analysis (FEA) is one of the most popular and widely used numerical methods. FEA facilitates construction, refinement, and optimization of a mechanical component before the start of manufacturing. This substantially reduces the cost and effort involved in production. The advantage of numerical simulation such as FEA is that it makes virtual testing possible in the early stages of production. Furthermore, the widespread availability of modern FEA software makes it very convenient to analyze the performance of the product and make the necessary improvements in the design.

The FEA formulation of the problem gives a system of algebraic equations. Solving this system of equations provides approximate values of the primary unknown, at a number of discrete points over the problem domain [2]. The resulting system of equations from FEA can be very conveniently represented in the form of matrices. Since FEA subdivides a large problem into smaller parts, due to local correlation in PDE discretization, the final matrix is sparse in nature [3].

There are two methods to solve the system of equations, the direct method and iterative method [1]. Direct methods are the procedures based on algebraic elimination. Gauss elimination and LU factorization are some of the most known algorithms in this category. Iterative methods start by assuming a trial solution or initial guess which is substituted in the system of equations, and then the difference between the obtained solution and the actual solution is determined, also known as an error or residual. Using the information from error value, the initial guess is improved iteration by iteration, so that it reaches closer to the actual solution. The conjugate gradient (CG) method is one of the most widely used iterative solvers [4].

In all practical FEA problems usually, the number of variables involved is very large, thus making the size of the sparse matrix huge. Solving a system of equations consisting of millions of variables becomes a very time-consuming exercise. In various studies, it has been observed that the solver may take anything between 70 and 90% of the total computation effort required (depending upon the problem size) [5]. Due to the facts discussed earlier, sparse matrix solver is an active research area.

74.1.1 Parallel Computing

One of the remedies to overcome a higher computational time of the solver is to use parallel computing, where a large problem is subdivided into small tasks and

executed by multiple computing units (processors) simultaneously, with or without communicating with each other [6].

There are different types of parallel computing architecture. The choice of architecture is subjected to many factors, such as the requirement of the problem, availability of hardware, and cost. One of the most popular architectures is the distributed memory architecture, where each processor has its own memory system at its disposal, which it can use directly. Another is the shared memory architecture, where each processor has access to a centralized memory system from which they have to share. Usage of memory by each processor depends on its current requirement. These two architectures are generally referred to as traditional architectures.

A distributed memory architecture setup generally consists of a large number of CPU cores. This results in high setup and maintenance cost, making it somewhat impractical for common use. Shared memory architecture does not need any separate communication channel. All the CPU cores in shared memory architecture access a common global memory. Here, the changes made in memory by one processor are visible to all other processors. This may lead to inaccurate memory operation when more than one processor is trying to access the same memory location. This issue is commonly known as race condition [7].

In the literature, the parallelization of solver has been performed using traditional computing such as the shared memory architecture [5, 8, 9] and also distributed memory architecture [10–13]. Considering the inconvenience posed by the traditional architectures, in recent years, the novel architecture of graphics processing unit (GPU) has gained popularity among the researchers. After the launch of general-purpose GPU (GPGPU) by NVIDIA in 2007, the GPUs can be used for general-purpose computing. GPU is advantageous in the sense that they are relatively cheaper, maintenance-free, and portable. They can be used for parallel computing using the programming interface like Compute Unified Device Architecture (CUDA) [14] or OpenCL [15].

GPU has a highly parallel, multithreaded, many-core processor architecture, which is well suited for problems that can be expressed as data-parallel computations with high arithmetic intensity. Hence, GPU can be a viable choice for a sparse matrix solver. It has been used for the parallelization of solver in a few research works [16–19]. Moreover, some libraries compatible with GPU are also developed for a sparse matrix solver. However, these libraries have not been compared on a common platform on the problems of different domains. Therefore, the main contributions of this paper are (i) to compare the popular linear sparse solver libraries that are developed for GPU, with respect to the multithreaded CPU implementation, and (ii) to analyze the performance of the popular libraries by solving some sparse matrices evolved from different applications that have different matrix sizes, sparsity pattern, and nonzero entries.

The paper is organized into four sections. Section 74.2 presents the existing sparse matrix solver libraries developed for GPUs. In Sect. 74.3, the matrices evolved from different problems are discussed and the libraries are used to find the solution for a system of the linear system. The computation time is reported, and the speedup is shown. Section 74.4 concludes the paper.

74.2 Sparse Matrix Solver Libraries

Given the importance of the large sparse linear equation system solver, there are several libraries available in the literature, which takes the sparse matrix as an input and provides the solution as an output. The majority of works in the literature focus on CPU implementation, while a few of them have been developed to work on GPU hardware so that the solution can be obtained in a lesser amount of time. Many of these solver libraries can work with a user-defined matrix-vector product, which is the basic operation of any iterative solver.

Solver libraries provide highly optimized functions that perform many times faster than CPU-only solvers. They can be very useful for compute-intensive applications such as structural optimization, computational mechanics, heat flow problems, earthquake simulations to name a few. In the following subsections, a few of GPU sparse linear equation system solver libraries are discussed.

74.2.1 CUSP

CUSP is an open-source library written in C++ for sparse linear algebra on GPUs with CUDA architecture. It provides a flexible programming interface for operating on sparse matrices and solving sparse linear systems. CUSP uses iterative methods and also includes multiple pre-conditioners. This library supports many popular sparse storage formats such as COO, CSR, DIA, ELL, and HYB. The library is developed by NVIDIA [20].

74.2.2 cuSolver

Also developed by NVIDIA, the cuSolver library provides a collection of direct solvers for both dense and sparse matrices. This library comes in-built with the CUDA toolkit. For dense matrices, it has common matrix factorization and

triangular solve routines, and for sparse matrices, a least square solver is used. This library also comes with an eigenvalue solver. It can be called from a C/C++ routine [21].

74.2.3 SuiteSparse

SuiteSparse is an open-source library, which has a Cholesky factorization accelerated by CUDA. It has a collection of a number of very popular sparse matrix packages such as UMFPACK, CHOLMOD, and SPQR [22].

74.2.4 PETSc

Portable, Extensible Toolkit for Scientific Computation (PETSc) is an open-source suite of data structures and routines for the parallel scalable solution of applications which are modeled by PDEs [23]. It has the capability of solving both linear and nonlinear systems of equations. It is available in C/C++ and Python. It uses MPI for message passing and supports GPU through CUDA or OpenCL. PETSc is the world's most widely used parallel numerical software library for partial differential equations and sparse matrix computations.

74.2.5 MAGMA

MAGMA is a collection of GPU-accelerated dense linear algebra libraries. It has been specifically developed to work with heterogeneous computing environment such as multicore CPUs and multi-GPUs [24].

74.3 Results and Discussion

In the present work, three libraries have been selected for solving the matrices using the direct and iterative methods on the GPU. The speedups of the GPU-based solvers are obtained with respect to a multithreaded implementation of Cholesky decomposition (CHOLMOD) provided by SuiteSparse library. The details of the three test libraries are provided in Table 74.1. The test matrices have been chosen

Table 74.1 Methods to solve a linear system of equations

	Library	Solution method	Architecture
Method 1	CUSP	Conjugate gradient (CG)	GPU
Method 2	cuSolver	Cholesky decomposition	GPU
Method 3	SuiteSparse	Cholesky decomposition	GPU
Ref. code	SuiteSparse	Cholesky decomposition	CPU multithreaded

from different applications of FEM and solved using libraries mentioned in Table 74.1. Five matrices with a different range of order and number of nonzeros have been selected from the Suitesparse matrix collection [3] for benchmarking. The details of these matrices are presented in Table 74.2. All the five test matrices are sparse and positive definite. The sparsity pattern of all test matrices is shown in Fig. 74.1. As can be seen from the figure, the sparsity pattern as well as the density of nonzeros can greatly vary for different applications. Since the focus of this paper is to analyze the performance of different solver libraries on matrices having different sparsity patterns, the problem formulation or mathematical modeling for the applications shown in Fig. 74.1 is not presented. However, the details of these applications can be found in [3], which is not included due to brevity. In the present work, the performance of solver libraries on matrices with significantly varying sparsity pattern is evaluated.

The hardware for the numerical experiment consists of Intel Xeon E5-2650 processor and NVIDIA Tesla K40 GPU. The CPU processor has 128 GB of memory and 12 physical cores clocked at 2.2 GHz. The GPU has 12 GB of memory, and the cores are clocked at 745 MHz.

The execution time for the matrices for different libraries is listed in Table 74.3. Execution time refers to elapsed GPU time for all three methods and CPU time for the reference code. The last column of Table 74.3 lists the time required by the reference multithreaded version. The speedups of the three solvers are computed with reference to this version and plotted in Fig. 74.2, using a semi-logarithmic scale. From the figure, it can be clearly seen that among all the libraries, the CG method in CUSP library takes the least amount of time for solving all the matrices. It can also be observed from the figure that the speedup for 'parabolic_fem' and 'thermal2' is significantly more pronounced than the other matrices. This can be

Table 74.2 Details of test matrices

Matrix	Application	Rows/columns	Nonzeros
thermal1	Steady-state thermal problem	82,654	574,458
thermal2	Steady-state thermal problem	1,228,045	8,580,313
nd6k	ND problem set	18,000	6,897,316
parabolic_fem	Diffusion–convection reaction	525,825	3,674,625
consph	FEM concentric spheres	83,334	6,010,480

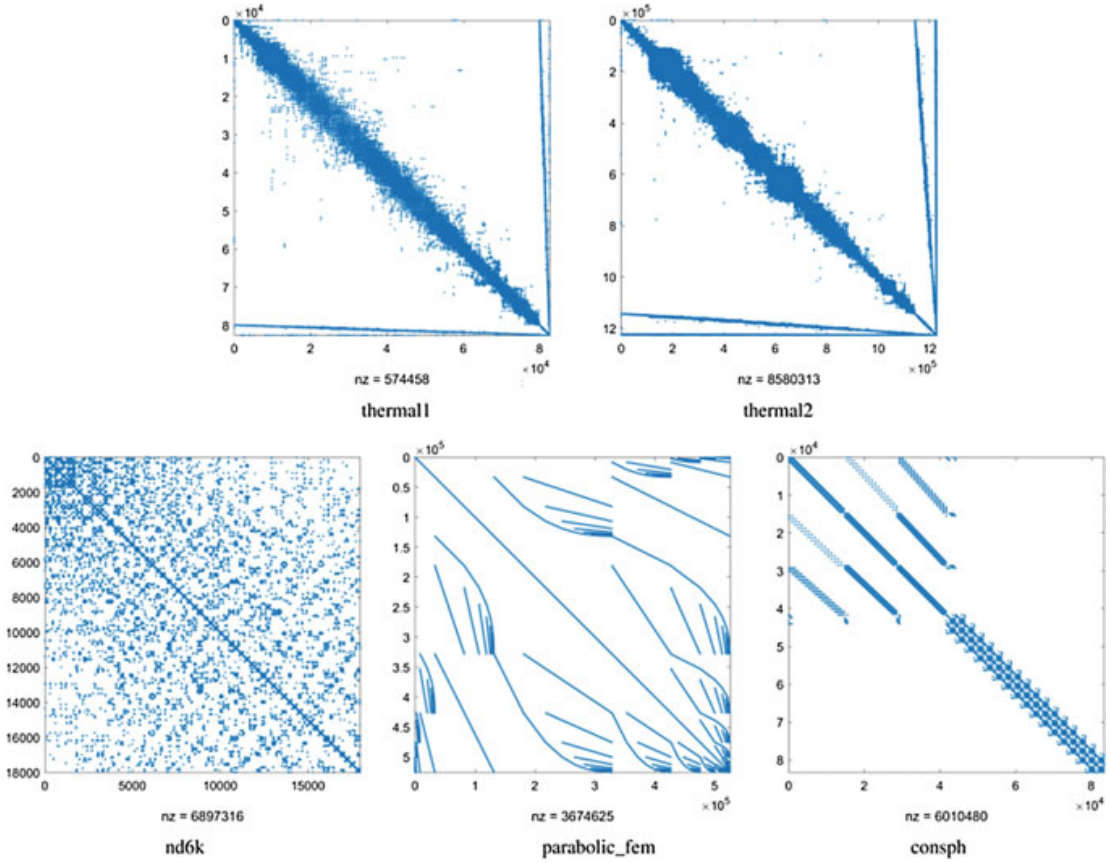


Fig. 74.1 Sparsity pattern of test matrices is represented by dots

Table 74.3 Execution time (in seconds)

Matrix	Method 1	Method 2	Method 3	Ref. code
thermal1	1.92	46.32	17.94	14.22
thermal2	18.11	–	195.92	262.74
nd6k	17.58	69.09	23.85	19.95
parabolic_fem	3.23	–	114.25	158.46
consph	27.84	–	153.85	177.08

attributed to the higher dimension of the said matrices. In the case of SuiteSparse library, the GPU implementation of CHOLMOD is seen to outperform the reference multithreaded CPU implementation for all matrices except 'thermal1' and 'nd6k'. This is due to the small size of the two matrices. It can be further concluded from the figures that the execution times of all the libraries are more dependent on the dimension of the matrix than the number of nonzeros in the matrices. For example, 'nd6k', despite having a high number of nonzeros as compared to the other matrices, suffers from a poor performance in all the methods as seen in Fig. 74.2.

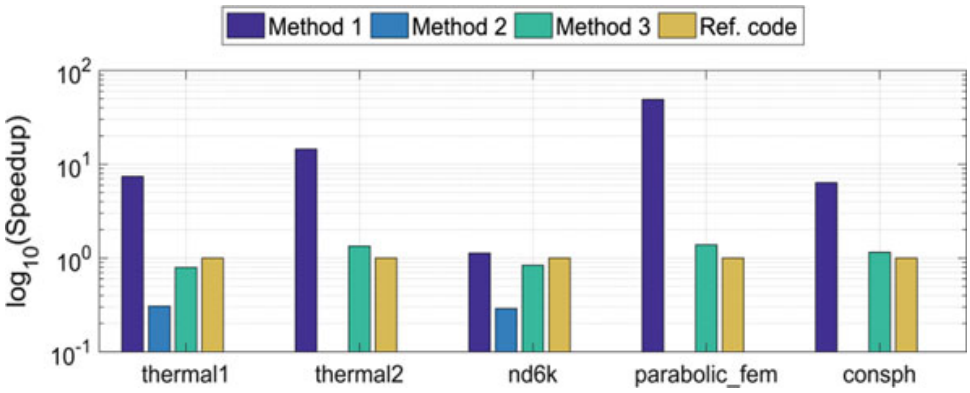


Fig. 74.2 Speedup of GPU solver libraries with respect to multithreaded ref. code

74.4 Conclusion

This paper presents a study of a solution of a sparse system of linear equations using a GPU. The performance of solver libraries, which use a GPU to accelerate the computation, is evaluated over a range of matrices. Five test matrices of varying sparsity pattern as well as a different number of nonzeros and dimension are taken. The matrices are solved by three popular libraries over GPU, and result is compared with a multithreaded CPU implementation. The GPU-based CUSP library achieves the least solution time for all the matrices. The other libraries that use GPU to accelerate computation also achieve lesser time than their CPU variants. It was observed that the effect of GPU acceleration becomes more apparent when matrix size increases. The highest speedup is achieved for the matrix of the largest dimension. The number of nonzeros is also found to be affecting the solution time.

References

1. Hoffman, J.D., Frankel, S.: Numerical methods for engineers and scientists. CRC Press (2001)
2. Zienkiewicz, O.C., Taylor, R.L.: The finite element method, vol. 36. McGraw-hill, London (1977)
3. Davis, T.A., Hu, Y.: The university of Florida sparse matrix collection. ACM Trans. Math. Softw. (TOMS) **38**(1), 1 (2011)
4. Shewchuk, J.R.: An Introduction to the Conjugate Gradient Method Without the Agonizing Pain. Technical Report. Carnegie Mellon University, Pittsburgh, PA, USA (1994)
5. Mahdavi, A., Balaji, R., Frecker, M., Mockensturm, E.M.: Topology optimization of 2D continua for minimum compliance using parallel computing. Struct. Multi. Optim. **32**(2), 121–132 (2006)
6. Pacheco, P.: Parallel Programming with MPI. Morgan Kaufmann (1997)
7. Farhat, C., Crivelli, L.: A general approach to nonlinear FE computations on shared-memory multiprocessors. Comput. Methods Appl. Mech. Eng. **72**(2), 153–171 (1989)
8. Farhat, C.: A simple and efficient automatic FEM domain decomposer. Comput. Struct. **28**(5), 579–602 (1988)

9. Babuška, I., Elman, H.C., Markley, K.: Parallel implementation of the hp-version of the finite element method on a shared-memory architecture. SIAM J. Sci. Stat. Comput. **13**(6), 1433–1459 (1992)
10. Borrvall, T., Petersson, J.: Large-scale topology optimization in 3D using parallel computing. Comput. Methods Appl. Mech. Eng. **190**(46–47), 6201–6229 (2001)
11. Masters, I., Usmani, A.S., Cross, J.T., Lewis, R.W.: Finite element analysis of solidification using object-oriented and parallel techniques. Int. J. Numer. Meth. Eng. **40**(15), 2891–2909 (1997)
12. Coelho, P.G., Cardoso, J.B., Fernandes, P.R., Rodrigues, H.C.: Parallel computing techniques applied to the simultaneous design of structure and material. Adv. Eng. Softw. **42**(5), 219–227 (2011)
13. Tezduyar, T., Aliabadi, S., Behr, M., Johnson, A., Mittal, S.: Parallel finite-element computation of 3D flows. Computer **26**(10), 27–36 (1993)
14. Sanders, J., Kandrot, E.: CUDA by Example: An Introduction to General-Purpose GPU Programming, 1st edn. Addison-Wesley Professional (2010)
15. Munshi, A., Gaster, B., Mattson, T.G., Ginsburg, D.: OpenCL Programming Guide. Pearson Education (2011)
16. Schmidt, S., Schulz, V.: A 2589-line topology optimization code written for the graphics card. Comput. Vis. Sci. **14**(6), 249–256 (2011)
17. Challis, V.J., Roberts, A.P., Grotowski, J.F.: High resolution topology optimization using graphics processing units (GPUs). Struct. Multi. Optim. **49**(2), 315–325 (2014)
18. Taylor, Z.A., Cheng, M., Ourselin, S.: High-speed nonlinear finite element analysis for surgical simulation using graphics processing units. IEEE Trans. Med. Imaging **27**(5), 650–663 (2008)
19. Goddeke, D., Buijssen, S.H., Wobker, H., Turek, S.: GPU acceleration of an unmodified parallel finite element Navier-Stokes solver. In: International Conference on High Performance Computing & Simulation, 2009. HPCS'09, pp. 12–21. IEEE (2009)
20. Bell, N., Garland, M.: Cusp: Generic parallel algorithms for sparse matrix and graph computations version 0.3. 0. Cusp-Library (2012)
21. NVIDIA Corporation, cuSOLVER Library User's Guide, https://docs.nvidia.com/cuda/pdf/CUSOLVER_Library.pdf
22. Davis, T.: Direct Methods for Sparse Linear Systems. SIAM Series on the Fundamentals of Algorithms, SIAM, Philadelphia (2006)
23. Balay, S., Abhyankar, S., Adams, M., Brown, J., Brune, P., Buschelman, K., Dalcin, L.D., Eijkhout, V., Gropp, W., Kaushik, D., Knepley, M.: Petsc user's manual revision 3.8 (No. ANL-95/11 Rev 3.8). Argonne National Lab. (ANL), Argonne, IL, United States (2017)
24. Tomov, S., Dongarra, J., Volkov, V., Demmel, J.: Magma Library. University of Tennessee and University of California, Knoxville, TN, and Berkeley, CA (2009)

Chapter 75
Hybridizing Evolutionary Multi-objective Algorithm Using Random Mutations and Local Searches

Riddhiman Saikia and Deepak Sharma

Abstract Multi-Objective Evolutionary Algorithms (MOEAs) have been successful in solving mathematical and real-world multi-objective optimization problems by evolving a set of optimal solutions, which are known as Pareto-optimal solutions. However, there are certain limitations with those algorithms such as slow convergence, lack of effective terminating condition to name a few. To address such challenges, hybrid MOEAs are being designed and studied where the global exploration power of MOEAs are combined with local exploitation modules of various numerical optimization techniques. However, hybridization itself brings new challenges in its implementation. In this work, a hybrid MOEA is presented in which random mutations are performed on the initial population to start with a better and diverse set of solutions. Moreover, a local search module is coupled to execute periodically on a least crowded non-dominated solution at a certain interval of generations. The proposed algorithm is tested on a set of benchmark multi-objective optimization problems and compared with the NSGA-II. The convergence plots demonstrate the superiority of the proposed algorithm over NSGA-II.

Keywords Hybrid multi-objective algorithm · Memetic algorithm · Mutation · Local search · NSGA-II

R. Saikia · D. Sharma (✉)
Department of Mechanical Engineering, Indian Institute of Technology Guwahati, Guwahati, Assam 781039, India
e-mail: dsharma@iitg.ac.in

R. Saikia
e-mail: saiki174103012@iitg.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_75

75.1 Introduction

In Operations Research, multi-objective optimization is a sub-field of multi-criterion decision making in which optimal solutions are generated by simultaneous optimization of two or more conflicting objectives. Mathematically, the formulation of a multi-objective optimization problem is

$$\begin{aligned} &\text{Min } F(x) = \{f_1(x), f_2(x), \ldots, f_M(x)\} \\ &\text{subject to } x \in S \subset R^n \end{aligned} \tag{75.1}$$

Here, M is the number of conflicting objectives. Any optimal solution x^* is considered as Pareto-optimal, if there does not exist another $x \subset S$ that dominates x*. Meaning, a solution x is said to dominate x^* if it is not worse than x^* in any objective, and better in at least one [1]. Symbolically, it can be represented as $x \prec x^*$ (x dominates x^*), if $f_i(x) \leq f_i(x^*)$ for all $i = 1, \ldots, M$ and $f_j(x) < f_j(x^*)$ for at least one objective $j \in (1, \ldots, M)$ for the problem given in Eq. (75.1).

No solution to any multi-objective optimization problem can be considered as best since trade-offs between conflicting objectives are inevitable. Hence it is necessary to generate a set of solutions. The optimization problem in Eq. (75.1) can be solved using any numerical optimization technique by converting it into a single-objective composite function or converting all but one objective into a set of constraints. However, such approaches fail many a time to generate a well-distributed Pareto-optimal front. Evolutionary algorithms (EAs), however, have been an ideal candidate for solving such optimization problems because they work on a population of solutions. With appropriate operators for convergence and diversity, these EAs can well evolve a set of Pareto-optimal solutions.

In the literature, many multi-objective evolutionary algorithms (MOEAs) have been suggested, such as NSGA-II [2], SPEA-II [3], DoD [4], MOPSO [5], ASREA [6] to name a few. Despite excellent performances, MOEAs are being criticized for their slow convergence toward the Pareto-optimal front, lack of proof of convergence, lack of proper termination condition etc. To address some of these issues, many authors added various modifications to the MOEAs, commonly being modules for local search (LS), which is exploitation of an already-found solution in an attempt to improve its fitness. It is performed on one or many solutions at a time which can lead to a faster convergence if applied correctly [7].

While LS has many advantages, it is not a straightforward task to synchronize it with MOEAs. The first challenge with LS is that it is computationally exhaustive, since extra function evaluations consume extra resources. Too much of LS demands more computational resources and too less can lead to a sub-optimal solution [8]. Therefore, the challenges are when to execute LS and how much to execute? Another challenge is about a selection of solutions on which LS is applied. Since MOEAs are population-based optimization algorithms, the basis of choosing solution(s) for LS poses additional challenges [9].

There have been numerous attempts at answering such challenges. A remedy to overcome the difficulty of choosing weights for weighted scalar or composite function approach was proposed in [10], where the authors used three populations for generation update: the current population, the one after genetic operations and also after LS. If a solution after LS is degraded, then a better choice from other two populations is likely to be chosen over it for the next generation. Secondly, the authors used an acceptance rule for LS. Ishibuchi et al. [11] addressed the issue on how to choose solutions for LS. The authors used tournament selection based on fitness for choosing the best solutions of a population and performed LS on them. The authors noted improved performances and concluded that choosing better individuals for LS had a significantly positive impact. On the other hand, in some studies [12, 13], all non-dominated solutions have been considered for LS. Sindhya et al. [14] made an attempt to obtain good a balance between local and global search. The authors developed a concurrent hybrid MOEA where a probability based LS model was proposed. It has been argued that initial local searches can be wasted as GA operators would change the schema of solutions. But in later stages, it is desired that the solutions converge quickly towards the Pareto-optimal front. In this paper, an attempt is made in hybridizing NSGA-II with random mutations on the initial population and also coupled with local search during the generations. The initial mutations can help to further diversify the initial population and the local searches can help for faster convergence.

The rest of the paper is organized into four sections. Section 75.2 discusses the framework of hybrid MOEA. In Sect. 75.3, the performance of the proposed algorithm is presented after solving some benchmark multi-objective optimization problems. Section 75.4 concludes this paper with plans for future research.

75.2 Framework for Hybrid MOEA (HMOEA)

The framework for hybrid MOEA is shown in Fig. 75.1. The non-dominated sorting genetic algorithm, which is known as NSGA-II, has been chosen for hybridizing with the proposed ideas of random mutations and local searches. The hybrid NSGA-II is referred to as HMOEA in this paper. As can be seen from Fig. 75.1, an initial random population of size N is generated at the beginning. Since the initial population can affect the convergence of MOEAs [15], HMOEA is equipped with a module called 'initial mutation' which uses the polynomial mutation operator on every solution of the initial population with a high mutation probability of $(p_m = 0.5)$. The purpose of 'initial mutation' module is to bring diversity among the solutions of the initial population. After mutation, the total number of solutions is greater than N, which needs to be reduced to N. The environment selection of NSGA-II is chosen for selecting the best N solutions in which all solutions from the initial population and after random mutation are ranked using the non-dominated sorting operator which assigns a rank to every solution and store them in different fronts. Thereafter, the crowding distance operator is

applied front-wise to choose less crowded solutions. Based on the rank and crowding distance, the best *N* solutions are selected. This is referred to as 'Evaluate population (P)' module in the HMOEA as shown in Fig. 75.1 after initialization.

At this stage, the termination condition is imposed on HMOEA in the decision box as shown in Fig. 75.1, which gets terminated after a fixed number of generations. It is noted that a sufficiently large number of generations are provided here so that the convergence of HMOEA can be observed with respect to the original NSGA-II. If the termination condition is not met, the population [which is now referred to as parent population (P)] is considered for 'selection' module. In the selection module, the binary tournament selection is used to choose better solutions for a mating pool. On this pool, the SBX crossover operator is applied to create new solutions in 'Crossover' module followed by polynomial mutation, which is referred to as 'Mutation' module in Fig. 75.1. The population consisting of new solutions is referred to as offspring population (Q). The parent population and offspring population are mixed together and the non-dominated sorting is applied to rank all solutions and store them into different fronts. The crowding distance is applied front-wise. The environment selection of HMOEA, which is same as NSGA-II, selects the best *N* solutions.

The random mutation brings diversity into the population while the environmental selection selects better fitness solutions in order to maintain a balance between diversity and selection pressure. The module for environment selection is shown as 'Environment Selection (P + Q)' in Fig. 75.1.

At this stage, a module of 'Local Search' is used which chooses a least crowded solution from the set of the non-dominated solution, that is, the front-1 solution.

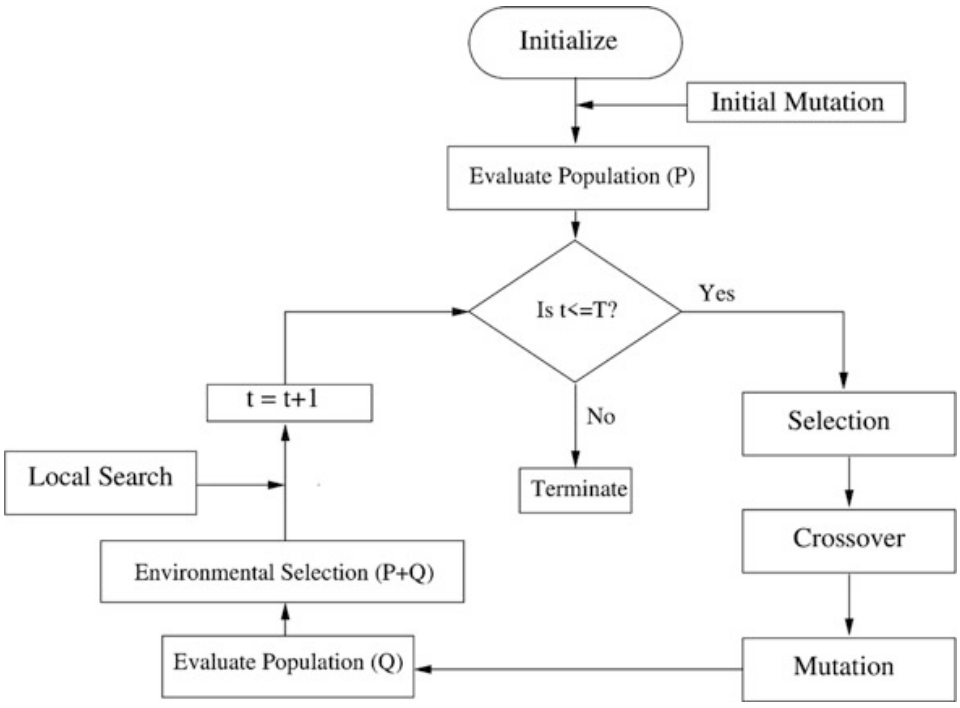


Fig. 75.1 Flowchart of hybrid multi-objective evolutionary algorithm

Thereafter, the multi-objective problem given in Eq. (75.1) is converted into a single-objective optimization problem using $\epsilon-$ constraint method [1]. The modified formulation for LS is given as

$$\begin{aligned} &\text{Minimize} f_i(x) \\ &\text{subject to } f_j \leq \epsilon_j, \forall j = (1, \cdots, M) \text{ and } j \neq i, \\ &x \in S \subset R^n. \end{aligned} \tag{75.2}$$

The single-objective problem of Eq. (75.2) is solved using the method of multipliers [16] in which the Cauchy's steepest descent method is used as the local search technique [16]. The solution after local search replaces its parent solution in the population. The challenge for the frequency of local searches is addressed by executing it once every $(T/5)$ generations, in which T refers to the fixed number of generations for which HMOEA is executed.

75.3 Results and Discussions

In this section, HMOEA is tested on five bi-objective benchmark problems, which are known as ZDT problems [17]. Among the ZDT problems, ZDT1, ZDT2, ZDT3, ZDT4, and ZDT6 problems are chosen. ZDT1, ZDT2, and ZDT3 have 30 number of variables and rest of ZDT problems have 10 number of variables. All variables are having the range from zero to one. ZDT5 problem is not considered because it has binary variables. The details of ZDT problems can be found in [17] which are not included in this paper due to brevity. HMOEA is also compared with the results of NSGA-II. Both MOEAs are run for the same population size of $N = 100$, number of generations $(T = 200)$, crossover probability of $p_c = 0.9$ and mutation probability of $(p_m = 0.1/$ number of variables). For local search, a maximum number of sequences for the method of multipliers is 40 and a maximum number of iterations for the Cauchy's steepest descent method is 20. The penalty parameter is kept 0.1 and the error for terminating the method of multiplier is 0.001.

75.3.1 Effect of Random Mutations

As part of HMOEA, random mutations were performed on the initial population for all test problems. Figure 75.2 shows the effect of initial mutation on the initial population. As can be seen, the population after random mutations is diverse and improved as compared to the randomly generated population after initialization. Due to the characteristics of ZDT6, many solutions are evolved at the extreme of 'f1' objective. As shown by [15], a better initial population helps MOEA for improved and faster convergence.

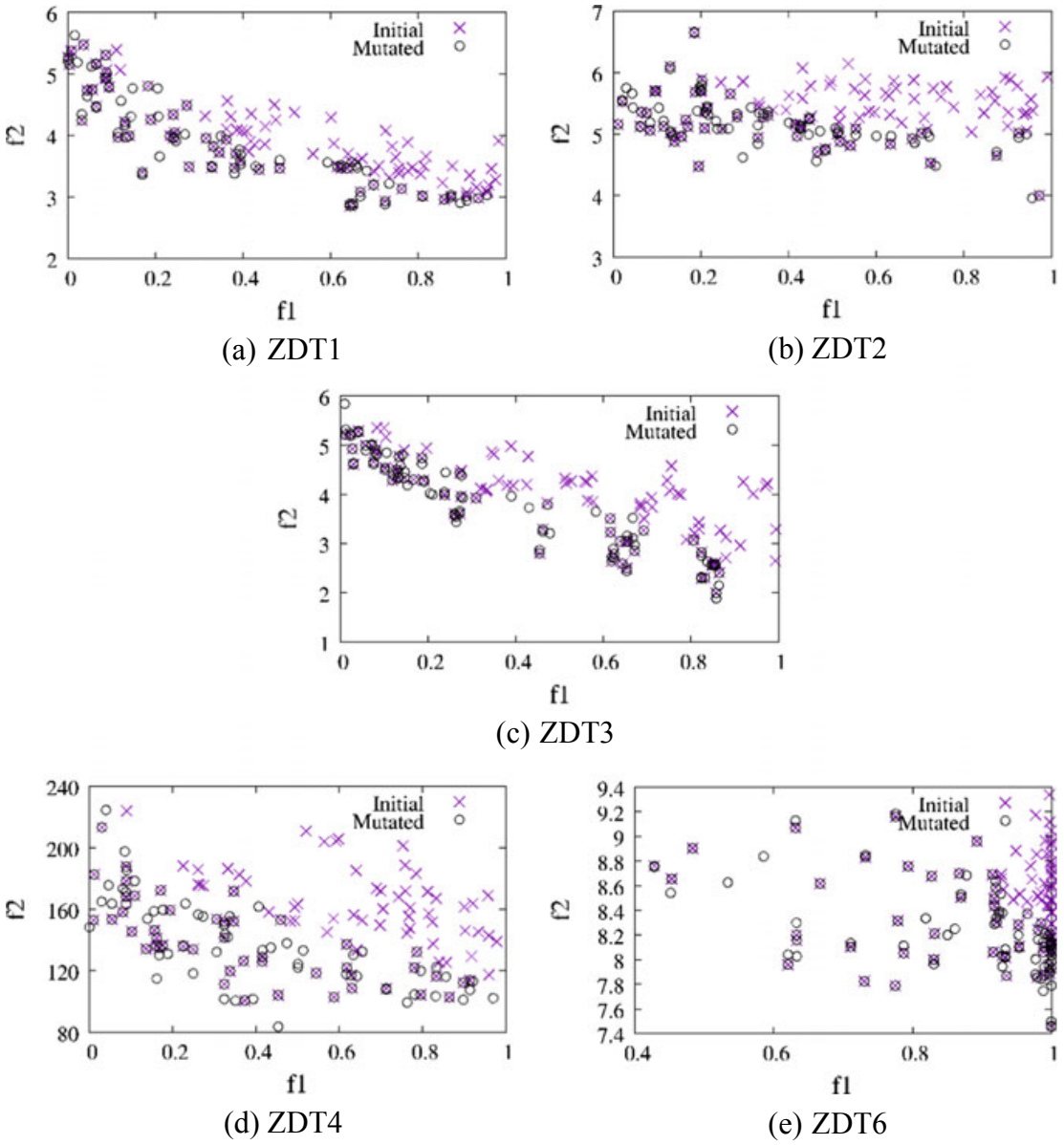


Fig. 75.2 Effect of random mutation on the initial population of HMOEA

75.3.2 Pareto-Optimal Solutions

The non-dominated solutions obtained by HMOEA for the chosen set of problems are shown in Fig. 75.3. It can be seen that HMOEA is able to generate non-dominated solutions that converge to the known Pareto-optimal fronts, for all the problems. It can be seen that ZDT1 has a non-convex Pareto-optimal front, ZDT2 has a convex Pareto-optimal front, ZDT3 has the discontinuous Pareto-optimal front, ZDT4 has a nonconvex Pareto-optimal front, and ZDT6 has nonlinearity along the convex Pareto-optimal front.

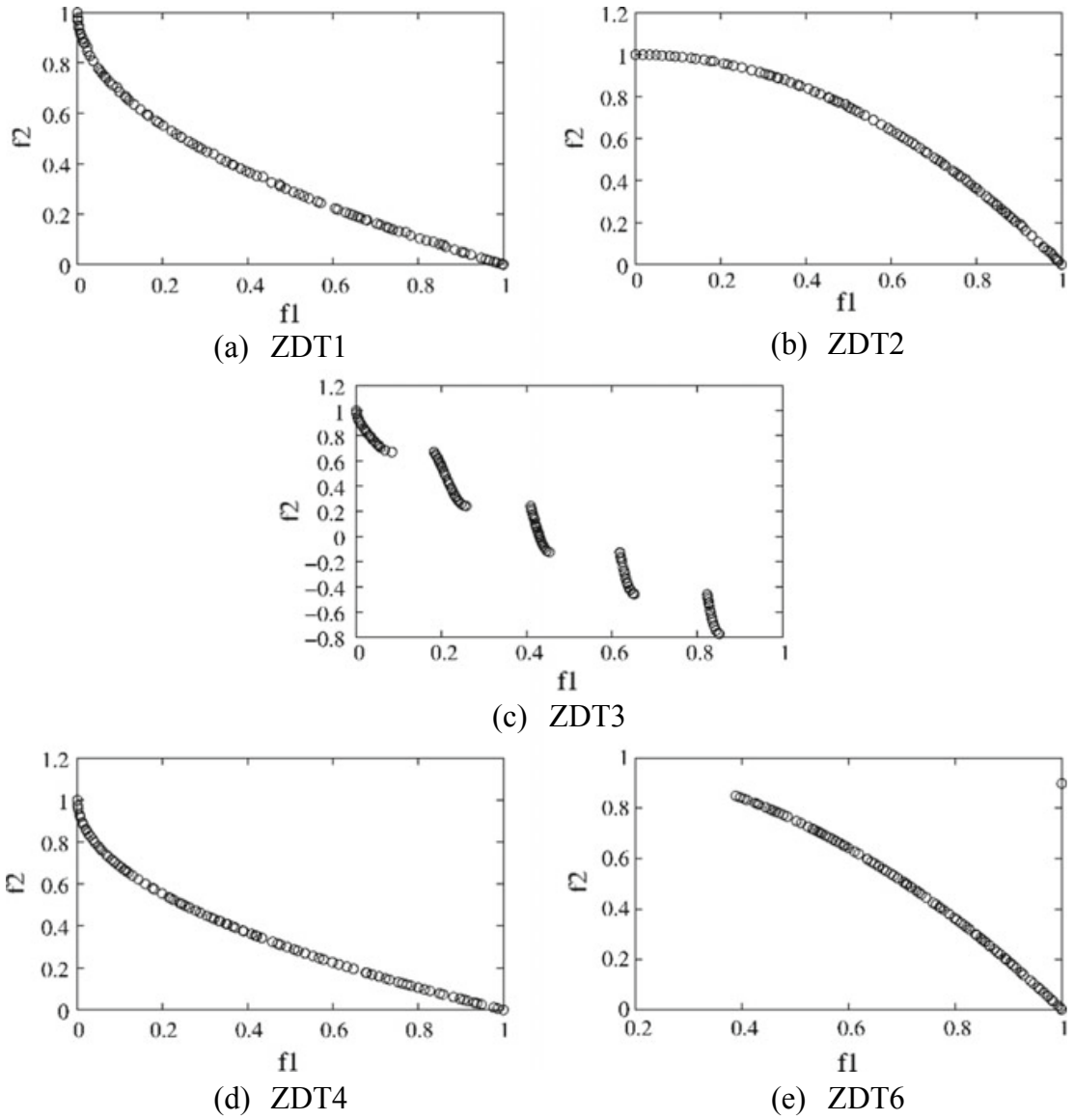


Fig. 75.3 The non-dominated solutions obtained from HMOEA

75.3.3 Convergence Plots

The convergence plots for HMOEA and NSGA-II for all five problems are shown in Fig. 75.4. These plots are made using the hypervolume (HV) indicator which signifies the volume occupied by the approximate set of non-dominated solutions generated by MOEAs and bounded by the reference point such that the reference point is weakly dominated by all the non-dominated solutions in the approximate set [1]. A semi-$\log_{10}$ scale is used for HV values over generations of MOEAs.

It can be seen from the figures that HMOEA has converged and stabilized faster than NSGA-II. The exact generation of convergence for both the algorithms is presented in Table 75.1, in which t_H and t_N denote generations required by

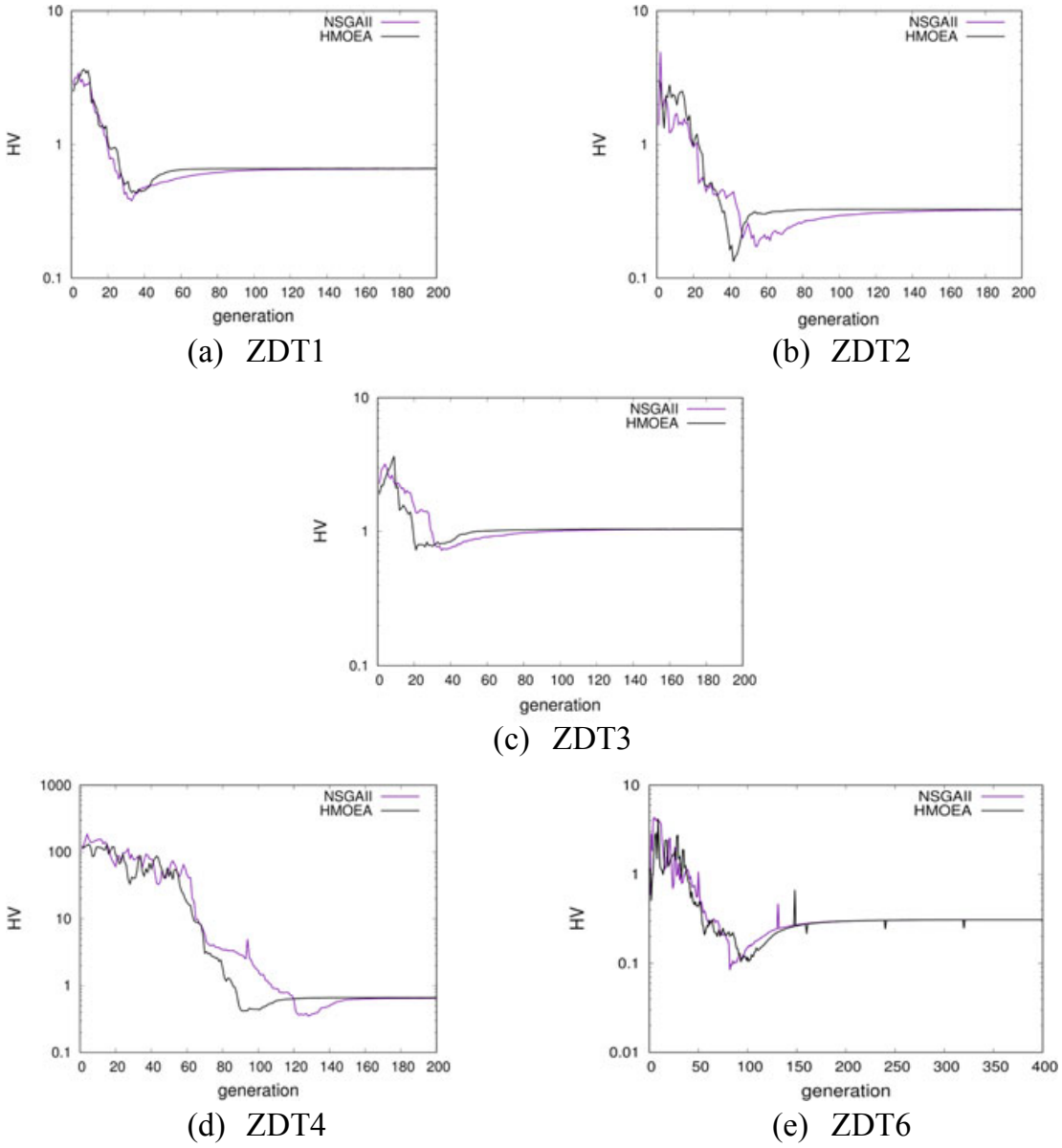


Fig. 75.4 Convergence plots of HMOEA over NSGA-II for all problems. HV values are plotted with $\log_{10}$ scale

Table 75.1 Number of generations required by HMOEA and NSGAII for solving ZDT problems

Problems	ZDT1	ZDT2	ZDT3	ZDT4	ZDT6
t_H	65	80	70	120	200
t_N	120	140	95	170	200
t_S	55	60	25	50	0

HMOEA and NSGA-II respectively, and t_S denotes the total generations saved by HMOEA over NSGA-II. Except for ZDT6, HMOEA is found to be superior over NSGA-II in saving a handful of generations.

75.4 Conclusions

This paper was an attempt to hybridize an existing MOEA with random mutations and local searches. From the results, it can be concluded that the random mutations with high mutation probability improved the diversity of the initial population. The convergence plots showed that local searches helped HMOEA to converge faster and saved a handful number of generations (or computations). Since the local search was performed on only one solution from the population and at the rate of $(T/5)$ generations, a well converged and nicely distributed set of non-dominated solutions was evolved that converged to known Pareto-optimal fronts. As a note on future work, HMOEA can be further improved by including an adaptive termination condition based on some performance indicators like HV. HMOEA also needs to be tested on other benchmark and engineering optimization problems and can be compared with other benchmark hybrid MOEAs.

References

1. Deb, K.: Multi-Objective Optimization Using Evolutionary Algorithms, 1st edn. Wiley, New York (2001)
2. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multi-objective genetic algorithm: NSGA-II. IEEE Trans. Evol. Comput. **6**(2), 182–197 (2002)
3. Zitzler, E., Laumanns, M., Thiele, L.: SPEA2: Improving the strength pareto evolutionary algorithm for multiobjective optimization. In: Giannakoglou, K., et al. (eds) Evolutionary Methods for Design, Optimization and Control with Application to Industrial Problems (EUROGEN 2001), pp. 95–100. International Center for Numerical Methods in Engineering (CIMNE) (2002)
4. Sharma, D., Basha, S.Z., Kumar, S.A.: Diversity over dominance approach for many-objective optimization on reference-points-based framework. In: Deb et al. (eds) 10th International Conference Proceedings on Evolutionary Multi-Criterion Optimization. Michigan State University, East Lansing, USA, (2019) (to appear)
5. Agarwal, D., Sharma D.: Experimental study on bound handling techniques for multi-objective particle swarm optimization. In: Snášel, V., Abraham, A., Krömer, P., Pant, M., Muda, A. (eds) Innovations in Bio-Inspired Computing and Applications, vol 424, pp. 555–564. Advances in Intelligent Systems and Computing, Cham (2016)
6. Sharma, D., Collet, P.: An archived-based stochastic ranking evolutionary algorithm (asrea) for multi-objective optimization. In: Proceedings of The 12th Annual Conference on Genetic and Evolutionary Computation (GECCO'10), pp. 479–486. ACM, New York (2010)
7. Sindhya, K., Deb, K., Miettinen, K.: A local search based evolutionary multi-objective optimization approach for fast and accurate convergence. In: International Conference on Parallel Problem Solving from Nature, pp. 815–824 (2008)
8. Adrian, T.A., Hopgood, E.-M., Nolle, L., Battersby, A.: Hybrid genetic algorithms: a review. Eng. Lett. **13**(2), 124137 (2006)
9. Sindhya, K., Miettinen, K., Deb, K.: A hybrid framework for evolutionary multi-objective optimization. IEEE Trans. Evol. Comput. **17**(4), 495–511 (2013)
10. Ichibuchi, H., Narukawa, K.: Some issues on the implementation of local search in evolutionary multi-objective optimization. In: Genetic and Evolutionary Computation Conference, pp. 1246–1258 (2004)

11. Ishibuchi, H., Hitotsuyanagi, Y., Wakamatsu, Y., Nojima, Y.: How to choose solutions for local search in multi-objective combinatorial memetic algorithms. In: International Conference on Parallel Problem Solving from Nature, pp. 516–525 (2010)
12. Sharma, D., Kumar, A., Deb, K., Sindhya, K.: Hybridization of SBX based NSGA-II and sequential quadratic programming for solving multi-objective optimization problems. In: The proceedings of IEEE Congress on Evolutionary Computation (CEC), pp. 3003–3010, Singapore (2007)
13. Kumar, A., Sharma, D., Deb, K.: A hybrid multi-objective optimization procedure using PCX based NSGA-II and sequential quadratic programming. In: The Proceedings of IEEE Congress on Evolutionary Computation (CEC), pp. 3011–3018, Singapore (2007)
14. Sindhya, K., Deb, K., Miettinen, K.: Improving convergence of evolutionary multi-objective optimization with local search: a concurrent hybrid algorithm. Nat. Comput. **10**(4), 1407–1430 (2011)
15. Sharma, D., Deb, K., Kishore, N.N.: Domain-Specific initial population strategy for compliant mechanisms using customized genetic algorithm. Struct. Multi. Optim. **43**(4), 541–554 (2011)
16. Deb, K.: Optimization for Engineering Design: Algorithms and Examples, 2nd edn. PHI Learning Pvt. Ltd., New Delhi, India (2012)
17. Zitzler, E., Deb, K., Thiele, L.: Comparison of multiobjective evolutionary algorithms: empirical results. Evol. Comput. J. **8**(2), 125–148 (2000)

Chapter 76
Local Search Heuristics-Based Genetic Algorithm for Capacitated Vehicle Routing Problem

Deepak Sharma, Sukhomay Pal, Aayush Sahay, Puneet Kumar, Gaurav Agarwal and Katravath Vignesh

Abstract Vehicle routing problem (VRP) is an NP-hard problem, which looks for the optimal route for vehicles to deliver goods to the customers with minimum transportation cost. Since transportation and logistics incur a good percentage of cost on to the basic price of the products delivered, VRP has been an interest in many organizations and also for many researchers. Due to the limitations of exact algorithms, the heuristics/meta-heuristics have been a choice for evolving an optimal solution for VRP. In this paper, various local search heuristics are coupled with genetic algorithm (GA), and two more intense local search heuristics are proposed. GA is tested on six different capacitated VRP instances. Results demonstrate the efficacy of GA in obtaining the solutions close to the best-known solutions that too in less computation time.

Keywords Local search Heuristics · Vehicle routing problem · Genetic algorithm

D. Sharma (✉) · S. Pal · A. Sahay · P. Kumar · G. Agarwal · K. Vignesh
Department of Mechanical Engineering, Indian Institute of Technology Guwahati, Guwahati, Assam 781039, India
e-mail: dsharma@iitg.ac.in

S. Pal
e-mail: spal@iitg.ac.in

A. Sahay
e-mail: aayush.sahay@alumni.iitg.ac.in

P. Kumar
e-mail: puneet.kumar@alumni.iitg.ac.in

G. Agarwal
e-mail: gaurav.agarwal@iitg.ac.in

K. Vignesh
e-mail: v.katravath@iitg.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_76

76.1 Introduction

Vehicle routing problem (VRP) is an NP-hard problem, which is difficult to solve in polynomial time [1]. It mainly finds optimal routes in transportation and logistics for people and organizations. In a typical supply chain, a requirement for VRP can be seen at different levels of the same organization ranging from transporting from its subsidiaries and manufacturing unit, delivering products to the users, or transporting employees from home to the organization and vice versa.

VRP has been solved by using sophisticated exact or heuristic algorithms in the literature. Exact algorithms aim to find a strict optimal solution, while heuristics are best in finding a good feasible solution. Another set of algorithms is hybrid approaches in which exact algorithm is coupled with heuristic to solve VRP. Due to the size of modern problems getting bigger day by day, heuristics and hybrid algorithms have been mostly used to solve VRP.

In specialized literature, the first GA to successfully solve the VRP was proposed by Prins [2] in which GA without trip-delimiters was hybridized with a local search procedure. GA showed competitive results with that of Tabu search (TS) and other heuristics. Nagata [3] proposed an edge-assembly crossover that was coupled with GA for solving capacitated VRP. Perboli et al. [4] proposed a hybrid GA in which the chain mutation was introduced with the *k*-chain local search. Masum et al. [5] in another effort represented trips of the route as chromosome of an individual. The authors used two types of crossover in which information from one individual was inserted into another to create a new child. Nazif and Lee [6] proposed a crossover operator, which was designed using a complete undirected bipartite graph. Vidal et al. [7] proposed a bi-criteria evaluation of individuals to maintain diversity in the population in order to avoid pre-mature convergence. The memetic algorithm proposed by Cattaruzza et al. [8] combined the split and local search procedure of Prins [2] along with the population diversity management of Vidal's approach [7]. Instead of single crossover and mutation, Contreras-Bolton and Parada [9] explored different combinations of crossover and mutation operators for better exploration and creation of new solutions. Pop and Horvat-Marc [10] considered theoretical aspects of seven neighborhoods motivated from VRP structure. These neighborhoods have been used in their local search approach such as intra-route and inter-route neighborhoods. For intra-route neighborhoods, the authors used the two-point move and three-opt move for improving single-vehicle routes. The shortest route was then selected. Similar steps were used for inter-route neighborhoods in which variations of opt moves to improve the solutions were performed. Kumar and Panneerselvam [11] studied different types of crossover operators used in GA in solving VRPs. Apart from that the authors also proposed 'sinusoidal motion crossover (SMC)' operator. Subramanian et al. [12] implemented a sequence of set partitioning in which an iterated local search-based heuristics and set partitioning were coupled. Wang and Lu [13] proposed a nearest addition method and sweep algorithm to solve the VRPs. Furthermore, an elitism conservation strategy was also used to hold superior chromosomes which can replace

inferior chromosomes in the population. Chen et al. [14] proposed a hybrid heuristic method called 'iterated variable neighborhood descent (IVND)' algorithm which used perturbation strategy designed by a crossover operator that helped the algorithm to move out from the local minima.

From the above studies, it can be observed that many efforts have been made to make GA effective for solving VRP by introducing heuristics either at crossover and mutation or by introducing local search and moves. Since in the study [10], various local search and moves have been discussed but not implemented together, and the contribution of this paper is to implement those local search moves with GA along with two proposed intense local search heuristics and test on the known set of capacitated VRPs.

The paper is organized into five sections. Section 76.2 presents a formulation for capacitated VRP. Section 76.3 presents the description of GA which is coupled with various local search heuristics. Section 76.4 presents results and discussion. In Sect. 76.5, the paper is concluded with future work.

76.2 Problem Formulation for Capacitated VRP

VRP is represented as a set of nodes with the depot at index 0, a heterogeneous fleet of vehicles with load capacity W, and individual cost constraint L. Here, L can represent any cost, monetary, or otherwise. It is used to represent distance in our problem statement since cost incurred by a vehicle is proportional to the distance covered by it.

The mathematical model is defined on a complete undirected network $G = (N, E)$. N consists of node 0 and C, where node 0 is the depot and C represents the customers with a non-negative demand d_i. E represents the set of all connected edges where every edge (i,j) has a travel cost c_{ij} associated with it. The set of vehicles is represented by V. We define a decision variable X_{ij}^v, which is 1 if there is a vehicle in V going from node i to j, and 0 otherwise. Our main objective is to minimize the cost as represented in (76.1), and the constraints are defined in (76.2)–(76.6).

$$\text{Minimize} \sum_{v \in V} \sum_{(i,j) \in E} c_{ij} X_{ij}^v \tag{76.1}$$

$$\sum_{v \in V} \sum_{j \in N} X_{ij}^v = 1, \quad \forall i \in N \tag{76.2}$$

$$\sum_{i \in \mathrm{N}} d_i \sum_{j \in N} X_{ij}^v \leq W, \quad \forall v \in V \tag{76.3}$$

$$\sum_{j \in \mathrm{N}} X_{0j}^{v} = 1, \quad \forall v \in V \tag{76.4}$$

$$\sum_{i \in N} X_{ik}^{v} - \sum_{j \in N} X_{kj}^{v} = 0, \quad \forall v \in V, \quad \forall k \in \mathrm{N} \tag{76.5}$$

$$X_{ij}^{v} \in \{0, 1\}, \ \forall (i,j) \in E, \quad \forall v \in V. \tag{76.6}$$

Equation (76.2) represents that each customer is visited by only one vehicle. Equation (76.3) imposes constraints on the capacity of a vehicle. Equation (76.4) represents that a vehicle can leave the depot only once. Equation (76.5) is designed such that the number of vehicles entering the depot or customer k is equal to the number of vehicles leaving. Equation (76.6) represents binary variable X_{ij}^{v}.

76.3 Local Search Heuristics for Genetic Algorithm

GA has been chosen for implementing local search heuristics for solving VRPs. Following are the steps of GA:

Step 1: Initialize random population (π) without duplicate. Set generation counter $t = 1$ and a maximum allowed generations $T = 1000$;
Step 2: Calculate fitness of each chromosome in π by using the split algorithm [2];
Step 3: Select two random parent chromosomes by using binary tournament selection from π;
Step 4: Perform crossover on the selected parent chromosomes to create a new chromosome by using specialized single-point, multi-point, and orderly crossover operators;
Step 5: Perform local search heuristics on the new chromosome with a probability p_m by using randomized or comprehensive local search method;
Step 6: Select new chromosome if it is unique (no duplicate) and has better fitness value than the worst fitness chromosome in π. Replace the worst fitness chromosome by the new chromosome in π. Else, do not select the new chromosome;
Step 7: Perform intense local search after every $\tau = 500$ generations;
Step 8: Increase count $t = t + 1$. If $t \leq T$, go to **Step 3**. Else, terminate.

GA starts with an initial random population (π) of size (σ) in Step 1. A real-valued chromosome is chosen to represent a solution. An example is shown

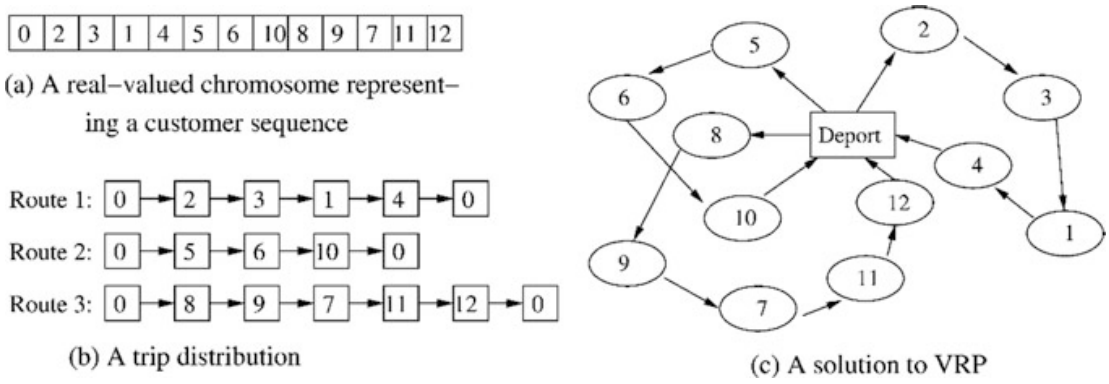


Fig. 76.1 A chromosome representation and trip distribution. Note that node 0 represents the depot

in Fig. 76.1a in which the customer sequence is represented the sequence in which vehicles are going to visit them. In Fig. 76.1b, the trips are distributed such that constraint on the capacity can be satisfied. Finally, a solution to VRP is shown in Fig. 76.1c.

The duplicate chromosomes are removed from the population for which an idea of spacing is used [2] in which the chromosomes should have the least similarity. Therefore, a split algorithm [2] is used that arranged the sequence of nodes in a chromosome such that least fitness of the chromosome can be found. An example is shown in Fig. 76.2 in which a chromosome, which is a solution to VRP, is shown before and after the split algorithm for making trips and calculating the least fitness for the given chromosome.

Once the initial random population is generated and the fitness is calculated for each chromosome, GA enters into its standard loop of generation at Step 3 at which two parent chromosomes are selected by using the binary tournament selection. The crossover operators are then applied to these two chosen chromosomes in order to create one new chromosome at Step 4. Three types of crossover operators are used with equal probability, which is referred to as single-point, multi-point, and orderly crossover operators. These operators are explained with the help of Fig. 76.3.

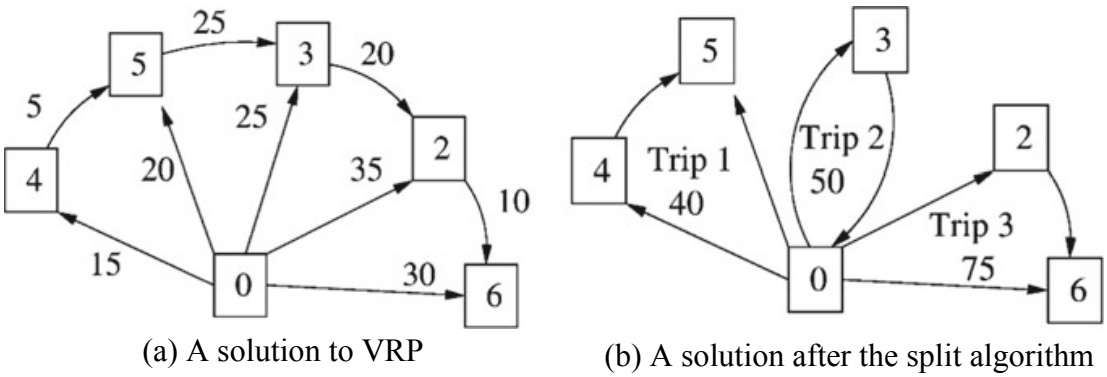


Fig. 76.2 Fitness evaluation of chromosome using the split algorithm [2]

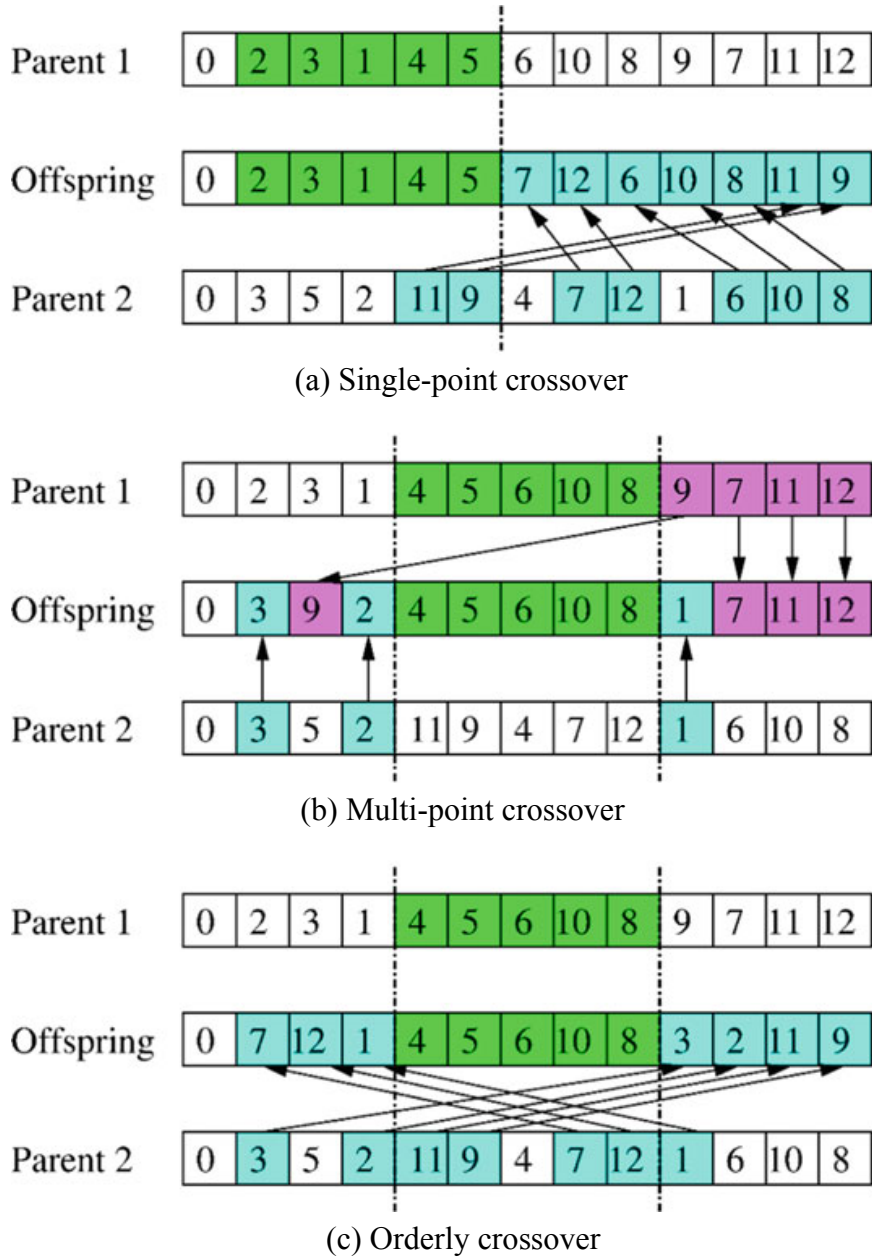


Fig. 76.3 Different crossover operators used with GA

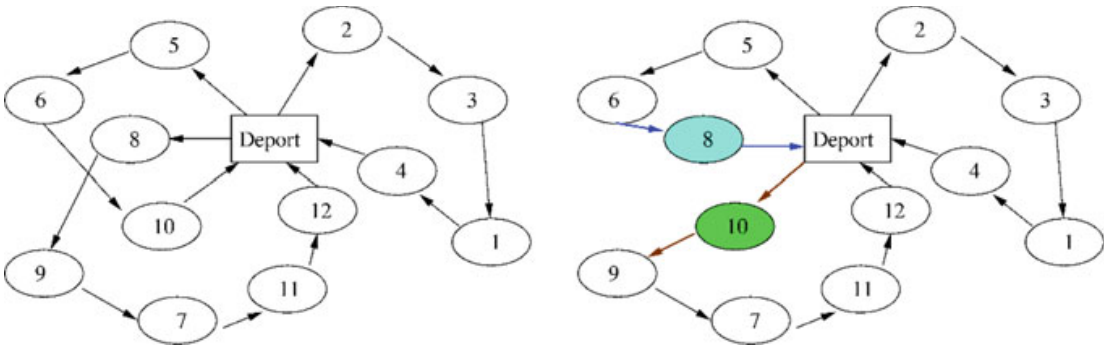
In a single-point crossover, a mating site is chosen randomly, say at node 5. All the nodes of parent 1 copied into the new chromosome. Thereafter, the remaining nodes are copied from parent 2 such that there is no duplicate node in the new chromosome, but in the given sequence of parent 2. In multi-point crossover operator, two random sites are chosen so that a sequence between these sites can be copied to the new chromosome as shown in Fig. 76.3b. At the empty places of a new chromosome, the nodes are copied in the same places as of parent 2, but without repetition of nodes in the new chromosome. There can be a few places in the new chromosome are not filled, which can be filled from the sequence of parent 1 as shown in the figure. In orderly crossover operators, two sites are chosen randomly, and then

the sequence from parent 1 is copied into the new chromosome at the places. Thereafter, the nodes, which do not appear in the new offspring, are copied in the given order/sequence as given in the parent 2.

After performing crossover, mutation is performed with a probability of p_m either by using randomized local search or comprehensive local search method at Step 5. Both of these random search methods are adopted from [2], in which the two nodes are selected randomly in the randomized local search. In the comprehensive local search method, two nodes are selected after scanning through all pairs of nodes. After selecting two nodes, six local search heuristics adopted from [10] are implemented which consist of (i) one-point move, (ii) two-point move, (iii) two-opt move, (iv) Or-opt move, (v) three-point move, and (vi) four-point move. Due to page restriction, the details for these moves are not given. However, the same can be found in [10].

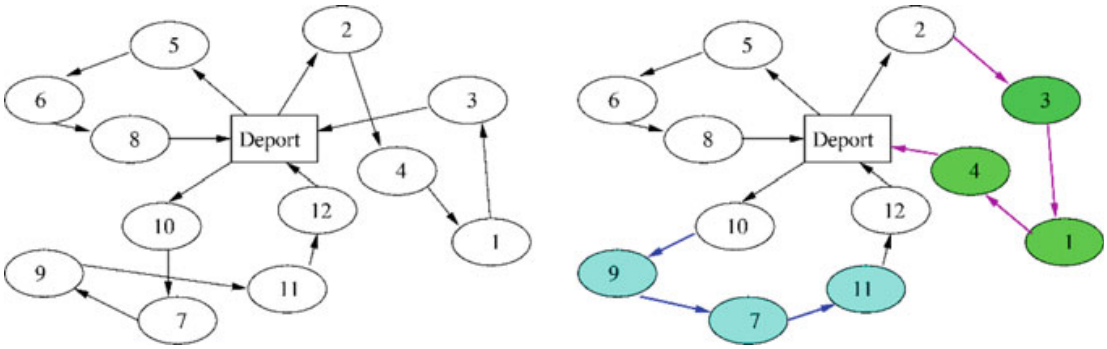
At Step 6, the new chromosome is checked for insertion in the population. Its fitness is found by using the split algorithm. If the fitness of the new chromosome is better than the worst fitness chromosome in the population and also there is no duplicate, then it is selected. Otherwise, it is removed.

The main contribution of this paper is the proposed intense local search which is performed at Step 7. It has been observed through simulation that the local search performed at Step 6 gets stuck at the local optima. In order to escape from these local optima, intensive local search is done on all the chromosomes in the population after every $\tau = 500$ generation. First intense local search is 'k-nearest neighbors local search' in which the nodes, which are close by based on Euclidean distance, bring closer in the sequence as well. We loop through all nodes one by one, find that particular node's k-nearest neighbors, and apply local search between that node and the ith nearest neighbor, where $1 \leq i \leq k$. Each iteration of this local search stops at the first improving move. For a graphical representation, the k-nearest neighbors local search is shown in Fig. 76.4. Another intense local search is 'trip-wise improvement procedure.' It has been observed through simulations that the improvement in the solution fitness is very low and by minor tweaking the sequence of a chromosome, the solution can converge to the global optimal



(a) Before local search (b) After local search

Fig. 76.4 An intense local search by using the k-nearest neighbors local search



(a) Before local search (b) After local search

Fig. 76.5 Trip-wise improvement procedure for local search

solution. In this method, every solution is scanned in every trip and the nodes' sequence in the given trips is tweaked. A graphical presentation is shown in Fig. 76.5. At last, the generation counter is increased by one and the termination condition is check. If not terminated, GA then starts from Step 3 till it gets terminated.

76.4 Results and Discussion

GA with local search heuristics is now tested on a set of benchmark CVRPs from the literature. The chosen set of the problems belong to four Augerat instances [15], one Christofide instance [16], and one Fisher instance [17]. These instances have a different number of nodes, which are presented in Table 76.1. The best-known solution from the literature is also presented in the table to compare the outcome of the proposed GA with various local search heuristics.

For all instances, GA starts with an initial population of size 15, and it runs for 1000 generations. The probability of mutation is set to $p_m = 0.1$. GA runs on Intel core i7 with 1.8 GHz processor.

Table 76.1 Details of Augerat, Christofide, and Fisher instances. The letter 'C' refers to Christofide instance, the letter 'A' refers to Augerat instance, and the letter "F' refers to Fisher instance. The number signifies a number of nodes

Instances	A33	A48	A54	A61	C76	F72
Nodes	33	48	54	61	76	72
Best-known solution	661	1073	1167	1035	832	238

Table 76.2 Solutions obtained for different instances of CVRP

S. No.	Problem instance	Solution obtained	Number of vehicles	Average time (s)
1	A33	662	5	0.945
2	A48	1170	7	2.494
3	A54	1428	7	4.189
4	A61	1040	9	2.083
5	C76	846	10	2.876
6	F72	241	4	3.389

76.4.1 Obtained Solutions and Computational Time

The solutions obtained by solving various instances are presented in Table 76.2. It can be seen that the local heuristics coupled with GA can generate the solutions close to the best-known solutions given in Table 76.1. Since a homogenous fleet of vehicles is assumed for solving CVRP, the minimum of vehicles are also presented in the table. The average computation time is also reported which is calculated after running GA 30 times from different initial population. It can be seen that the developed GA is not expensive for solving CVRP having a maximum of 76 nodes. It is noted that the instances are different from each other; therefore, the computation time is not linear with the number of nodes in the problem instances.

76.4.2 Convergence Plots

The convergence plots for all test instances are shown in Fig. 76.6. It can be seen that the best fitness solution is improved drastically in very early generations and

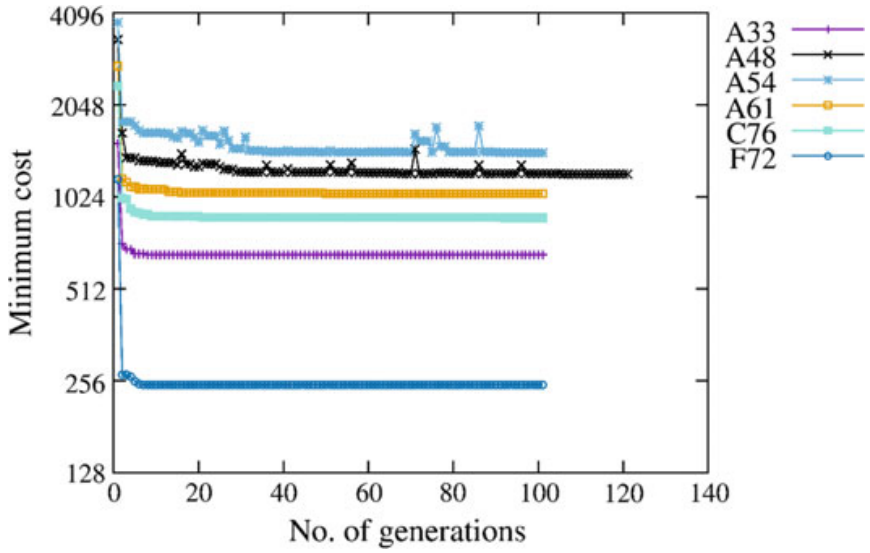


Fig. 76.6 Convergence plot of GA for different instances of CVRPs. The minimum cost values are plotted on a $\log_2$ scale

further their performance improved very slowly. The local search heuristics coupled with GA are able to make it efficient by improving chromosome drastically in the early generations.

76.5 Conclusions

In this paper, various local search heuristics were coupled with GA to improve its performance for solving various test instances of CVRP. From the results, it can be concluded that the hybrid GA can be used for generating a solution for CVRP as it was able to find the solution close the best-known solutions of the test instances. Since the computation time of GA was small, it can also be used for a large test instance and for application problems of CVRP. In the future, GA can be improved further so that it can generate the best-known solutions. Also, a large set of CVRP test instance needs to solve for assessing its performance.

References

1. Lenstra, J.K., Rinnooy Kan, A.H.G.: Complexity of vehicle routing and scheduling problems. Networks **11**, 221–227 (1981)
2. Prins, C.: A simple and effective evolutionary algorithm for the vehicle routing problem. Comput. Oper. Res. **31**(12), 1985–2002 (2004)
3. Nagata, Y.: Edge assembly crossover for the capacitated vehicle routing problem. In: European Conference on Evolutionary Computation in Combinatorial Optimization, pp. 142–153. Springer (2007)
4. Perboli, G., Pezzella, F., Tadei, R.: EVE-OPT: a hybrid algorithm for the capacitated vehicle routing problem. Math. Methods Oper. Res. **68**(2), 361–382 (2008)
5. Masum, A.K.M., Shahjalal, M., Faruque, F., Sarker, I.H.: Solving the vehicle routing problem using genetic algorithm. Int. J. Adv. Comput. Sci. Appl. **2**(7), 126–131 (2011)
6. Nazif, H., Lee, L.S.: Optimised crossover genetic algorithm for capacitated vehicle routing problem. Appl. Math. Model. **36**(5), 2110–2117 (2012)
7. Vidal, T., Crainic, T.G., Gendreau, M., Prins, C.: A hybrid genetic algorithm with adaptive diversity management for a large class of vehicle routing problems with time-windows. Comput. Oper. Res. **40**(1), 475–489 (2013)
8. Cattaruzza, D., Absi, N., Feillet, D., Vidal, T.: A memetic algorithm for the multi-trip vehicle routing problem. Eur. J. Oper. Res. **236**(3), 833–848 (2014)
9. Contreras-Bolton, C., Parada, V.: Automatic combination of operators in a genetic algorithm to solve the traveling salesman problem. PLoS ONE **10**(9), e0137724 (2015)
10. Pop, P.C., Horvat-Marc, A.: Local search heuristics for the generalized vehicle routing problem. In: 2012 International Conference on System Modeling and Optimization (ICSMO 2012), IPCSIT, vol. 23, pp 84–87 (2012)
11. Kumar, S.G.V., Panneerselvam, R.: A study of crossover operators for genetic algorithms to solve VRP and its variants and new sinusoidal motion crossover operator. Int. J. Comput. Intell. Res. **13**(7), 1717–1733 (2017)
12. Subramanian, A., Uchoa, E., Ochi, L.S.: A hybrid algorithm for a class of vehicle routing problems. Comput. Oper. Res. **40**(10), 2519–2531 (2013)

13. Wang, C., Lu, J.: A hybrid genetic algorithm that optimizes capacitated vehicle routing problems. Expert Syst. Appl. **36**(2), 2921–2936 (2009)
14. Chen, P., Huang, K., Dong, X.: Iterated variable neighborhood descent algorithm for the capacitated vehicle routing problem. Expert Syst. Appl. **37**(2), 1620–1627 (2010)
15. Augerat, P., Belenger, J.M., Benavent, E., Corberan, A., Naddef, D., Rinaldi, G.: Computational results with a branch and cut code for the capacitated vehicle routing problem. Technical Report RR 949-M, University Joseph Fourier, Grenoble, France (1995)
16. Christofides, N., Eilon, S.: An algorithm for the vehicle dispatching problem. J. Oper. Res. Soc. **20**, 309–318 (1969)
17. Fisher, M.L.: Optimal solution of vehicle routing problems using minimum k-trees. Oper. Res. **42**, 626–642 (1994)

Part VIII
Design

Chapter 77
Performance Evaluation of Adhesive Spur Gear with the Influence of Properties of Adhesive

R. Shailesh, M. Ramu, M. Govindaraju, K. Karthikeyan and V. Satheeshkumar

Abstract The effect of adhesive properties on performance of spur gear made of adhesive system referred to as adhesive spur gear is investigated. In addition, the repair ability and reusability of adhesive spur gear towards sustainability are also investigated. Two-part epoxy adhesive system constituting resin and hardener was used to fabricate spur gear samples. In order to modify adhesive properties, the hardener to resin ratio was varied. The repair ability and reusability of spur gear made by the adhesive system were checked by breaking a few teeth of the adhesive spur gears and repaired by adhesive. The repaired gears were tested and compared with the performance of actual spur gears made of adhesive. The gears were tested by a gear test rig developed indigenously. The performance of the gears with different hardener to resin ratios was evaluated by monitoring the number of cycles. The results show that the number of cycles increases with decrease in hardener to resin ratio. This is due to the better strength of the adhesive system with the resin-rich formulation. There is not much difference in performance between repaired gear and the corresponding actual gear. This ensures the sustainability of adhesive gears over conventional polymer gears.

Keywords Spur gear · Polymer gear · Epoxy adhesive · Gear performance · Hardener to resin ratio · Gear repair ability

R. Shailesh · M. Ramu · M. Govindaraju
Department of Mechanical Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore, India

K. Karthikeyan
Department of Mechanical Engineering, Sri Ramakrishna Engineering College, Coimbatore 641022, India

V. Satheeshkumar (✉)
Department of Production Engineering, National Institute of Technology, Tiruchirappalli 620015, India
e-mail: satheeshv@nitt.edu

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_77

77.1 Introduction

Performance evaluation of the spur gears made of the adhesive system with the influence of adhesive properties and checking repair ability and reusability of the gears made of adhesive are the key activities of the present study. In this regard, White [1] invented a composite plastic metal gear in which different plastic materials were used for having gear mesh, and the gear hub was made of metal. In this case, the hybridization of plastic gear and metal gear hub together eliminates distortion and slip. Beddoe [2] invented a composite gear employing a set of deformable teeth disposed slightly in advance of the hard teeth to transfer load through hard teeth only when it exceeds a critical level. Santi [3] invented composite gears made from a rigid material like metal and from a thermoplastic material for the abnormal conditions like abrupt stop occurs. Kurokawa et al. [4] investigated the performance of injection-moulded spur gears with different types of composites made of poly-ether-ether-ketone with different types of carbon fibres (CFs). It was found that a composite gear reinforced with CF of the highest density indicated the highest load carrying capacity irrespective of the test conditions. Kim [5] proposed a method to evaluate the wear and durability of plastic spur gears made of acetal and nylon materials. In the results, acetal gears exhibited a large amount of wear with the development of plastic deformation and increased tooth stiffness due to brittle nature of the material. In the case of nylon gears, reduced amount of wear and prolonged service life with less gear temperature were observed due to the polymer softer in nature. Mao [6] proposed a design method for polymer composite gear by relating polymer gear surface temperature and wear rate. It was found that polymer gear wear rate increases with an increase in critical load for a specific geometry and then decreases. Nozawa et al. [7] investigated the tribological behaviour of polymer sheets adhered to steel gear teeth. It was found that noise suddenly increased when only one of the polymer sheets was failed from the gear surface due to low adhesive strength.

Kansal et al. [8] addressed the problem of evaluating the temperature distribution and thermal residual stresses that are developed due to the non-uniform cooling of the injection-moulded polystyrene gear. Senthilvelan and Gnanamoorthy [9, 10] investigated the influence of the gear rotational speed on performance of polymer composite gear tooth. It was found that gear rotational speed affects the performance of gears made of both the materials at different high running speeds and torques. The glass fibre reinforced nylon six gears showed better performance when compared with unreinforced gears due to its better mechanical strength and thermal deformation resistance. It was reported that the rise in the gear temperature under loading is due to hysteresis effects and friction. Kirupasankar et al. [11] reported the effect of applied torque on the transmission efficiency of polyamide nanocomposite spur gears. The increase in gear tooth temperature affects the tooth shape and thereby affects the gear performance. The improvement in mechanical properties of polyamide nanocomposite gears results in better power transmission efficiency. Anand Mohan and Senthilvelan [12] evaluated fatigue performance of

polypropylene symmetric gears and asymmetric gears with and without the reinforcement of glass fibre. The results showed better load-carrying capacity of asymmetric gears when compared with symmetric gears due to the increased gear tooth width at the root region. Chaubey and Jain [13] conducted a detailed review on meso- and micro-gears used in various micro-systems and devices used for transmitting power and/or motion due to their reduced weight, dimensional accuracy, zero backlash, compact ability, better functional and operating characteristics, longer service life and sustainable performance. The review reveals about the manufacturing of meso- and micro-spur gear by various advanced manufacturing processes and the directions towards manufacturing quality gears. Hoskins et al. [14] generated a sound frequency spectrum with the influence of various polymeric gear materials and operating conditions. Results showed that there is a significant effect of surface characteristics, wear and temperature corresponding to the respective sound power levels. The effects of material, reinforcement and geometry show that the tribological properties of gear materials could be used to predict noise generated and the inaccuracy significance in polymer gear trains. Dearn et al. [15] described friction and wear by controlling the running temperatures by using a series of solid lubricant coatings like molybdenum disulphide, graphite flake, boron nitride (hexagonal) and poly-tetra-fluoro-ethylene (PTFE) deposited on flanks of the polymer gear teeth. Results showed a significant reduction in frictional forces with PTFE.

In line with adhesive materials, mostly the adhesive systems with two parts, namely resin and hardener are used for structural applications. In a compression testing of an adhesive system that is made of epoxy group resin and amine group hardener with different mixing ratios, the results show brittle behaviour with the epoxy-rich systems. This is due to the evolution of a rigid macromolecular structure. The adhesive mixture with hardener-rich formulation shows ductile behaviour [16]. Satheeshkumar and Ganesh Narayanan [17] demonstrated the mechanical performance of adhesive with two parts, namely resin—bisphenol A epichlorohydrin and hardener—polyamidoamine. It was postulated that resin-rich formulation turns the adhesive sample stronger, and hardener-rich formulation turns the adhesive sample more ductile.

In the above detailed literature review, it is observed that numerous investigations have been carried out on the manufacturing of polymer gears with different polymer materials. Different types of filler materials and reinforcements have been used for improving the gear performance. The polymer gears have been used predominantly, and efforts have been taken to achieve equivalent and better performance than metal gears. The polymer gear conventionally manufactured by the injection moulding process is replaced with a new gear once it is damaged. The damaged thermoplastic gears may be melted and reused. In the case of thermoplastic gears with typical fillers and reinforcements, repair ability is one of the major issues. The idea here is to develop polymer gears by using adhesive materials by choosing suitable properties for having better performance and to check the repair

ability and reusability of adhesive gears in line with sustainable gear manufacturing. In the present preliminary study, the influence of hardener to resin ratio of adhesive on the performance of adhesive gear is investigated with actual and repaired gears.

77.2 Methodology

77.2.1 Materials Used and Mechanical Properties

Commercially available structural adhesive was used as a polymer matrix in this study. The adhesive system is a combination of two parts, namely resin (bisphenol A epichlorohydrin) and hardener (polyamidoamine). The suitable mixing ratio of the adhesive system is of 1:0.8 by weight as per supplier's instruction.

Gear Test Rig. Figure 77.1 shows the schematic of the test rig for evaluating the adhesive spur gear performance. It has a 0.5 HP DC shunt motor with maximum speed up to 1500 rpm. It is coupled with an alternator having the same capacity. The input to the motor and the field of the alternator is controlled externally by two separate autotransformers. The autotransformer is used to adjust speed and torque. There is an additional rectifier circuit consisting of diodes and capacitor for the AC to DC conversion. A 100 Ω rheostat is used as the load across the alternator. The testing gear is connected to the alternator shaft while the standard steel gear is connected to the motor shaft.

Evaluation of Adhesive Spur Gear Performance. Figure 77.2 shows the schematic of spur gear used for fabrication. Table 77.1 shows the parameters of the adhesive spur gear (Anand Mohan and Senthilvelan [12]). The adhesive mixture with different properties by varying hardener to resin (H/R) ratio like 0.6:1, 0.7:1, 0.8:1, 0.9:1 and 1:1 was cast into the suitable mould with a thickness of 6 mm (Fig. 77.3a). The cast adhesive mixture into the mould was allowed to cure for more than 24 h and removed from the mould carefully. Figure 77.3b shows fabricated spur gear made of epoxy adhesive system and a steel gear for a reference.

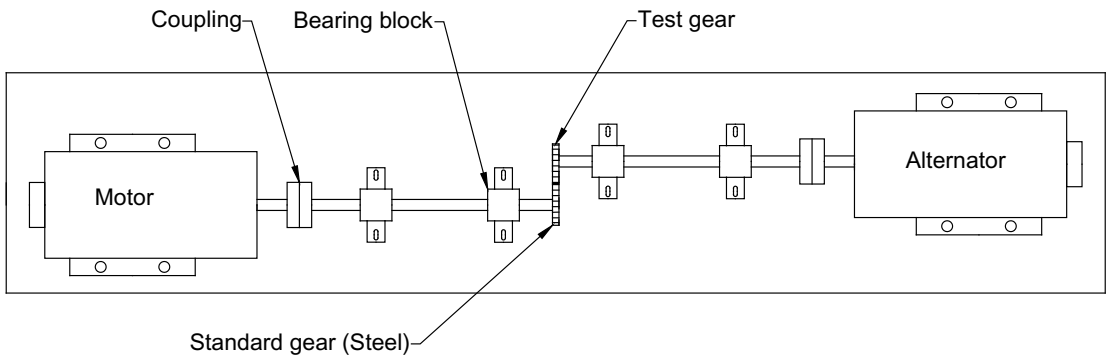
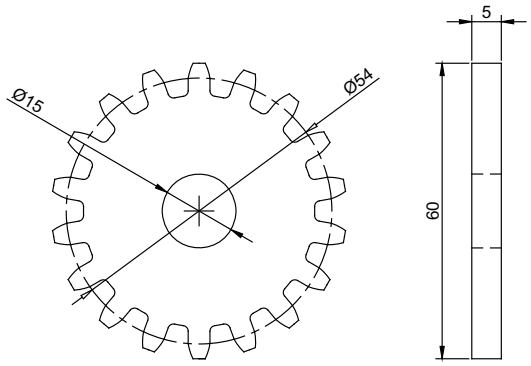


Fig. 77.1 Schematic of gear test rig



All dimensions are in 'mm'

Fig. 77.2 Schematic of spur gear

Table 77.1 Test gear parameters

	Adhesive spur gear
Pitch circle diameter	54 mm
Addendum	3 mm
Module	3 mm
Pressure angle	25°
Number of teeth	18
Face width	6 mm
Root circle diameter	47 mm
Fillet radius	1.14 mm

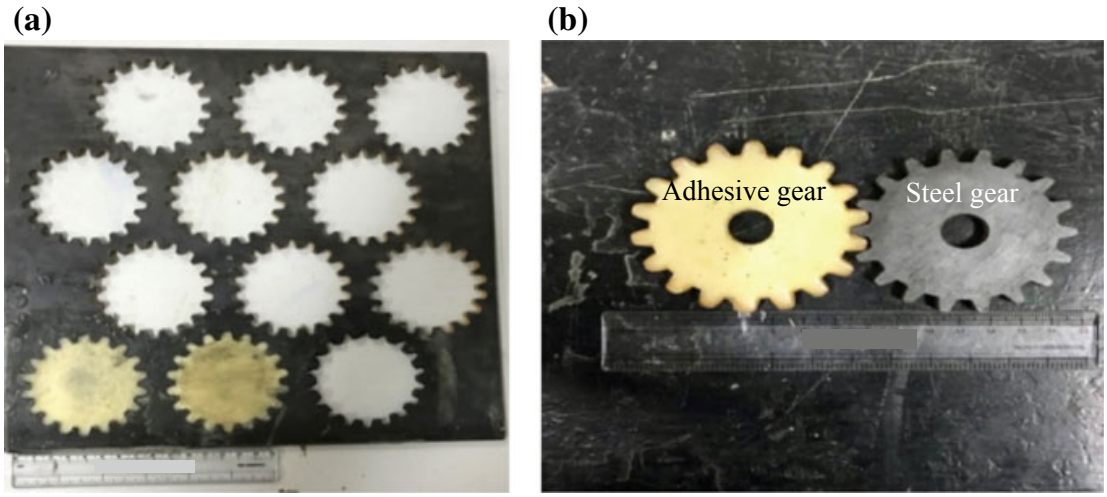


Fig. 77.3 **a** Making of adhesive gear using mould and **b** fabricated adhesive gear and steel gear

Further, a tooth in each fabricated adhesive gear with varying hardener to resin ratio was broken deliberately and inserted into the mould. The damaged location was repaired with the adhesive mixture maintaining the corresponding hardener to resin ratio. In the present work, the gear tested directly from the moulding process is referred to as actual gear, and the gear deliberately broken and repaired is referred to as repaired gear.

The adhesive gears were tested at a constant motor speed of 300 rpm and applied torque of 2 Nm. The platform of the testing setup was made rigid and heavy to minimize vibration. Initially, using a pre-test gear attached to the alternator, electrical equivalent of all torque conditions were calculated by measuring the voltage and current across the input of DC motor. The voltage and current were measured to be 35 V and 1.8 A, respectively. The voltage, current and speed transformation to equivalent torque were calculated by the Eqs. (77.1), (77.2).

$$P = V \times I \tag{77.1}$$

where P—power, V—voltage in V and I–current in A.

Also

$$P = \frac{2\pi NT}{60} \tag{77.2}$$

where N—Motor speed in rpm and T—torque in Nm.

During testing of gears, gear life was evaluated by monitoring the number of cycles. The failure of gear was ensured by sound difference and followed by tooth breakage. Three gears were tested in each case, and the fourth gear was tested in the case of not observing good repeatability.

77.3 Results and Discussion

77.3.1 Effect of Hardener to Resin Ratio of Adhesive on Failure Pattern of Gears

Figure 77.4a, b shows the failure pattern of actual gears and repaired gears with varying H/R ratios of adhesive. It is found that 1 or 2 teeth have failed in all cases. There is not much difference in failure pattern between actual gears and repaired gears. In the case of repaired gears, the failure of gear teeth is observed randomly, and not at the repaired location always.

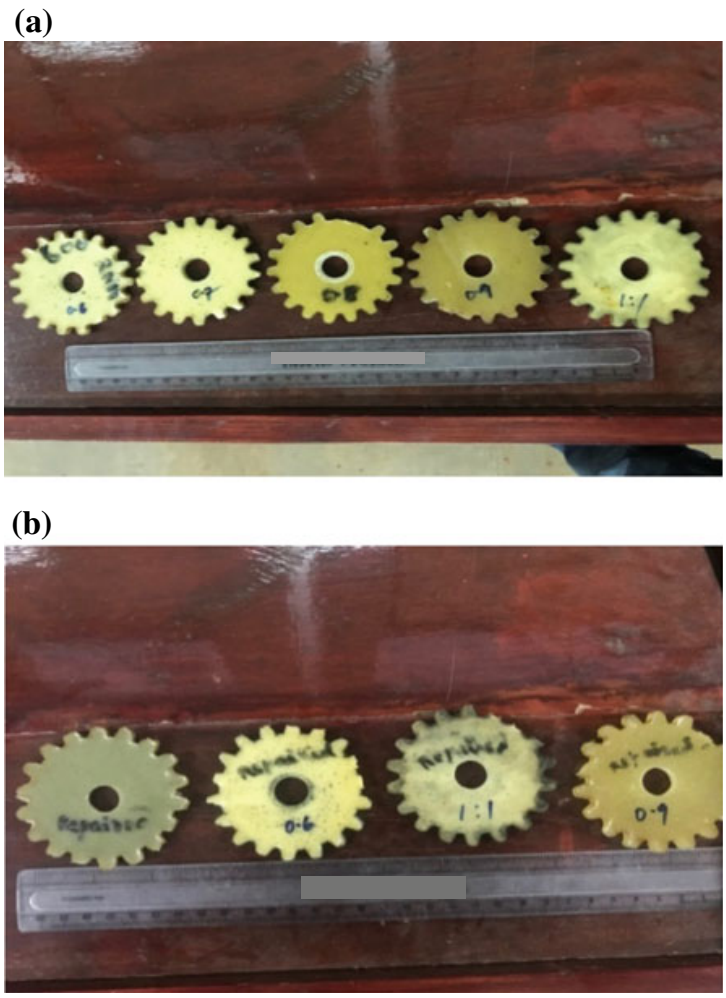


Fig. 77.4 Failure pattern of adhesive gears: **a** actual gears and **b** repaired gears

77.3.2 Effect of Hardener to Resin Ratio of Adhesive on Gear Performance

Figure 77.5 shows a comparison of maximum number of cycles achieved by the adhesive gears with different hardener to resin ratios. The comparison of maximum number of cycles between actual and repaired adhesive spur gears is also shown. It is observed that the number of cycles of adhesive spur gear increases with decrease in H/R ratio of adhesive. The maximum number of cycles of about 40200 is observed for H/R ratio of 0.6:1. Based on the maximum number of cycles, the decrease in the percentage of maximum number of cycles is calculated as 19.4, 28.4, 44.8 and 60.4% with hardener to resin ratios of 0.7:1, 0.8:1, 0.9:1 and 1:1, respectively. This is due to the rich formulation of the resin increases the strength of the adhesive (Satheeshkumar and Ganesh Narayanan [17]) and thereby increases the gear performance. There is not much difference between actual and repaired adhesive spur gears.

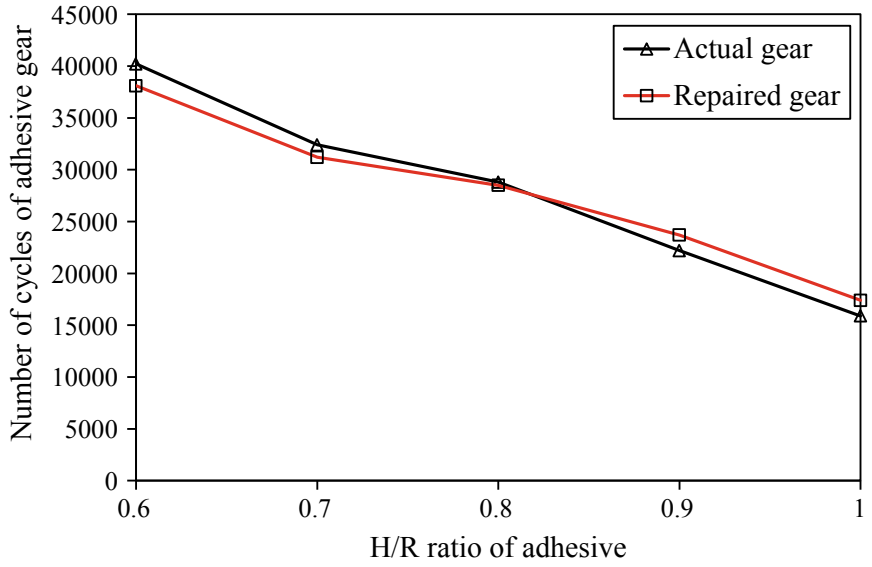


Fig. 77.5 Comparison of maximum number of cycles of adhesive spur gear between actual and repaired cases with different hardener to resin ratios (Data Variation: ±1200 cycles)

77.4 Conclusions

From the present study, the conclusions are summarized as follows:

- The performance of adhesive spur gears in terms of number of cycles increases with decrease in hardener to resin ratio of the adhesive system. This is due to the rich formulation of resin that improves the strength of adhesive. It is understood that improved strength of adhesive with a rich formulation of resin increases the strength of gear and thereby increases the number of cycles.
- There is not much difference in the performance of actual and repaired adhesive spur gears irrespective to hardener to resin ratio. This ensures that the adhesive gear has good repair ability and reusability. In the case of conventional thermoplastic gears made by injection moulding process, reusability is possible by melting the damaged gears and reinjection moulding process. But, the repair ability of those gears is difficult by injection moulding process and not suggestible as well.
- It is suggested that the adhesive with resin-rich formulation could be fabricated in sheet or plate form, and water jet cutting process could be used for spur gear cutting process in the case of mass production.

References

1. US3013440, White, J.T. [1961]
2. US3602058, Beddoe, B. [1971]
3. US005852951 A, Santi, J.D. [1998]
4. Kurokawa, M., Uchiyama, Y., Nagai, S.: Performance of plastic gear made of carbon fiber reinforced polyether-ether-ketone: Part 2. Tribol. Int. **33**, 715–721 (2000)
5. Kim, C.H.: Durability improvement method for plastic spur gears. Tribol. Int. **39**, 1454–1461 (2006)
6. Mao, K.: A new approach for polymer composite gear design. Wear **262**, 432–441 (2007)
7. Nozawa, J.I., Komoto, T., Kawai, T., Kumehara, H.: Tribological properties of polymer-sheet-adhered metal hybrid gear. Wear **266**, 893–897 (2009)
8. Kansal, G., Rao, P.N., Atreya, S.K.: Study: temperature and residual stress in an injection moulded gear J. Mater. Process. Technol. **108**, 328–337 (2001)
9. Senthilvelan, S., Gnanamoorthy, R.: Effect of gear tooth fillet radius on the performance of injection molded Nylon 6/6 gears. Mater. Des. **27**, 632–639 (2006)
10. Senthilvelan, S., Gnanamoorthy, R.: Effect of rotational speed on the performance of unreinforced and glass fiber reinforced Nylon 6 spur gears. Mater. Des. **28**, 765–772 (2007)
11. Kirupasankar, S., Gurunathan, C., Gnanamoorthy, R.: Transmission efficiency of polyamide nanocomposite spur gears. Mater. Des. **39**, 338–343 (2012)
12. Anand, Mohan N., Senthilvelan, S.: Preliminary bending fatigue performance evaluation of asymmetric composite gears. Mech. Mach. Theory **78**, 92–104 (2014)
13. Chaubey S.K., Jain, N.K.: State-of-art review of past research on manufacturing of meso and micro cylindrical gears. Precis. Eng. (2017) (In Press)
14. Hoskins, T.J., Dearn, K.D., Kukureka, S.N., Walton, D.: Acoustic noise from polymer gears —a tribological investigation. Mater. Des. **32**, 3509–3515 (2011)
15. Dearn, K.D., Hoskins, T.J., Petrov, D.G., Reynolds, S.C., Banks, R.: Applications of dry film lubricants for polymer gears. Wear **298–299**, 99–108 (2013)
16. d'Almeida, J.R.M., Monteiro, S.N.: The effect of the resin/hardener ratio on the compressive behaviour of an epoxy system. Polym. Test **15**, 329–339 (1996)
17. Satheeshkumar, V., Ganesh Narayanan, R.: Investigation on the influence of adhesive properties on the formability of adhesive bonded steel sheets. Proc. Inst. Mech. Eng. Part C: J. Mech. Eng. Sci. **228**(3), 405–425 (2014)

Chapter 78
Design Analysis and Modification of Sugarcane Fibrizer Hammer at Wonji Shoa Sugar Factory (WSSF) Ethiopia

Tesfaye Mathewos, Perumalla Janaki Ramulu and Esmael Adem Esleman

Abstract Wonji Shoa Sugar Factory (WSSF) is one of the top sugar factories in Ethiopia which produces sugar and electricity from sugarcane using three critical plants: mill plants, sugar processing plant, and a cogenerating power plant. Among three plants, the mill plant is the most important plant used to crush the cane for separating sugar juice from bagasse. In this plant, Fibrizer hammer device is used for cruising sugarcane. The Fibrizer performs the major share in milling process for fibrizing the chopped cane after set of cane knives using rotating knives. Most of the times, WSSF is facing breakages of Fibrizer hammer by which losses incurring to the factory. The main aim of the paper is to analyze the failure causes of existing Fibrizer hammer with respect to design and modify the necessary changes in design for Fibrizer hammer at WSSF. By which, get high preparation index (PI) and optimize in power consumption and cost. In addition, hardness of the existing Fibrizer hammer was tested to know the mechanical strength. For modification of existing Fibrizer hammer: comparison was made in terms of design, and load along with harmonic simulation analyses performed using finite element analysis. The results from finite element analysis revealed that the modified Fibrizer hammer

T. Mathewos
School of Mechanical, Chemical and Materials Engineering, Adama Science and Technology University, Adama, Ethiopia & Wonji Shoa Sugar Factory, Wonji, Adama, Ethiopia
e-mail: tesfamathewos@gmail.com

T. Mathewos
Wonji Shoa Sugar Factory, Wonji, Adama, Ethiopia

P. J. Ramulu (✉)
Centre of Excellence for Advanced Manufacturing Engineering, School of Mechanical, Chemical and Materials Engineering, Adama Science and Technology University, Adama, Ethiopia
e-mail: perumalla.janaki@astu.edu.et

E. A. Esleman
Program of Mechanical Design and Manufacturing Engineering, School of Mechanical, Chemical and Materials Engineering, Adama Science and Technology University, Adama, Ethiopia
e-mail: esm99a@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_78

having minimum deformation, minimum shear stress, and more centrifugal force has a relation to achieve better preparation index.

Keywords Fibrizer hammer · Harmonic analysis · Preparation index (PI) · FEM · Wonji Shoa Sugar Factory

78.1 Introduction

For completing the preparation of cane juice in the sugar industries, Fibrizer hammer is used after knifing. Formerly to break the cane, initial knifing was followed by heavy crusher. In this method, cane was not crushed properly in most of the time which leads to more power consumption and increase miscellaneous tasks. To overcome these issues, a few researchers studied and suggested the upgradation of technology to the sugarcane industries, in which Fibrizer hammer is the upgradated technology to all the sugarcane industries for crushing the cane. With this technology, betterment in level of preparation of index (PI) and reduction in the power consumption are attained during processing. For the best operations, less maintenance, low power consumption, and slip ring motors with high capacities are used, whereas steam turbines were used formerly for roller crushing instead of electric motors [1–4]. There were some of the studies carried out based on the Fibrizer hammer, i.e., design of hammer, modification, and optimization of Fibrizer hammer, etc. Podany et al. [5] dealt hammer construction for cane shredder, in which the hammer functions in shredder comprises crushing sugarcane stalks for further sugar production processing. For protection of hammer against wear and dynamic loading, hammer loading and other related calculations were made and suggested the weld deposit materials for strengthening the hammer. Ransing et al. [6, 7] studied about cane roller crusher for attaining good cane preparation for extraction of juice from cane in favor of cane Fibrizer. Cane knifing in chopper, cane crushing in Fibrizer, and extraction of juice in milling like three processes were performed on it. Based on the plant capacity, Fibrizer was changed in size and model, resulting in high PI. Nikam et al. [8] designed and analyzed a new flexible blade-type sugarcane leveler by numerical analysis. They concluded that the efficiency of cane leveler increased by changing design from fixed bitter to flexible blades. By modifying design cane lever, one can overcome the limitations of existing cane leveler. It also increases the productivity of milling section. Efficiency of cane leveler increases, and then it also increases the production.

From the above literature, it is observed that many of the researchers were carried out the research on the functional aspects of Fibrizer hammer, parameters effect, shredder design, power consumption, etc. No researchers discussed the Fibrizer hammers breakages and its consequences in the breakdown of the industry. In the present work, design, analyze and modification of the existing Fibrizer hammer for improvement of sugar production process at Wonji Shoa Sugar Factory (WSSF). In the case industry, i.e., at WSSF, frequent hammer breakage was

observed at the hard-faced welded part and at the tail part of the Fibrizer hammer. Also, at hammer head tip wear, shearing of tip fixing bolt leads to huge crushing downtime. The recorded downtime during 2014–2017 period was 121 h in terms of cost 14,82,421.3 USD. The other problem in the factory was overtime payments during repair of Fibrizer hammer and electrical loss due to interruption of cane bagasse from mill plant to boiler. Based on these issues, the significance of this work is to improve the current Fibrizer hammer to increase juice extraction by improving the preparation index from 85 to 92%; to reduce overtime payments caused by Fibrizer hammer breakage; to supply continuous electrical power by continuous crushing; to save excess bagasse by continuous crushing which is used as a row material for pulp.

78.2 Experimental Methodology

To realize the stated objective, pertinent methodology was followed. The relevant data were collected from both WSSF and existing literature. The detailed specifications of the existing Fibrizer hammer at WSSF are shown in Table 78.1. The material used for Fibrizer hammer was structural steel. Fibrizer hammer bolt was made of mild steel with diameter, length, and thread length 22, 130 and 85 mm,

Table 78.1 Specifications of existing Fibrizer at Wonji Shoa Sugar Factory

S. No.	Description	Unit	Quantity	Remark
1	No's of Fibrizer hammer head	pcs	176	
2	Weight of single Fibrizer hammer head with tip	kg	29	
3	Speed of Fibrizer	RPM	750	
4	Type of Fibrizer			Swing type
5	Hub outer diameter	mm	1670	
6	Number of Hubs	pcs	23	
7	Hub inner diameter	mm	450	
7	Thickness of Hub	mm	50	
8	Material type of Hub			MS IS 2062
9	Design for preparation index	%	92	
10	Current preparation index	%	85	
11	Types of materials			Mild steel
12	Types of tip	kg	3.95 kg/ pcs	Detachable
13	Types of bushing			Stainless steel
14	Materials of tip			Carbide tip

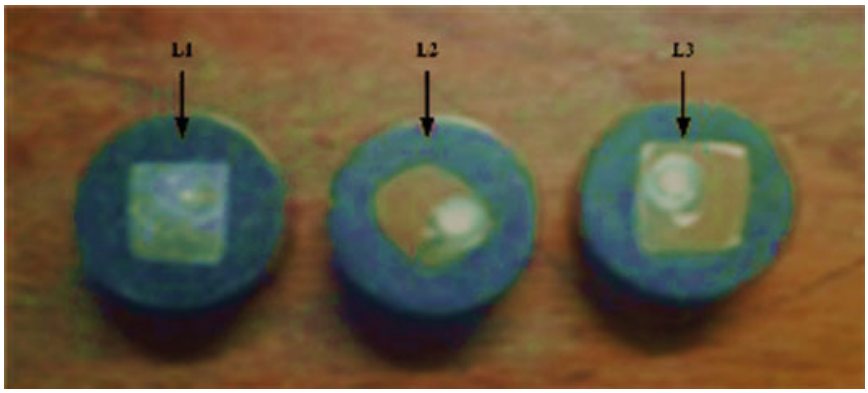


Fig. 78.1 Cut pieces from current Fibrizer hammer

respectively. The material of the Fibrizer hammer was tested at three different locations as L-1, L-2, and L-3 as shown in Fig. 78.1.

78.2.1 Experimental Analysis

Indian made Fibrizer hammer was submitted for failure cause analysis. The hammer was broken at the tail patron during operation. Failure causes were examined with following techniques.

78.2.1.1 Visual Examination

The part was broken and detached from rest of the part. The fracture surfaces were found corroded; the broken part was cleaned using a hydrochloric acid solution fetching from sugar processing department to enable them to be examined more closely. The nature of the failure parts was the same and similar. The welders were welding the worn out Fibrizer hammers and cane knife continuously to use like a spare part.

78.2.1.2 Stress Concentration Factor Analysis

In any machine, machine parts and structures' effect of stress concentration can arise from internal holes or voids formed from any manufacturing process [9].

Stress Concentration Factor Analysis of the Existing Fibrizer Hammer

Figure 78.2 is indicating the complete dimensions of the existing Fibrizer hammer. Center of gravity for load analysis was calculated at different zone of areas (A1, A2, A3, and A4 shown in Fig. 78.2). Parameters considered for existing Fibrizer

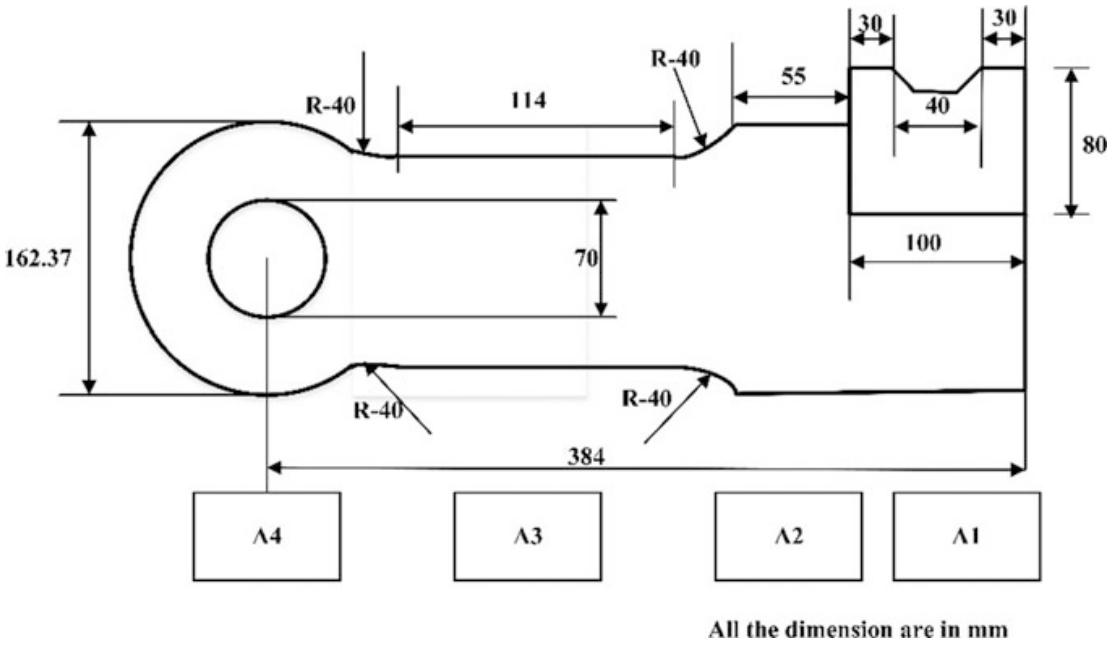


Fig. 78.2 2D schematic representation of current Fibrizer hammer

hammer design analysis were height at tail part of hammer H = 180 mm; web height h = 110 mm; and radius of the fillet of Fibrizer hammer r = 40 mm. From these parameters r/h; H/h; moment of inertia (I) and C which is distance from the neutral axis to the point of interest of hammer. By which, bending moment (M) and maximum shear stress $\left(\delta_{\max} = k_t \frac{MC}{I}\right)$ were evaluated considering a constant factor (k_t) value as 1.58.

78.2.1.3 Load Analysis of Current Fibrizer Hammer

Load analysis of current Fibrizer hammer was evaluated by using speed of Fibrizer (N) = 750 rpm and motor power = 2 MW. By calculating the center of gravity ($Y = (A1Y1 + A2Y2 + A3Y3 + A4Y4)/A1 + A2 + A3 + A4)$), angular speed ($\omega = 2\pi N/60$); linear velocity ($v = \omega \times Y$) and Torque ($T = 60P/2\pi N$).

Figure 78.3 indicates the forces which are acting on Fibrizer hub and hammer during the process. There were three major forces acting on the hub and hammers such as force created by the pin (F_p), centrifugal force (F_c), and force generated by cane being shredded (F_s). Centrifugal force (F_c) was calculated by $mr\omega^2$, where m is mass of hammer and r is distance of gravity center from rotation axis. Loading of one hammer during normal operation was value of tangential force, i.e., force created by pin (F_p) during conventional operation of Fibrizer. For this the force the following presumption were taken such as first hammer carries 75% of loading, second 20% and third 5% and evaluated the exact F_p.

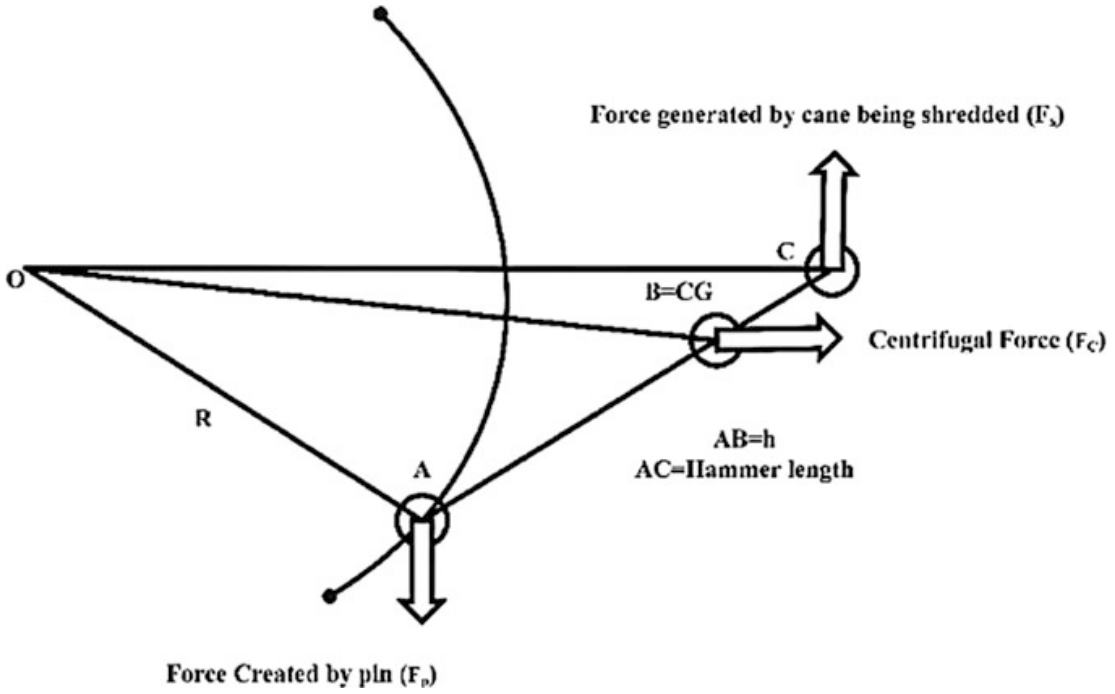


Fig. 78.3 Schematic of skeleton force diagram of Fibrizer hub and hammer

78.2.1.4 Hardness Measurement

The sample was taken from Fibrizer hammer at three places such as L-1, L-2, and L-3 indicated in Fig. 78.4. The hardness of the Fibrizer hammer was tested by Vickers hardness testing machine.

Fig. 78.4 Broken parts of Fibrizer hammer

78.2.1.5 Design Modification of Fibrizer Hammer

Necessity for Modification

The current Fibrizer hammer breakage problems at WSSF are happening frequently. The breakages locations were observed to be at the hard-faced welded part, at the tail part, and sometimes at the hole in which the Fibrizer hammer was fixed with rotatable shaft. The other problem is geometrical structure which exposed the Fibrizer hammer. This problem is not only at WSSF but also other sugar industries in Ethiopia and may be in the world sugar industries where this technology is under implementation. The details of the cost downtime losses were indicated in Table 78.2 in US dollars.

Load analysis of Modified Fibrizer Hammer

Figure 78.5 indicated the complete dimensions of the modified Fibrizer hammer. For this hammer also, the center of gravity for load analysis at different zone of areas such as A1, A2, A3, and A4 as shown in Fig. 78.5 was evaluated. Similar to existing Fibrizer hammer, Center of gravity (C.G), angular velocity (ω), linear velocity (v), torque (T), Centrifugal force (F_c), Work done in shredding per revolution, Force generated by pin (F_p), Force generated by cane being shredded (F_s) and Stress concentration analysis by which maximum shear stress were calculated for modified Fibrizer hammer. In addition to that, cost comparison was also made between existing and modified Fibrizer hammer.

Table 78.2 Loss through downtime

Year	Month	Nature of failure	Recorded downtime (h)	Loss due to downtime in Quintal	Cost of production in USD	Total in USD
2014	December–April	Fibrizer Hammer breakage	60	13,636.4	750,000	1,512,503
2015	April	Fibrizer hammer breakage & shearing of bolts	40.5	9204.6	506,253	
2016	May	Fibrizer hammer breakage & shearing of bolts	20.5	4659.1	256,250	

Note For 1 Qt = 55USD

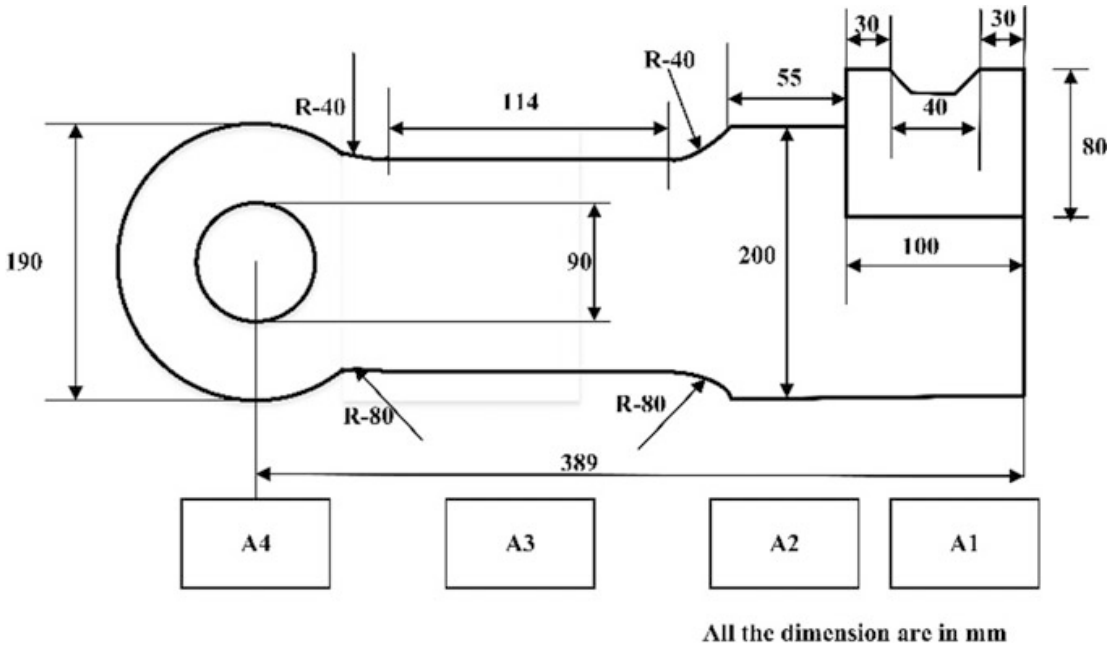


Fig. 78.5 2D schematic diagram of modified Fibrizer hammer

78.2.1.6 Harmonic Analysis Using Finite Element Simulation

A three-dimensional model of Fibrizer hammer was developed in a modeling software SOLDWORKS by comprehensively considering the structural feature, processing requirements, and solution calculation amount for current and modified Fibrizer hammers. This model was imported into ANSYS 18.0, a finite element software as shown in Fig. 78.6. The hammer part adopted the structural steel as per the records. Automatic mesh generation was activated and obtained that the number of nodes were 81,074, unit number was 48,701 for current Fibrizer, whereas 42,601 and 24,487 nodes and units were for modified Fibrizer.

One end of hammer where circular hole existed was fixed and other end of Fibrizer hammer was given the point load of 7136.6 N as boundary conditions. Load was applied along positive y-axis. Simulation results were recorded on harmonic response, total deformation, von Mises stresses, maximum shear stress, and

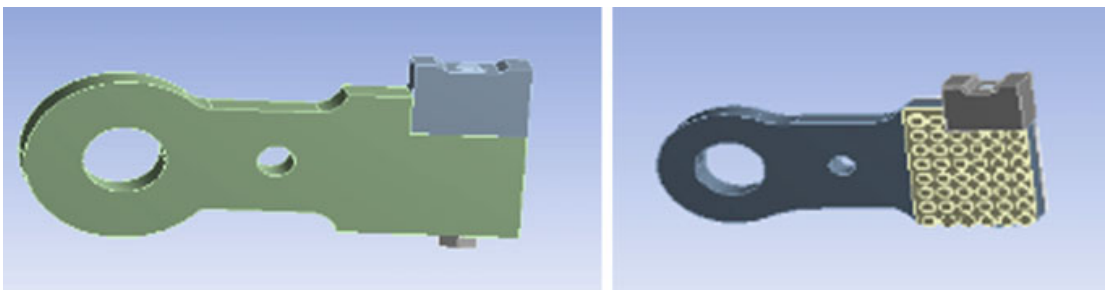


Fig. 78.6 3D model of from solid works to ANSYS current and modified with hard-facing Fibrizer Hammer

maximum amplitude. Young's modulus and passions ratios were 210 GPa and 0.3 considered.

78.3 Results and Discussion

78.3.1 Stress Concentration Factor Analysis Result of the Existing and Proposed Fibrizer Hammer

Stress concentration factor is evaluated with respect to the dimension of exiting Fibrizer hammer and proposed hammer. The final results are tabulated in Tables 78.3 and 78.4.

As the above results shows that the maximum shear stress of the new modified Fibrizer hammer is reduced compared to existing, it indicates that there is less shear stress and less stress concentration factor in modified Fibrizer hammer. The large H/h and the small r/h values give worse stress concentration (k_t).

78.3.2 Load Analysis Result

The result of load analysis for existing and proposed Fibrizer hammer results has been indicated in Table 78.5. The results show that maximum load can be derived with the modified Fibrizer hammer in all directions.

Table 78.3 Stress concentration factor analysis result for existing Fibrizer hammer

H (mm)	h	r	r/h	H/h	k_t	$I = 1/12bc$	$\delta_m = k_tMC/I$ (MPa)
180	100	25	0.25	1.9	1.58	0.0000042	22.6
180	110	40	0.27	1.7	1.34	0.0000055	16.1
180	120	50	0.33	1.58	1.28	0.0000072	13
180	130	60	0.38	1.46	1.23	0.0000092	10.9

Table 78.4 Stress concentration factor analysis for proposed

H (mm)	h	r	r/h	H/h	k_t	$I = 1/12bc$	$\delta_m = k_tMC/I$ (MPa)
190	150	80	0.53	1.27	1.2	0.0000114	8.8

Table 78.5 Load analysis comparison between existing and modified Fibrizer hammer

Hammer	F_c (N/Hammer)	F_p (N)	F_s (N)
Existing Fibrizer hammer	6623	406	314.2
Modified Fibrizer hammer	7170.6	529	342.5

In Table 78.5, the result of centrifugal force of the modified Fibrizer hammer is increased which does work. The higher the centrifugal force, the more would be the preparation index.

78.3.3 Vickers Hardness Results

The test is carried out on each sample at five different locations with three reputations. These results show that the hardness of the material is not the same throughout the hammer. At the hard-facing welded part, the hardness is high and less at the tail part. Due to this, the Fibrizer hammer frequently breaks at the welded part. The welding heat input is important welding parameter, which effects on the structure and properties of the weld metal. At location-1, sample harness result shows that the hardness varied between 110 and 124 VHN. At the location-2, Fibrizer hammer hardness is between 146 and 151 VHN, and at the location-3, hardness is between 152 and 167 VHN.

78.3.4 Cost Comparison Between Existing and Modified Fibrizer

Table 78.6 indicates the cost comparison between existing and modified Fibrizer. The difference in cost between them is $11,088.00.

78.3.5 Harmonic Analysis Results of Existing and Modified Fibrizer Hammers

ANSYS finite element software is used for the simulation to performing harmonic analysis. From the simulation results, total deformation, direction deformation, von Mises stresses, maximum shear stresses, and maximum amplitude are evaluated for existing and modified Fibrizer hammers.

Table 78.6 Cost Comparison

Hammer	No. of hammers	Unit cost in USD	Total cost in USD
Imported Fibrizer hammer	176 pcs	$167	$29,392.00
New fabricated Fibrizer hammer	176 pcs	$104	$18,304.00
	Difference	$63	$11,088.00

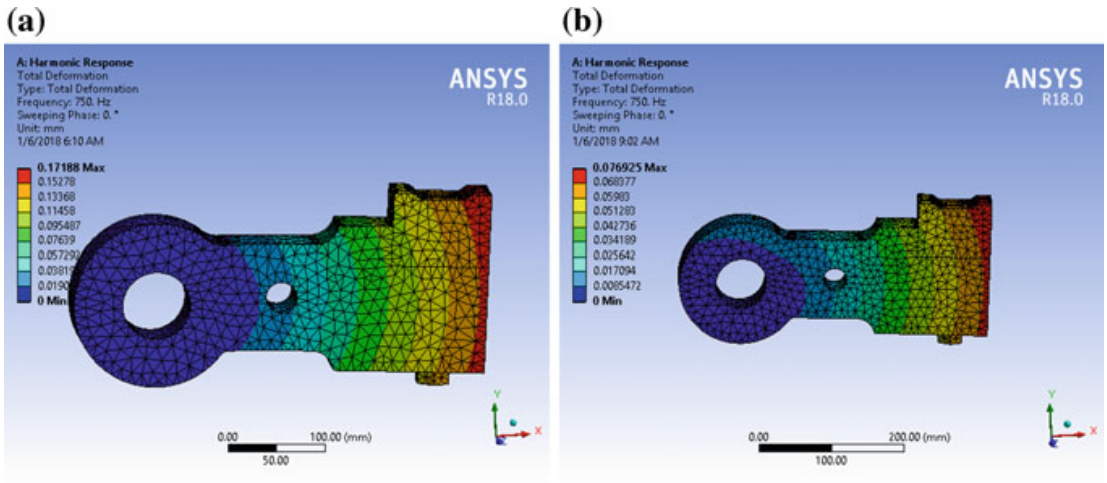


Fig. 78.7 Total deformation of **a** existing and **b** modified FH

From Figs. 78.7, 78.8, 78.9, and 78.10 analysis, results are tabulated in Table 78.7. From the results, it is understood that the total deformation in the modified Fibrizer hammer is lesser than the total deformation of the existing hammer. This shows that the modified Fibrizer hammer is more reliable than the current one. As the result shows that the maximum equivalent (von Mises) stress of current FH is increased by double, the higher the stress, the more the material will be exposed to be broken. The maximum shear stress result in the current Fibrizer hammer is more than the modified Fibrizer hammer. The smaller the radius (r) and web (h) in stepped plate of the Fibrizer hammer, the higher will be the stress.

In Table 78.7, analysis results show that there is less total and directional deformation in modified Fibrizer hammer. The equivalent (von Mises) stress and max. shear stress result also much less by half from the existing.

Similar manner frequency responses are verified for other surfaces also and tabulated in Table 78.8. From Table 78.8 data, the output of harmonic response has been taken in the form of amplitude which can be understood as mean stress development in any engineering components. Modified FH has more amplitude initially and gradually decreased compared to current FH. The overall behavior modified FH is taking less deformations and stresses by which life span of FH definitely will improve.

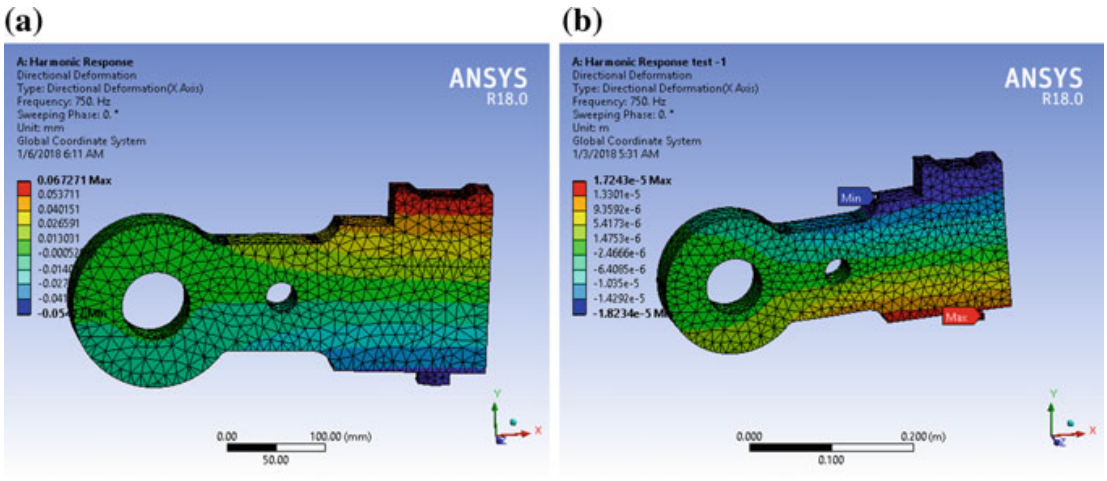


Fig. 78.8 Directional deformation of **a** existing and **b** modified FH

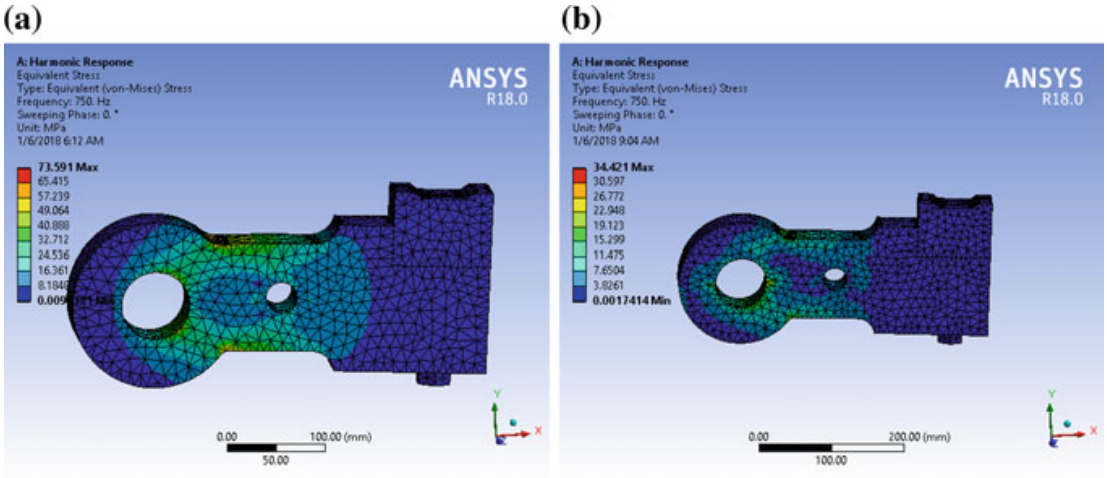


Fig. 78.9 Equivalent (von Mises) stresses of **a** existing and **b** modified FH

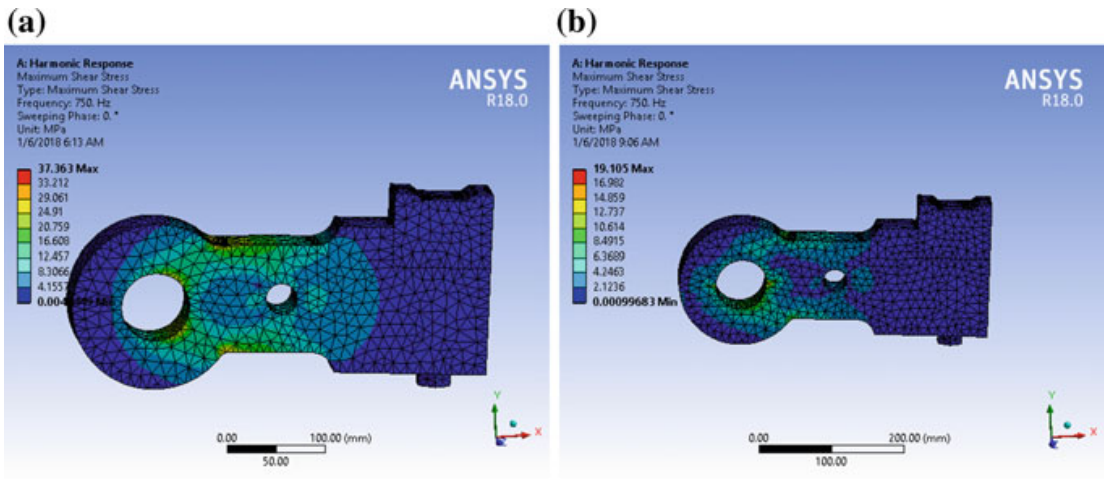


Fig. 78.10 Max. shear stresses of **a** existing and **b** modified FH

Table 78.7 Simulation results on current and modified FH

Parameter	Existing FH	Modified FH
Total deformation	0.1788 mm	0.076925 mm
Max. directional deformation	0.0672 mm	1.7243e − 5 = 0.0116 mm
Equivalent (von Mises) stress	73.591 MPa	34.421 MPa
Max. shear stress	37.363 MPa	19.105 MPa

Table 78.8 Frequency response of modified and current FH

Frequency responses at different surfaces	Modified FH max. amplitude (MPa)	Current FH max. amplitude (MPa)
1	114.9	127.91
2	216.1	184.94
3	28.1	37.25
4	4.5	5.97
5	0.29	0.44
6	0.26	0.24
7	0.59	0.28
8	0.59	0.44
9	0.2	12.07
10	0.38	43.24

78.4 Conclusions

Load and stress analysis of the existing Fibrizer hammer and the new modified Fibrizer hammer are performed. From the results, the stress concentration factor analysis revealed that there is high stress concentration factor on the fillet radius of the existing Fibrizer hammer and less at web thickness (h) that is the cause of the hammer to be broken at the tail part. It shows that the current Fibrizer hammer having design problem and stress concentration factor on the fillet radius of the modified designed Fibrizer hammer is reduced by half. The experimental result of hardness analysis at different locations reveals that the average Vickers hardness of existing Fibrizer is 140.76 VHN. From this experimental result, one can understand that the material property of the Fibrizer hammer is not uniform throughout the hammer. Total deformation, directional deformation, equivalent von Mises stress, max. shear stress and harmonic response analysis of current, and modified Fibrizer hammer were analyzed and compared using FEM software. The result confirmed that the modified Fibrizer hammer having less total deformation, directional deformation, equivalent von Mises stress, max. shear stress, and harmonic response analysis than the current Fibrizer hammer. Using the modified Fibrizer hammer will solve Wonji Shoa Sugar Factory Fibrizer hammer breakage.

References

1. Anderson, J.G., Camuglia, S.I.J., Trapp, N.: Enhancing the life of shredder hammer tungsten carbide tips. In: Proceedings of the Australian Society of Sugar Cane Technologists, pp. 508–518. Scribe Consulting (2005)
2. Anderson, T.L.: Fracture Mechanics: Fundamentals and Applications. CRC press (2005)
3. Hugot, E.: Handbook of Cane Sugar Engineering, 3rd edn. (1986–2014)
4. Rein, P.: Cane Sugar Engineering, 2nd ed. Verlag Dr. Albert Bartens KG (2016)
5. Podany, P., Džugan, J., Vacik, J.: Construction of hammer for sugarcane shredder. In: Advanced Materials Research, vol. 811, pp. 308–313. Trans Tech Publications (2013)
6. Ransing, K.A., Gajjal, D.S., Saidpatil, V.V.: Redesign and optimization of hub profile of Fibrizer in sugar cane industry. Int. Eng. Res. J. (IERJ), Special Issue **2**, 5790–5793 (2015a)
7. Ransing, K.A., Gajjal, D.S., Saidpatil, V.V.: Review of design and optimization of Fibrizer in sugar cane industry. Int. J. Innov. Sci. Eng. Technol. **2**(4) (2015b)
8. Nikam, P.P., Nimbalkar, A.A., Thorat, A.S., Nanaware, A.A., Taware, O.H.: Design modification and analysis of sugar cane leveler at sugar cane processing plant. Imperial J. Interdisc. Res. **3**(4) (2017)
9. Norton, R.L.: Machine Design: An Integrated Approach. Prentice-Hall Inc., NJ, USA (1996)

Chapter 79
FEA Validation of Experimental Results of First Ply Failure of Composite Structure

Ishan Sonawane and A. Muthuraja

Abstract Laminated composite has gained popularity (despite of their higher cost) in high-performance products that need to have high corrosion resistance, good thermal stability, excellent fatigue resistance, lightweight and high specific strength to take harsh loading conditions such as aerospace components. Based on the literature survey, it is observed that Tsai–Wu criterion is widely used because of its accuracy. In the current study, finite element analysis is being carried out by using MSC Nastran and MSC Patran to identify the critical ply location and strength of first ply failure within laminated composite material plates. Tsai–Wu failure criterion is used to predict the maximum failure indices on each ply of laminated composite material. The results obtained from FEA are validated by comparing with the experimental result.

Keywords Laminates · Composite materials · First ply failure · Finite element analysis

79.1 Introduction

Laminated composite materials have recently gained popularity for different applications in the construction of mechanical, aerospace, marine and automotive structures. In general, these structures require high reliability. Within the laminated structure, plies are oriented in a different direction so as to make the composite structure tailor able when multi-directional loads are applied. This makes the laminated structure highly anisotropic and structurally inhomogeneous. So for reliability conformance, we need to predict accurately that maximum load a ply can sustain before failure. During the design phase of aerospace composite structures,

I. Sonawane · A. Muthuraja (✉)
Department of Mechanical Engineering, Sandip University, 422213 Nashik, India
e-mail: a.muthuraja2010@gmail.com

I. Sonawane
e-mail: s.ishan.r@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_79

first ply failure is considered as critical phenomena in composite failure. However, after the first ply failure other ply will carry the loads till the complete failure of composites happens. First ply failure analysis of laminated composite has drawn close attention in recent years. Turvey [1] used analytical methods to study the linear and nonlinear first ply failure loads of simply supported symmetrically and anti-symmetrically laminated composite plates based on the classical lamination theory. Tsai–Hill failure criterion has been used to facilitate an initial flexural failure analysis [2, 3]. Initial flexural failure was analyzed under uniform or non-uniform loading [2]. It was obtained that the initial failure loads decreased with corresponding plate center deflections as plate patch size increases during loading at Z-direction. Tsai–Hill failure criterion was used for angle ply laminates of GFRP and CFRP [3]. It was concluded that the initial failure loads decrease rapidly as theta increases, and the simply supported GFRP strips were obtained to be insensitive to the number of constituent laminas under initial failure loads. Turvey [4] obtained that the initial failure loads were direction-dependent, and simply supported strips were shown superior and increase with increasing number of laminas. Kam and Jan [5] formulated a finite element model for a thick laminated composite plate on the basis of layer wise linear displacement theory. Finite element results were compared between experimental and analysis results for displacements and stresses. Reddy and Pandey [6] developed a finite element computational procedure for the first ply failure analysis of laminated composite based on the first-order shear deformation theory and for a tensor polynomial failure criterion like maximum stress, maximum strain, Tsai–Hill, Tsai–Wu and Hoffman. Kam et al. [7] used a nonlinear finite element method to study the deformation and first ply failure of thin laminated composite plates. This study was based on the principle of minimum total

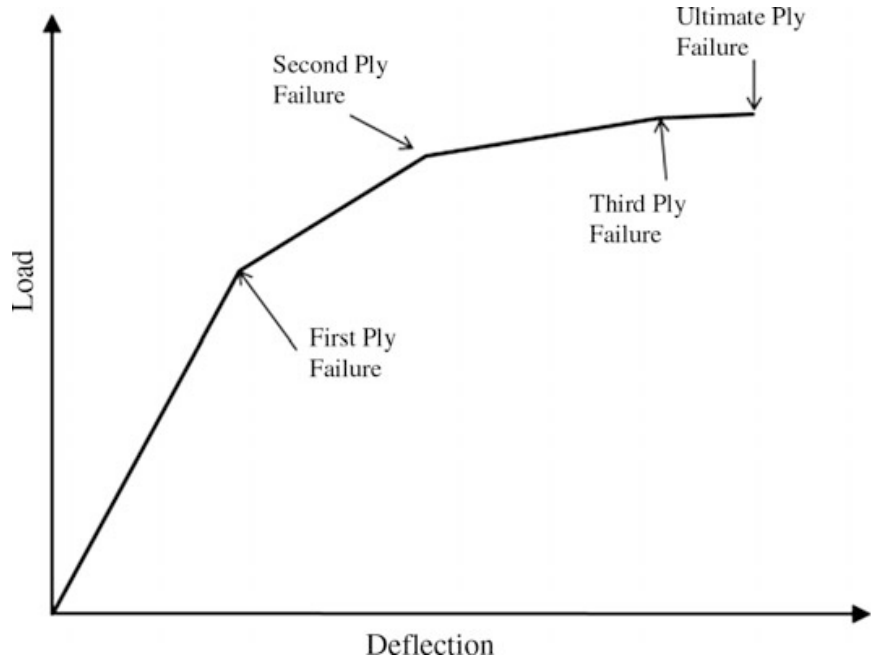


Fig. 79.1 Progressive ply failure

potential energy and von Karman–Mindlin plate theory. Reddy and Reddy [8] developed a finite element computational procedure to find first ply failure loads of composite laminates subjected to in-plane and transverse loads. This analysis was carried out for linear and nonlinear loading conditions, based on the first-order shear deformation theory and other failure criteria.

In the current study, finite element analysis software MSC Patran/Nastran is used to analyze the composite failure. Patran is used as pre-/post-processing, and Nastran is used as a solver. Nastran program was originally developed for NASA (Fig. 79.1).

79.2 Finite Element Method

79.2.1 Finite Element Modeling

In this paper, the finite element method is used to determine the first ply failure strength of laminated composite plate of 100 × 100 mm dimensions. The laminated composite plate is made with equal ply thickness of 0.121 mm and different orientations. Stacking sequence of $\left[45_3^0/0_3^0/-45_3^0/90_3^0\right]_S$ is used to construct the composite laminate. The lamina (graphite/epoxy) properties are considered from the experiments conducted [9] (Table 79.1).

The plate is modeled with 2D elements as the thickness is very less in comparison with other two dimensions. Also, the 2D elements have the ability to capture the bending stiffness. Elements are created on the mid surface and thickness is assigned. Vertically downward concentrated loads of 500, 1000 and 1335 N are applied independently at the center of the plate. It is observed from the result that the first ply failure occurs at a load of 1335 N. Thus, in this paper, a load of 1335 N

Table 79.1 Composite material properties

Material properties (linear elastic)	
Elastic modulus 1 (E11)	139,400 MPa
Elastic modulus 2 (E22)	7650 MPa
Shear modulus 12 and 13 (G12 and G13)	4350 MPa
Shear modulus 23 (G23)	1020 MPa
Poisson ratio (ϑ_{12})	0.29
Strength properties	
Tension stress limit (X_T)	1537.2 MPa
Tension stress limit (Y_T)	42.7 MPa
Compressive stress Limit (X_C)	1722.4 MPa
Compressive stress limit (Y_C)	213.9 MPa
Shear stress limit (S and T)	102.4 MPa
Interaction term	−3.215E − 6
Bonding shear stress limit	5000 MPa

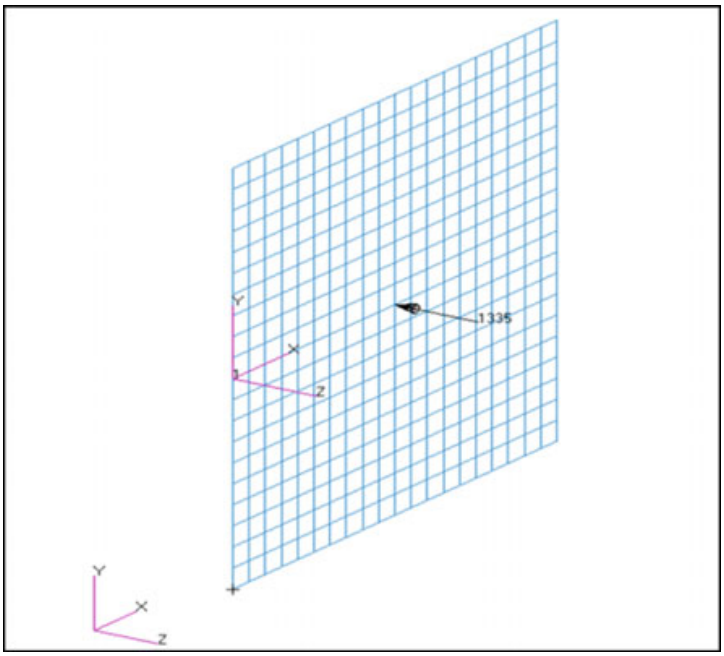


Fig. 79.2 Applied load

is applied as shown in Fig. 79.2, and all results are extracted at 1335 N applied load.

In accordance with experimental method carried out by [5], the model is constrained in all six directions at all four edges as shown in Fig. 79.3.

79.2.2 Results and Discussion

The strength of the composite plate is determined by using Tsai–Wu failure criteria. Figure 79.4 shows that the maximum failure index for non-layered plate occurs at the center of plate. Figure 79.5 shows that Tsai–Wu failure index of 1 is observed at the bottom ply location at the center location. During experimental tests, T. Y. Kam and F. M. Lai found that the first ply failure occurs at the bottom location with applied load of 1320 N. The FEM result is compared with the experimental result [5] and found that there is 1% difference of failure load between the experimental and FEA results. In both cases, first ply failure occurs at the bottom location.

79.3 Theoretical Approach

The constitutive model is considered in the linearly elastic region, and Tsai–Wu criterion is used to predict the first ply failure of the laminated plate using the maximum failure index. A number of failure criteria for composite laminate proposed over the years are maximum stress, maximum strain, Hill–Tsai, Tsai–Wu. Stress interaction between the laminates is ignored for maximum stress and

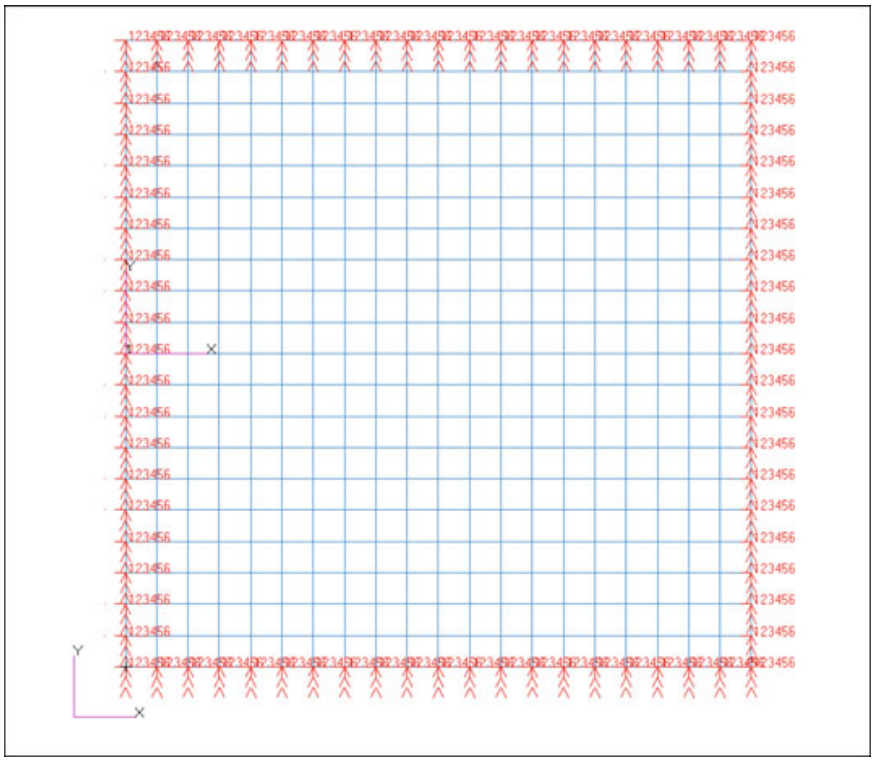


Fig. 79.3 Boundary conditions

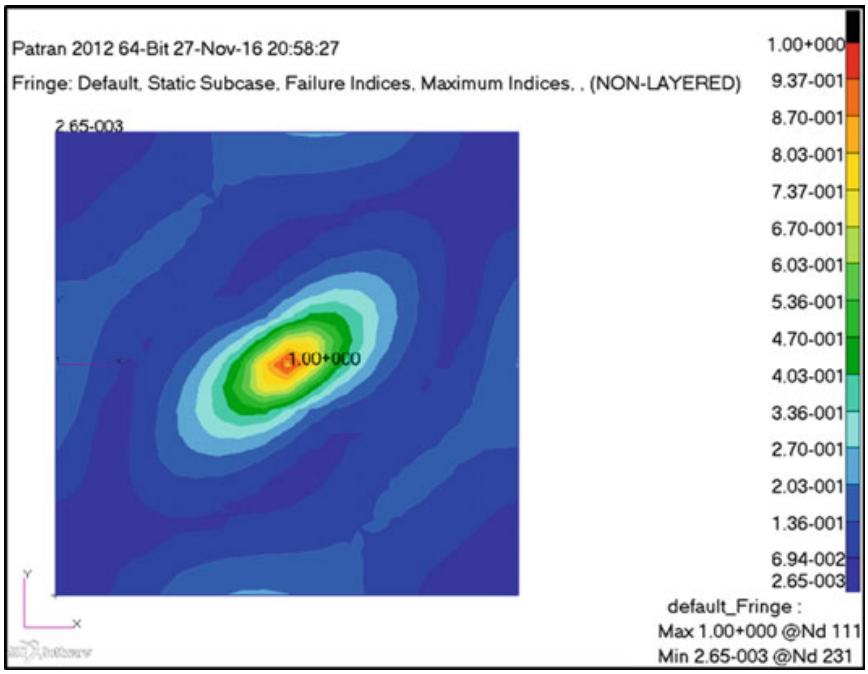


Fig. 79.4 Maximum failure index

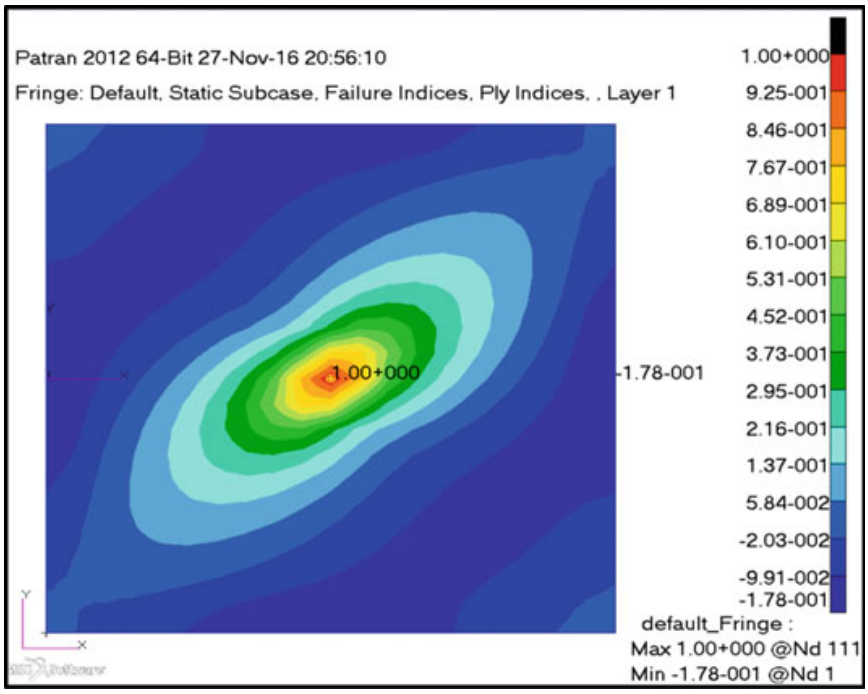


Fig. 79.5 First ply failure

maximum strain criteria; whereas, Hill–Tsai and Tsai–Wu criteria include full stress interaction.

The Tsai–Wu failure criterion is a macro-mechanical failure theory which is prominently used for anisotropic composite materials which have different strengths in tension and compression. Tsai–Wu uses strength tensors (F_i and F_{ij}) established through experiment. Theoretical predictions of Tsai–Wu criterion match very well compared with the experimental data.

Tsai–Wu theory predicts that the failure occurs when below criteria is satisfied

$$F_i\sigma_i + F_{ij}\sigma_i\sigma_j \geq 1 \text{ where } i,j = 1,2\ldots6$$

(Summation is done for repeated indices)

$$F_1 = \frac{1}{X_T} - \frac{1}{X_C} \quad F_2 = \frac{1}{Y_T} - \frac{1}{Y_C} \quad F_3 = \frac{1}{Z_T} - \frac{1}{Z_C} \quad F_{11} = \frac{1}{X_T X_C} \quad F_{22} = \frac{1}{Y_T Y_C}$$

$$F_{33} = \frac{1}{Z_T Z_C} \quad F_{44} = \frac{1}{R^2} \quad F_{55} = \frac{1}{S^2} \quad F_{66} = \frac{1}{S^2}$$

$$F_{12} = -\frac{1}{2\sqrt{X_T X_C Y_T Y_C}} \quad F_{13} = -\frac{1}{2\sqrt{X_T X_C Z_T Z_C}} \quad F_{23} = -\frac{1}{2\sqrt{Y_T Y_C Z_T Z_C}}$$

X_T, Y_T, Z_T (X_C, Y_C, Z_C) Lamina normal tensile (compressive) strength in 1, 2 and 3 directions.

R, S	Shear strengths in the 23 and 12 planes, respectively.
σ_1, σ_2, σ_3	Normal stress components; σ_4, σ_5, σ_6: shear stress components

For the analysis, shear strengths in the 12 and 13 planes are assumed to be same and $Y_T = Z_T$.

79.4 Conclusions

First ply failure strengths of laminated composite plates with different layups were determined using FEA, and the results were validated by using experimental data. The Tsai–Wu failure criterion is used to calculate the failure index.

1. It is observed that there is 1% difference of failure load between the experimental and FEA results. Bottom most layer shows the maximum failure index under 1335 N loading condition comparatively other layers. It shows the accuracy of first ply failure prediction from both the methods.
2. In composite laminate, the stacking sequence plays an important role in improving the strength with different loading conditions. During the design phase, stacking sequence changed frequently so as to obtain feasible design. By experimental methods, we can evaluate the first ply failure strength and location. But experimental setup and procedure results are in high cost. During design phase, it is difficult to go for experimental results every time, which results in

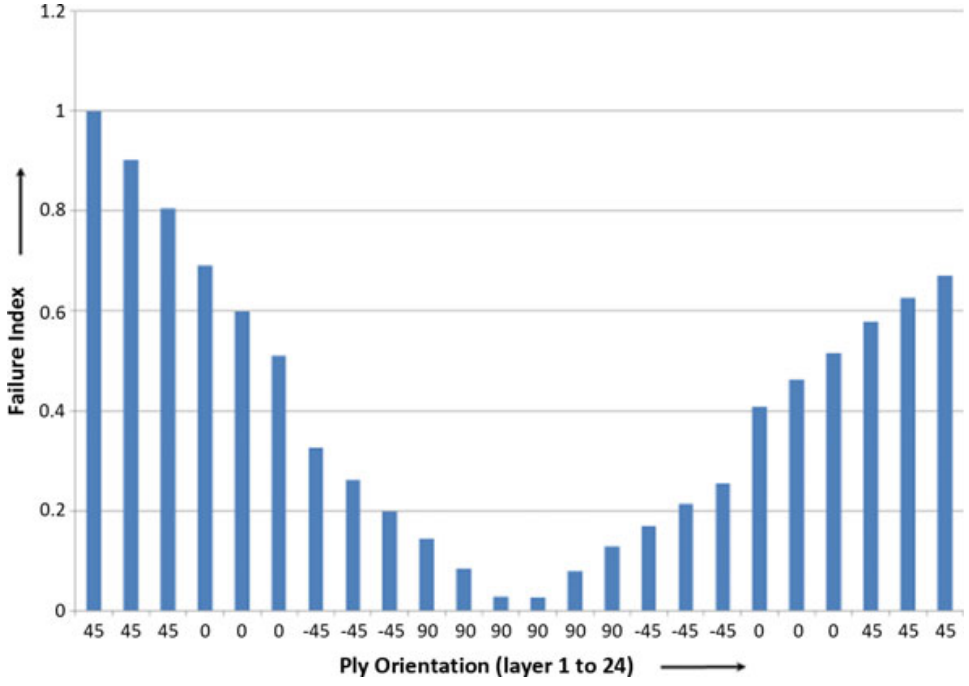


Fig. 79.6 Failure indices of different ply orientation

high cost. By FEA, we can predict the first ply failure strength and location. Also, we can predict the strength of other critical plies as shown in Fig. 79.6.

3. After the first ply failure, the composite will not fail suddenly, and the loads are transferred to other plies. It is important to detect the remaining strength in the structure till final failure. In many cases, composite structures are connected with metallic joints. Composites are weaker at the joint location. So, further study needs to be done on the first and progressive ply failure analysis at composite joint location.

Acknowledgements First and foremost, we praise and thank our almighty God whose blessings have bestowed on us the will power and confidence to carry out this paper. We would also like to thank all our friends and our parents who have prayed and helped us during the project work.

References

1. Turvey, G.J.: An initial flexural failure analysis of symmetrically laminated cross-ply rectangular plates. Int. J. Solids Struct. **16**, 451–463 (1980)
2. Turvey, G.J.: Initial flexural failure of square, simply supported, angle-ply plates. Fibre Sci. Technol. **15**, 47–63 (1981)
3. Turvey, G.J.: Flexural failure analysis of angle-ply laminates of GFRP and CFRP. J. Strain Anal. Eng. Des. **15**, 43–49 (1980)
4. Turvey, G.J.: A study of the onset of flexural failure in cross-ply laminated strips. Fibre Sci. Technol. **13**, 325–336 (1980)
5. Kam, T.Y., Jan, T.B.: First ply failure of laminated composite plates based on the layerwise linear displacement theory. Compos. Struct. **32**, 583–591 (1995)
6. Reddy, J.N., Pandey, A.K.: A first-ply failure analysis of composite laminates. Comput. Struct. **25**, 371–393 (1987)
7. Kam, T.Y., Chang, R.R., Sher, H.F., Chao, T.N.: Prediction of first-ply failure strengths of laminated composite plates using a finite element approach. Int. J. Solids Struct. **33**, 375–398 (1996)
8. Reddy, Y.S.N., Reddy, J.N.: Linear and nonlinear failure analysis of composite laminates with transverse shear. Compos. Sci. Technol. **44**, 227–255 (1992)
9. ASTM Standards and Literature References for Composite Material, 2nd ed. (1990)

Chapter 80
Triply Periodic Minimal Surface-Based Porous Scaffold Design and Analysis Subjected to Hard Tissue Reconstruction

Verma Rati, Nishant Singh, Sanjay Rai and Shekhar Kumta

Abstract Porous structures based on triply periodic minimal surfaces (TPMS) have occurred as appropriate candidates for scaffold design with high level of porosity and promising strength for bone replacement. This work describes a suitable procedure for design and modeling of 3D architecture of TPMS-based gyroid and primitive structures and identifying the optimal architecture for scaffold designing. Different models with varying porosity were reconstructed and analyzed to access the effective elastic moduli and compressive strength of each model. An optimal model that can be utilized in bone tissue replacement is identified eventually. Ti6Al4V which is considered as the best material for producing implants in conjunction with biocompatibility and strength is used in the study. The compression test performed by FEM revealed that the scaffold models with porosity level of 65 and 60% are best suited for cortical bone replacement and the model with 90% porosity can be used on the anatomical location, which are more inclined to cancellous bones.

Keywords Tissue engineering · Porous scaffolds · TPMS · Gyroid · Primitive · Finite element analysis

V. Rati · S. Rai (✉)
School of Biomedical Engineering, IIT BHU, Varanasi, India
e-mail: skrai.bme@iitbhu.ac.in

V. Rati
e-mail: rati.verma14@gmail.com

N. Singh
Department of Biomedical Engineering, NIT Raipur, Raipur, Chhattisgarh, India
e-mail: nksingh.bme@gmail.com

S. Kumta
Department of Orthopaedics and Traumatology, Prince of Wales Hospital, the Chinese University of Hong Kong, Shatin, New Territories, Hong Kong
e-mail: shekharkumta@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_80

80.1 Introduction

The field of tissue engineering aims to develop an efficient biological substitute in order to preserve, maintain, or enhance the function of tissue. Scaffolds play a significant part in deciding mechanical characteristics of desired tissue formed in it at a macroscopic or microscopic scale and also provides the vascularization needed by the newly forming tissues [1]. Scaffold must be biocompatible, bio-inert, bioactive, bioresorbable, and biodegradable so as to work efficiently to possess long term results [2]. Porous scaffold, therefore, provides necessary temporary atmosphere for adhesion, propagation, and differentiation of cell [3], together with the transport phenomena and degradation behavior [4]. Other advantages of these porous scaffolds include its flexibility in terms of mechanical properties, a large surface area that enables bio-functionalization and infection prevention and accelerates bone ingrowth as well as facilitates drug delivery in implants [5]. To establish a balance between the mechanical properties of scaffold and porosity is a challenge [1, 6]. Tissue regeneration efficiency is greatly affected by scaffold pore size [3].

Triply periodic minimal surface (TPMS) is an emerging area of interest for designing the scaffolds because of its unique properties like the ability to mimic the organic system and to build self-supportive structures for additive manufacturing [7, 8]. Another reason for TPMS-based scaffolds to draw the attention toward it is because of its capability to provide site-specific biological performance [9]. TPMS is defined in mathematical terms as surface curvatures which are infinite and periodic in 3D Euclidian space that owes to a surface capable of generating highly controllable and homogeneous scaffold design [10]. TPMS have implicit periodic surfaces in three independent Cartesian directions with mean curvature equal to zero [11]. These minimal surfaces play a very essential role in guiding chemical, cellular, and biochemical processes [10]. The TPMS-based porous structures have the potential to attain the lightweight bone models and medical requirements of bone tissue, thus providing a solution for the production of biomimetic bone in biomaterials and tissue engineering [12].

Observing the TPMS topology for Schoen's gyroid that perfectly suits the porous-solid structural configuration ideal for tissue engineering scaffolds [4, 13–16], this paper emphasizes on TPMS surface geometry and gyroid surface, particularly.

The present study aims to develop a porous scaffold based on TPMS geometry that possesses the mechanical properties (elastic modulus and compressive strength) proximate to bone. The study utilizes Schwarz's primitive (P) and Schoen's gyroid surface (G) structures to design bone scaffolds as their macroscopic structure is analogous to bone. Ti6Al4V (Grade 5) regarded as a biocompatible material is employed with its mechanical properties with scaffold porosities within the range of 60–90%. Unit cell length and pore size are varied with wall thickness kept constant at 102 μm for G and 133 μm for P, respectively, to achieve porosities within the desired range for both the structures (P and G). Finite element analysis was performed in ANSYS® software with static loading and boundary conditions for the

determination of compressive strength based on von-Mises criteria to study the deformation behavior and to estimate the effective elastic moduli and compressive yield strength.

80.2 Materials and Method

80.2.1 TPMS-Based Porous Scaffold-Primitive (P) and Gyroid (G) Surface Modeling

Both the primitive (P) as well as gyroid surface are designed from the basic parametric equations with implicit functions described in Eqs. (80.1) and (80.2).

$$\text{Primitive}\,(P)\text{: } \cos(x) + \cos(y) + \cos(z) = 0 \tag{80.1}$$

$$\text{Gyroid: } \sin(x) * \cos(y) + \sin(y) * \cos(z) + \sin(z) * \cos(x) = 0 \tag{80.2}$$

The unit cell of TPMS is designed with a MATLAB code, executed in the MATLAB® (v. R2016a) software. "MESHGRID" command along with "ISOSURFACE" command was used to extract the isosurface 3D data from the volume 3D data generated by "MESHGRID" command.

The unit cell was extended periodically in x, y, and z directions by modulating the Eqs. (80.1) and (80.2) as described in Eqs. (80.3) and (80.4) [17].

$$\text{Primitive}\,(P)\text{: } F_P = \left[\cos\left(\frac{2\pi Nx}{L}\right) + \cos\left(\frac{2\pi Ny}{L}\right) + \cos\left(\frac{2\pi Nz}{L}\right)\right] - P \tag{80.3}$$

$$\begin{aligned}\text{Gyroid: } F_G(x,y,z) = \Bigg[&\sin\left(\frac{2\pi Nx}{L}\right) * \cos\left(\frac{2\pi Ny}{L}\right) + \sin\left(\frac{2\pi Ny}{L}\right) * \cos\left(\frac{2\pi Nz}{L}\right) \\ &+ \sin\left(\frac{2\pi Nz}{L}\right) * \cos\left(\frac{2\pi Nx}{L}\right)\Bigg] - P\end{aligned} \tag{80.4}$$

where N defines the number of unit cells, L is length of the unit cell, and P is the strut thickness parameter that controls the volume fraction in a unit cell. For a scaffold with 3 × 3 × 3 unit cells (Fig. 80.1), value of N was fixed at $N = 3$, unit cell length L was varied due to which the pore diameter and total scaffold length were altered accordingly as listed in Table 80.1. Seven models for each surface (primitive: P_1–P_7 and gyroid: G_1–G_7) type were designed. As a result, a total of fourteen models were designed. P is kept constant at $P = 0$ to fix the volume fraction at 50%. "STLWRITE" command was added to the MATLAB code to export the geometry as stl file which can be imported to other software for finite element analysis (FEA).

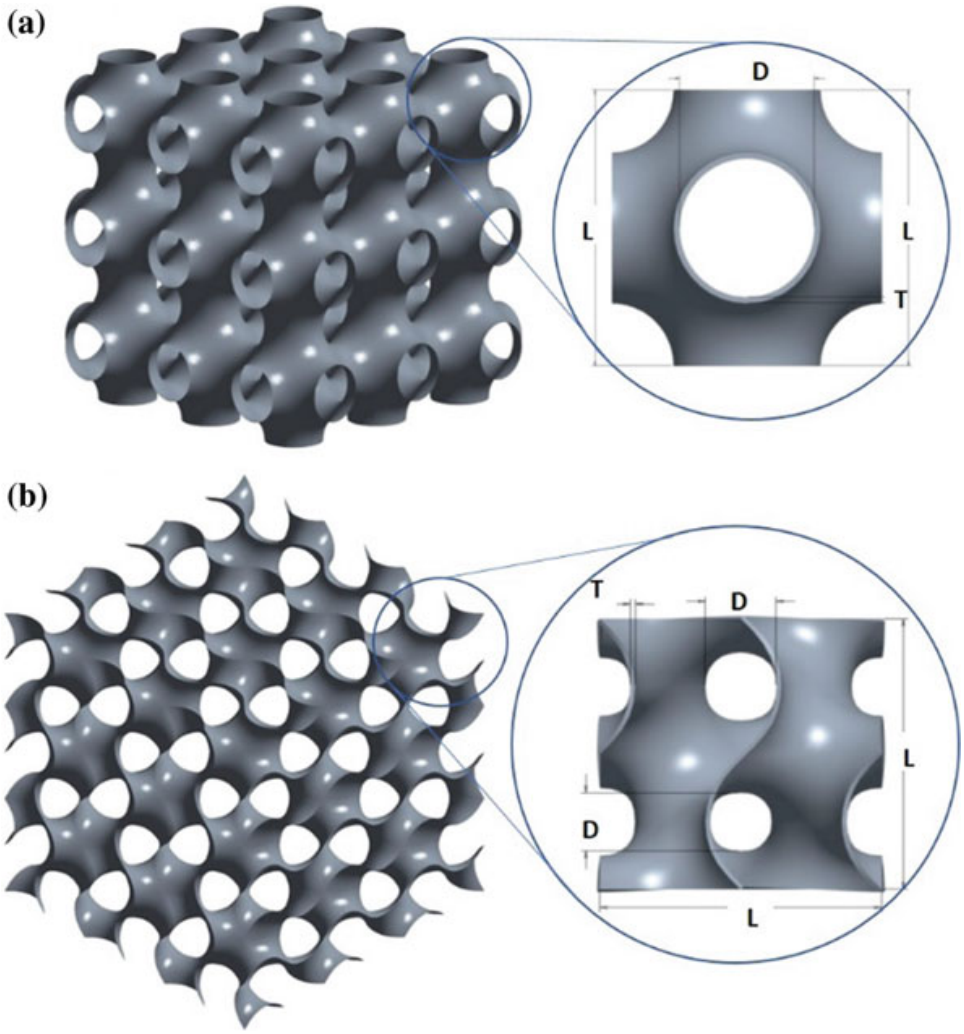


Fig. 80.1 Designed TPMS-based porous surface models. Isometric view with 3 × 3 × 3 unit cells for **a** primitive surface, **b** gyroid surface (Zoom in view inside the circle presents the unit cell length (L), pore diameter (D) and wall thickness (T) of each model)

Table 80.1 Parametric values of TPMS-based primitive (P_1–P_7) and gyroid (G_1–G_7) structures

	P_1, G_1	P_2, G_2	P_3, G_3	P_4, G_4	P_5, G_5	P_6, G_6	P_7, G_7
Porosity (%)	90	85	80	75	70	65	60
Unit cell length (L) (μm)	3141	2104	1570	1256	1055	911	791
Pore size (D) (μm)	1654	1074	776	600	488	407	340

80.2.2 Design and Formulation of Finite Element Model for FEA

The surface models of P surface and gyroid were imported in 3-matic software as stl files and fixed for any errors in the surface. Wall thickness was developed and the model was exported to ANSYS software for FEA after performing surface and volume meshing. The complete procedure is explained in the following subsections.

Developing wall thickness and preparing Finite Element Models

Stl files generated through MATLAB software were imported in 3-matic software for assigning wall thickness to the surface and carrying out surface and volume meshing. The surface of the models of gyroid and P were thickened with a uniform offset of 102 µm and 133 µm, respectively (Fig. 80.1). Since, offsetting the surface disturbed the surface geometry (mesh) of gyroid structures, surface meshing was performed with specific 3-matic tools to improve the quality of surface mesh. Additionally, to reduce the density of the triangles and improve the quality of mesh, quality preserving reduce triangles tool followed with uniform meshing with a target triangle edge length of 0.2 was performed in a manner that element quality is preserved. The shape-quality threshold and maximum geometrical error values were fixed to 0.5 and 0.05, respectively [7].

The generated models were checked for errors and fixed accordingly. Consequently, the number of inverted normal, noise shells, bad contours, planar holes, overlapping triangles, and intersecting triangles present in each model were fixed manually to obtain a single shell with no errors. Volume meshing was done to solidify the surface models with voxel size of Tet4 elements. The volumetric models of all scaffolds generated in 3-matic software were imported to Ansys software as stl files.

Assigning Material properties, Boundary conditions, and Load parameters

Ti6Al4V (Grade 5) material considered suitable in terms of a compatible biomaterial for implants is used for FEA [18]. The elastic isotropic material properties of Ti6Al4V adopted in the study include Young's modulus of 114 GPa, Poisson's ratio of 0.33, density of 4.43 g/cc, and compressive yield strength of 1070 MPa [19].

Remeshing with 4-node SHELL181 hexahedral element was performed for each model (Fig. 80.2). Two rigid plates are fixed at the top and bottom of each scaffold to assign the boundary conditions and loading. The bottom plate is constrained in all directions while a static compressive load of 75 N is applied in negative Y direction at the top plate (Fig. 80.3).

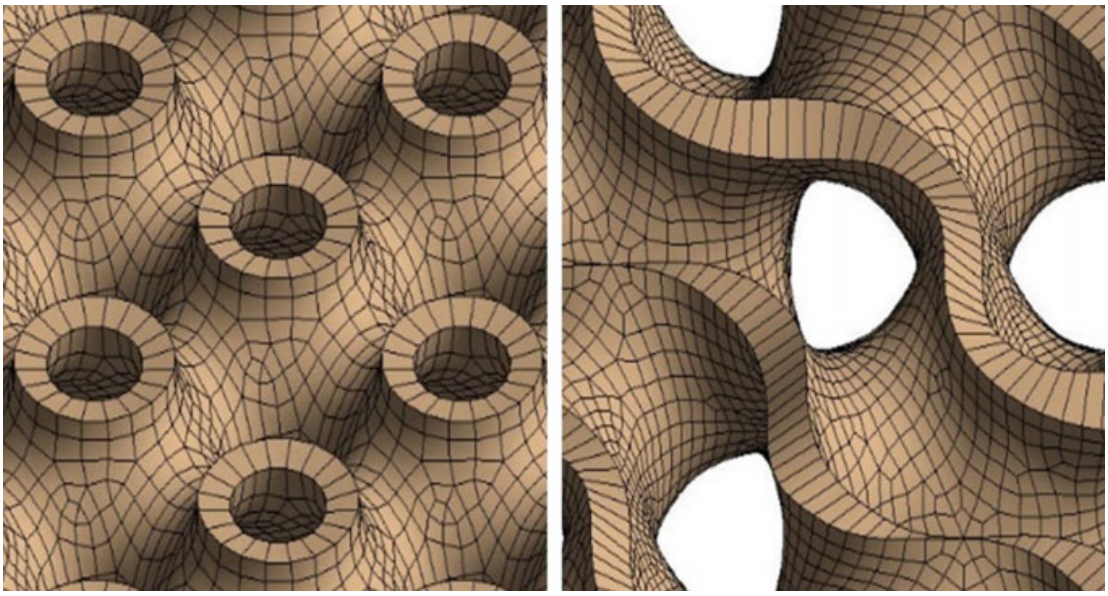


Fig. 80.2 Remeshed models with four-node hexahedral elements for primitive (left) and gyroid (right) surfaces

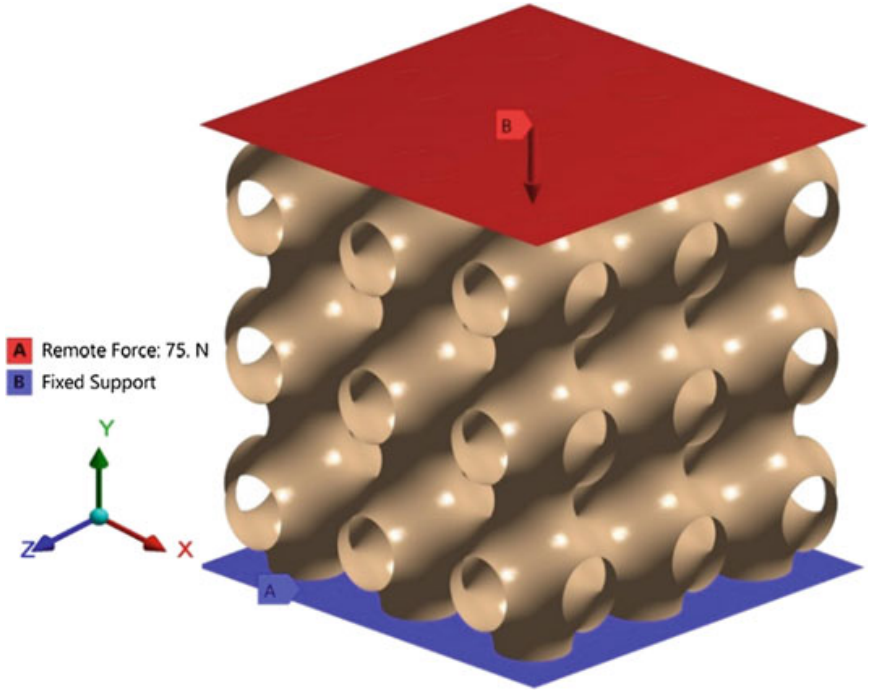


Fig. 80.3 Application of boundary conditions and load parameters. Two rigid plates, red and blue are fixed at top and bottom of designed scaffold structure and compressive load of 75 N is applied in negative Y direction

Table 80.2 Porous volume (V_P) and solid volume (V_S) of primitive (P_1–P_7) and gyroid (G_1–G_7) models

	P_1	P_2	P_3	P_4	P_5	P_6	P_7
(V_P) (μm^3)	2897.9	1287.9	763.62	495.08	348.26	260.81	195.89
(V_S) (μm^3)	28333.5	8395.1	3832.3	2000.8	1180.2	764.9	497.9
	G_1	G_2	G_3	G_4	G_5	G_6	G_7
(V_P) (μm^3)	3123.7	1402.2	780.92	499.79	352.65	262.7	198.37
(V_S) (μm^3)	31006490	9325.9	3875.8	1984.2	1176.2	756.207	496.191

80.3 Porosity

Porosity was calculated from the relation expressed in Eq. (80.3) [20].

$$N = 1 - \left(\frac{V_P}{V_S}\right) \tag{3}$$

where V_S is the solid volume and V_P is the volume of porous construct. The solid volume constitutes the total volume of cubic construct. Solid volume (V_S) and porous volume (V_P) were calculated all primitive (P_1–P_7) and gyroid surface (G_1–G_7) models with different porosity levels (60–90%) as listed in Table 80.2.

80.4 Results

The effective elastic modulus as well as compressive strength was calculated for each model. Effective elastic moduli were determined from Hooke's law ($\sigma = \varepsilon E$) and compressive strength was obtained from von-Mises criteria. The performance characteristics of scaffold structure with 3 × 3 × 3 unit cells were observed under static compressive load of 75 N. Total von-Mises stress distribution and maximum deformation were measured for scaffolds with varying length and porosity (Fig. 80.4). Compressive strength for each model was determined by the load of 75 N that caused the maximum stress of 909.24 and 1102.5 MPa in P surface model and gyroid surface model, respectively, with 70% porosity (Fig. 80.4a, b). Maximum deformation of 28.542 and 22.875 µm was obtained for P surface model and gyroid surface model, respectively, with 70% porosity (Fig. 80.4c, d). Variation in elastic modulus and compressive strength with respect to porosity has been plotted (Fig. 80.5).

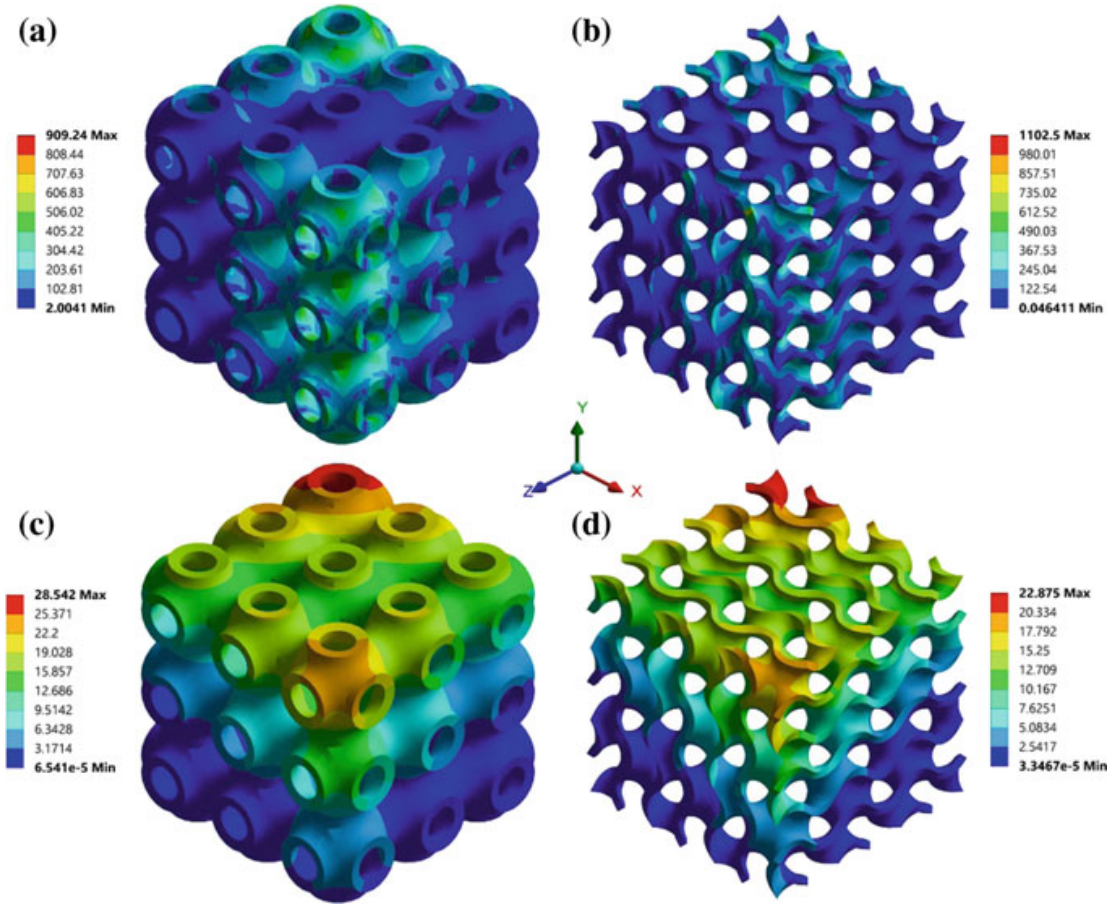


Fig. 80.4 Mechanical properties of scaffolds with 70% porosity obtained from FEA simulation. Von-Mises stress distribution (MPa) for **a** primitive structure: P_5, **b** gyroid structure: G_5 and total deformation (μm) for **c** primitive structure: P_5, **d** gyroid structure: G_5

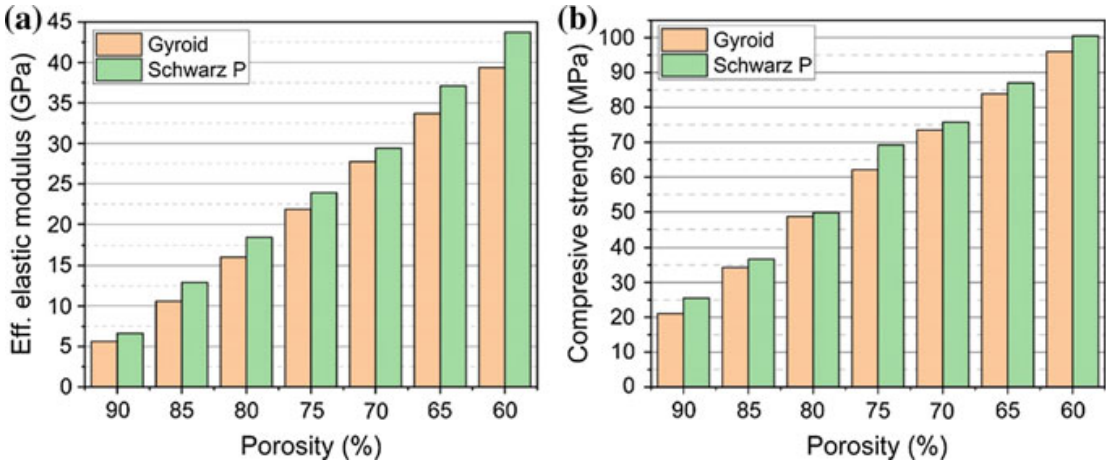


Fig. 80.5 Graphical representation of mechanical properties plotted against porosity. Plot of **a** effective elastic moduli and **b** compressive strength with respect to porosity for Schwarz *P* surface and gyroid surface

80.5 Discussion

Numerous studies have been carried out the design and manufacturing of TPMS-based porous scaffold materials because of the challenges imposed in this field [9, 20–22]. In order to obtain a scaffold that perfectly mimics mechanical properties similar to bone, the effective elastic moduli in addition to compressive strength of the scaffolds are obtained proximate to the bone [23, 24]. Elastic modulus for cortical bone has been noted in the range of 3–30 GPa, whereas for cancellous bone, it is in the range of 0.02–0.2 GPa. Also, the compressive strength for cortical and cancellous bone is reported as 100–230 MPa and 2–12 MPa, respectively [25]. Although, the mechanical characteristics of bone with compressive load show a discrepancy with according to various factors (age, gender, and anatomical locations) within a human body [26–28]. Therefore, a scaffold is designed in such a way so that a balance between the porosity level and mechanical properties can be maintained [29–31]. This prevents the bone from stress shielding as well as reabsorption due to bone–scaffold interaction [32, 33]. TPMS-based porous scaffolds also enhance the bone ingrowth and weight-bearing capacity due to their exceptional porous architecture that are lightweight in comparison with the conventional solid scaffolds [12, 33–35].

Figure 80.5 characterizes the mechanical properties comprising the effective elastic moduli along with stress distribution for primitive and gyroid surface models with varying porosity under a compressive load of 75 N in inverse Y direction. Effective elastic modulus and compressive strength for both structures (primitive and gyroid) was maximum for minimum porosity of 60% and followed a pattern to decrease with increasing porosity.

Observing the characteristics of effective elastic moduli (Fig. 80.5a), a maximum effective elastic modulus of 43.73 MPa for *P* surface and 39.36 MPa for gyroid surface was obtained in case of 60% porosity. However, a minimum value of 6.625 MPa for *P* surface and 5.628496 MPa for gyroid surface was obtained in case of maximum porosity of 90%.

Likewise, the compressive strength (Fig. 80.5b) was also observed to be maximum in case of 60% porosity for both primitive surface (100.519 MPa) and gyroid surface (96.029 MPa). A minimum value of von-Mises stress was obtained in case of 90% porosity for primitive surface (25.52178 MPa) and gyroid surface (21.06333 MPa). However, it can be noted that, for 80% porosity, the compressive strength was approximately equal for *P* surface (49.78 MPa) and gyroid surface (48.68 MPa).

As a result, it can be concluded that effective elastic moduli of gyroid surface models were lower than *P* surface models for all the porosity levels. Similar to the pattern of effective elastic moduli, compressive strength was also observed to be low for gyroid surface models as compared to *P* surface models for all the porosity levels ranging from 60 to 90%.

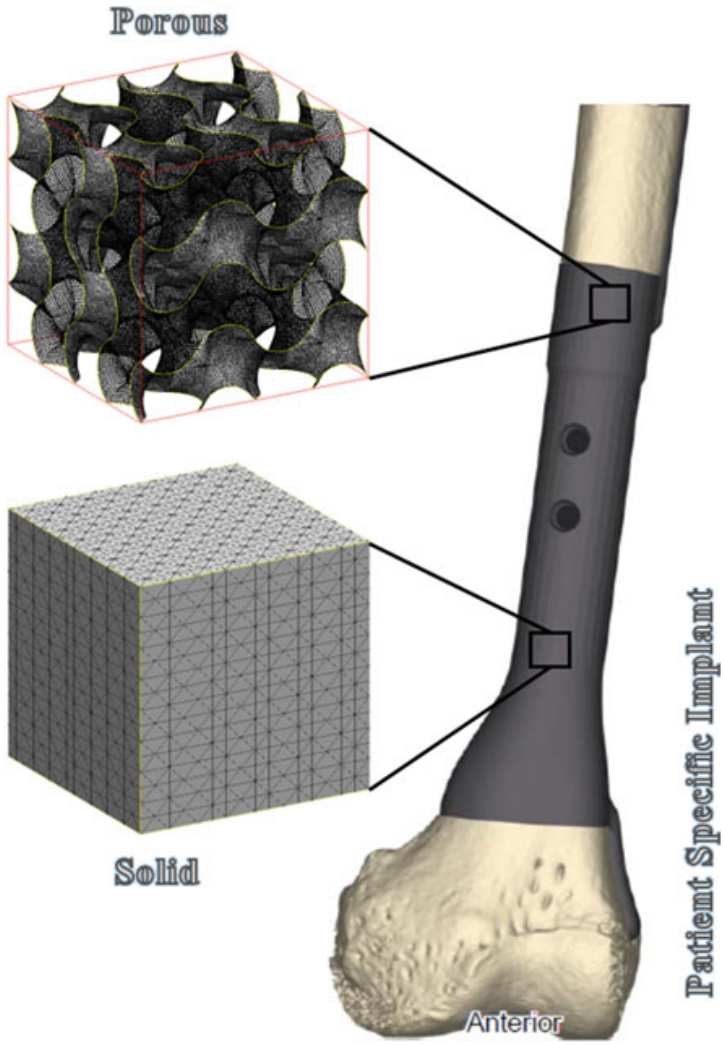


Fig. 80.6 Illustration of mapping of developed porous scaffold to treat large segmental bone defect

The conclusions of this pre-clinical analysis study elucidated the proficiency of design criterion for selecting an efficient scaffold structure in order to obtain the preferred value of the mechanical properties intended for patient-specific implant manufacturing that can be used for treating large segmental bone defects [36], where high porosity scaffolds are greatly needed (Fig. 80.6).

The overview of outcomes achieved in accordance with the mechanical properties of TPMS-based porous scaffold recommended their applicability at different anatomical locations according to the porosity level and suggested them to be appropriate structures for applications in bridging the bone gaps. However, the result is effective merely for desired scaffold structure requirements. The working fundamentals can be followed for other porous scaffold structures and biomechanical bone properties.

References

1. Almeida, H.A.: Design of tissue engineering scaffolds based on hyperbolic surfaces: structural numerical evaluation. Med. Eng. Phys. **36**(8), 1033–1040 (2014)
2. Zhang, X.Y.: Additively manufactured scaffolds for bone tissue engineering and the prediction of their mechanical behavior: a review. Materials **10**(1), 50 (2017)
3. Yang, N.: Mathematically defined gradient porous materials. Mater. Lett. **173**, 136–140 (2016)
4. Melchels, F.P.: Mathematically defined tissue engineering scaffold architectures prepared by stereolithography. Biomaterials **31**(27), 6909–6916 (2010)
5. Bobbert, F.S.L.: Additively manufactured metallic porous biomaterials based on minimal surfaces: a unique combination of topological, mechanical, and mass transport properties. Acta Biomater. **53**, 572–584 (2017)
6. Mehboob, H.: Finite element modelling and characterization of 3D cellular microstructures for the design of a cementless biomimetic porous hip stem. Mater. Des. **149**, 101–112 (2018)
7. Mohammed, M.I.: Design of three-dimensional, triply periodic unit cell scaffold structures for additive manufacturing. J. Mech. Des. **140**(7), 071701 (2018)
8. Mohammed, M.I.: Design and fabrication considerations for three dimensional scaffold structures. KnE Eng. **2**(2), 120–126 (2017)
9. Zheng, X.: Minimal surface designs for porous materials: from microstructures to mechanical properties. J. Mater. Sci. **53**(14), 10194–10208 (2018)
10. Blanquer, S.B.: Surface curvature in triply-periodic minimal surface architectures as a distinct design parameter in preparing advanced tissue engineering scaffolds. Biofabrication **9**(2), 025001 (2017)
11. Feng, J.: Porous scaffold design by solid T-splines and triply periodic minimal surfaces. Comput. Methods Appl. Mech. Eng. **336**, 333–352 (2018)
12. Wang, S.: Lightweight of artificial bone models utilizing porous structures and 3D printing. Int. J. Performability Eng. **13**(5), 633 (2017)
13. Olivares, A.L.: Finite element study of scaffold architecture design and culture conditions for tissue engineering. Biomaterials **30**(30), 6142–6149 (2009)
14. Melchels, F.P.: A poly (D, L-lactide) resin for the preparation of tissue engineering scaffolds by stereolithography. Biomaterials **30**(23–24), 3801–3809 (2009)
15. Gauvin, R.: Microfabrication of complex porous tissue engineering scaffolds using 3D projection stereolithography. Biomaterials **33**(15), 3824–3834 (2012)
16. Giannitelli, S.M.: Current trends in the design of scaffolds for computer-aided tissue engineering. Acta Biomater. **10**(2), 580–594 (2014)
17. Maskery, I.: Insights into the mechanical properties of several triply periodic minimal surface lattice structures made by polymer additive manufacturing. Polymer **152**, 62–71 (2018)
18. Li, X.: Fabrication and characterization of porous Ti6Al4V parts for biomedical applications using electron beam melting process. Mater. Lett. **63**(3–4), 403–405 (2009)
19. Welsch, G., Boyer, R., Collings, E.W. (eds.): Materials properties handbook: titanium alloys. ASM international (1993)
20. Walker, J.M.: Design and mechanical characterization of solid and highly porous 3D printed poly (propylene fumarate) scaffolds. Prog. Addit. Manuf. **2**(1–2), 99–108 (2017)
21. Dias, M.R.: Optimization of scaffold design for bone tissue engineering: a computational and experimental study. Med. Eng. Phys. **36**(4), 448–457 (2014)
22. Murr, L.E.: Next-generation biomedical implants using additive manufacturing of complex, cellular and functional mesh arrays. Philos. Trans. R. Soc. Lond. A: Math. Phys. Eng. Sci. **368** (1917), 1999–2032 (2010)
23. Al-Ketan, O.: Mechanical properties of a new type of architected interpenetrating phase composite materials. Adv. Mater. Technol. **2**(2), 1600235 (2017)
24. Montazerian, H.: Porous scaffold internal architecture design based on minimal surfaces: a compromise between permeability and elastic properties. Mater. Des. **126**, 98–114 (2017)

25. Wang, X.: Topological design and additive manufacturing of porous metals for bone scaffolds and orthopaedic implants: a review. Biomaterials **83**, 127–141 (2016)
26. Choi, K.: The elastic moduli of human subchondral, trabecular, and cortical bone tissue and the size-dependency of cortical bone modulus. J. Biomech. **23**(11), 1103–1113 (1990)
27. Rho, J.Y.: Mechanical properties and the hierarchical structure of bone. Med. Eng. Phys. **20**(2), 92–102 (1998)
28. Rho, J.Y.: Young's modulus of trabecular and cortical bone material: ultrasonic and microtensile measurements. J. Biomech. **26**(2), 111–119 (1993)
29. Li, D.: Optimal design and modeling of gyroid-based functionally graded cellular structures for additive manufacturing. Comput. Aided Des. **104**, 87–99 (2018)
30. Ataee, A.: Anisotropic Ti-6Al-4V gyroid scaffolds manufactured by electron beam melting (EBM) for bone implant applications. Mater. Des. **137**, 345–354 (2018)
31. Kadkhodapour, J.: Failure mechanisms of additively manufactured porous biomaterials: effects of porosity and type of unit cell. J. Mech. Behav. Biomed. Mater. **50**, 180–191 (2015)
32. Arabnejad, S.: Fully porous 3D printed titanium femoral stem to reduce stress-shielding following total hip arthroplasty. J. Orthop. Res. **35**(8), 1774–1783 (2017)
33. Arabnejad, S.: High-strength porous biomaterials for bone replacement: a strategy to assess the interplay between cell morphology, mechanical properties, bone ingrowth and manufacturing constraints. Acta Biomater. **30**, 345–356 (2016)
34. Kadkhodapour, J.: Investigating internal architecture effect in plastic deformation and failure for TPMS-based scaffolds using simulation methods and experimental procedure. Mater. Sci. Eng., C **43**, 587–597 (2014)
35. Taniguchi, N.: Effect of pore size on bone ingrowth into porous titanium implants fabricated by additive manufacturing: an in vivo experiment. Mater. Sci. Eng., C **59**, 690–701 (2016)
36. Reichert, J.C.: The challenge of establishing preclinical models for segmental bone defect research. Biomaterials **30**(12):2149–2163 (2009)

Chapter 81
Design of Medical Device Product Using Multiphysics Simulations

Rajesh Kapuganti

Abstract The primary function of kidneys is to remove waste and other impurities from the blood of a human being on a daily basis. When the kidneys stop functioning properly, the toxins from the blood have to be removed through an external device, dialyser. The dialyser mimics the kidney and draws the blood out of the patient, purifies the blood and infuses them back into the patient's body. When the blood is infused back to the human body, it has to be at 37.6 °C failing which would lead to hypothermia, a fatal condition. To avoid this fatal condition, the body is heated to human body blood temperature 37.6 °C. The blood flow rate in the bloodline can vary over a wide range but still the blood has to be heated to 37.6 °C which is human body blood temperature. To raise the human body temperature to 37.6 °C, a sub-component in the dialyser machine called blood warmer takes care of heating. Since it is a medical device which interacts with the patient, the precision of 0.5 °C is expected over a range of flow rates based on patient's medical condition. The scope of this work is the design of a precise blood warmer which is effective than existing. The multiphysics simulations helped the engineers to visualize, observe and understand the physics involved during the design stage.

Keywords Multiphysics · Blood warmer · Product design · Flow rate and heat supply

81.1 Introduction

Dialysis is a process through which waste products artificially removed from the blood to compensate for the presence of an abnormally functioning kidney. In a healthy kidney, excess fluids and electrolytes such as sodium and potassium are removed from the body in the form of urine. In the case of kidney disease, the loss

R. Kapuganti (✉)
HCL Technologies, Chennai 600101, India
e-mail: rajesh-kapuganti@hcl.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_81

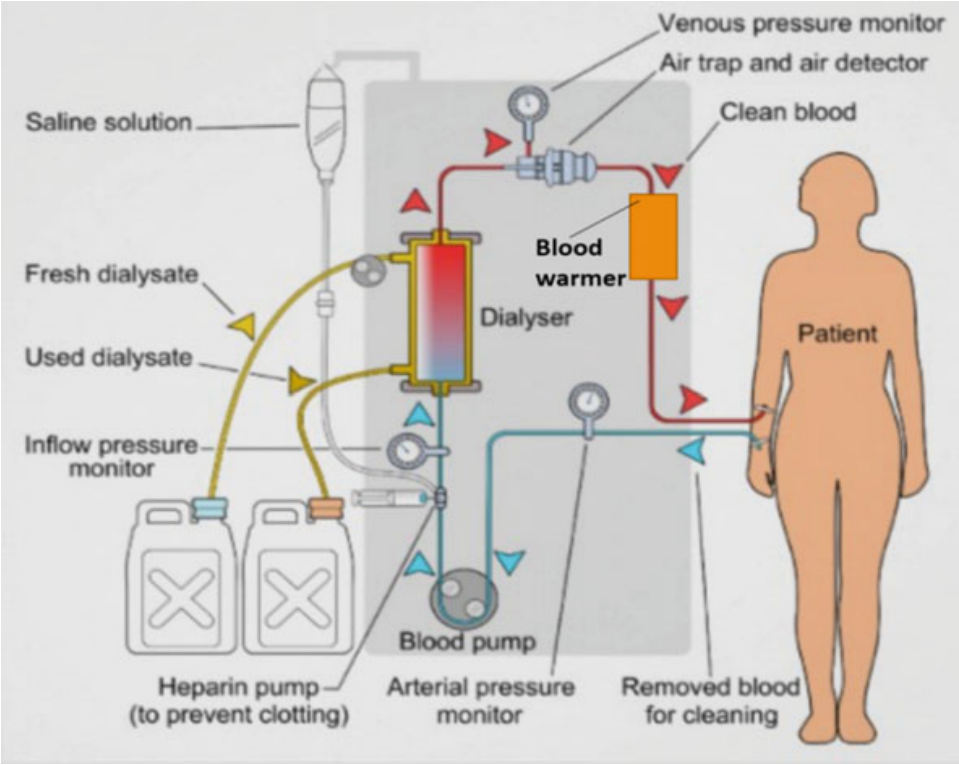


Fig. 81.1 Blood warmer

of kidney function results in the accumulation of excess waste, toxins or fluids in the body thereby causing other complications such as hypertension, heart failure, metabolic acidosis and hypokalemia [1]. Dialysis is performed in the situations of acute kidney failure and to remove toxins in blood from the body (Fig. 81.1).

81.1.1 Blood Warmer

Blood warmer is a sub-component in dialyser machine which is intended to deliver the blood or fluids before to transfusion to a human body at 37 °C which is the set temperature. Malfunctioning of the blood warmer device could lead to hypothermia. The blood warmer heats the blood to the desired temperature (37 °C) that is non-toxic for infusion.

The resistance heating technology which has a silicone heater patch and aluminium plates between which the fluid tubing passes was selected based on the thermal efficiency results. In designing the heating plate in the blood warmer, the data needed are type of plate material, plate dimensions, the effective area of the heating plate, the upper and lower temperature limits of the heating plate and the discharge of blood flow [1].

Though current blood warmers can heat blood to the human body blood temperature, the challenges are temperature management, heat mapping across the plate, optimum sensor location and the determination of the right compensation factor for dynamic control algorithm. This study was performed to improve the performance and functional aspects of the blood warmer design.

The device manufacturer supplied the complete list of product requirement specifications which had all the details of constraints, functionalities and geometric details. The design and development of blood warmer subassembly involved a multidisciplinary team which consisted of simulation, software, hardware, CAD, electrical, testing, industrial design and materials engineers.

However, in this paper, only the electro-thermo-fluid interactions of the multiphysics simulations are discussed. The multiphysics simulation procedure and the results are discussed.

81.2 Multiphysics Simulation Procedure and Results

81.2.1 Simulation Procedure

The multiphysics simulations were performed to determine the power input required for eight different flow rates to reach a desired set temperature of 37.6 °C from the inlet temperature of 30 °C. The range of flow rates used in the simulation are 100, 150, 200, 250, 300, 350, 400 and 450 ml/min. By varying the power input applied on the heating coils of heater plates, the power required to achieve the desired set temperature was determined. A convective heat transfer coefficient of 10 W/m^2 K was assumed considering that the heating unit is placed in an enclosure. The operating fluid was assumed as water instead of blood to avoid the nonlinearities. Humidity effects were not considered in the simulation.

The voltage difference was applied on the coil path of the top plate and the bottom plate (to generate joule heating in the plates). A perfect thermal contact was assumed between the disposable and heater plates.

The blood warmer assembly used in the dialysis machine consists of four components, the fluid domain, blood flow path, the heater coil and the two aluminium heater plates. The material properties of the different components used in the simulations are as listed in Table 81.1 (Fig. 81.2).

81.2.2 Simulation Results

The power required for the flow rates 100, 150, 200, 250, 300, 350, 400, 450 (ml/min) are 60, 89, 118, 148, 177, 206, 234, 264 (W), respectively, as listed in the Table 81.2.

Table 81.1 Material properties of the blood warmer assembly

S. No.	Component	Material	Property	Values	Unit
1	Fluid	Water	Specific heat	4181.8	J/kg K
			Density	1000	kg/m^3
			Dynamic viscosity	1.002	(Pa-s, N-s/m^2) $\times 10^{-3}$
			Kinematic viscosity	1.004	(m^2/s) $\times 10^{-6}$
			Thermal conductivity	0.58	W/(m K)
2	Blood flow path	Polyethylene	Specific heat	1800	J/kg * K
			Density	1500	kg/m^3
			Thermal conductivity	0.209	W/(m K)
3	Heater plate	Aluminium	Specific heat	956	J/kg * K
			Density	2699	Kg/m^3
			Thermal conductivity	165	W/(m K)

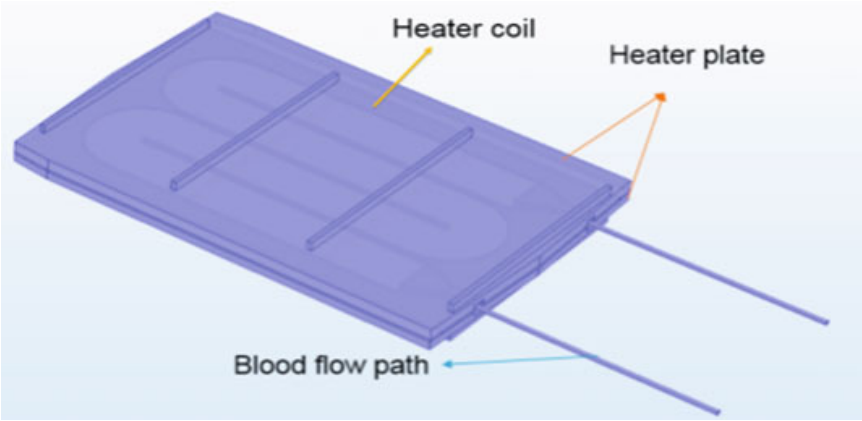


Fig. 81.2 Blood warmer assembly

Table 81.2 Heat energy required versus mass flow rate

S. No.	Mass flow rate (ml/min)	Heat energy required (W)
1	100	60
2	150	89
3	200	118
4	250	148
5	300	177
6	350	206
7	400	234
8	450	264

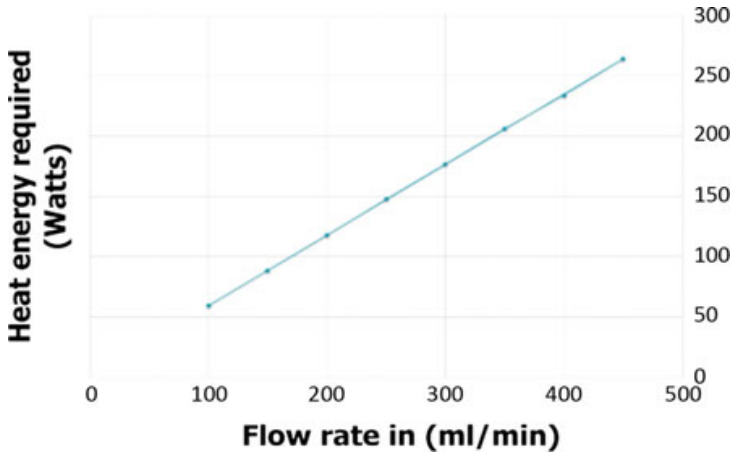


Fig. 81.3 Relationship between heat energy and flow rate

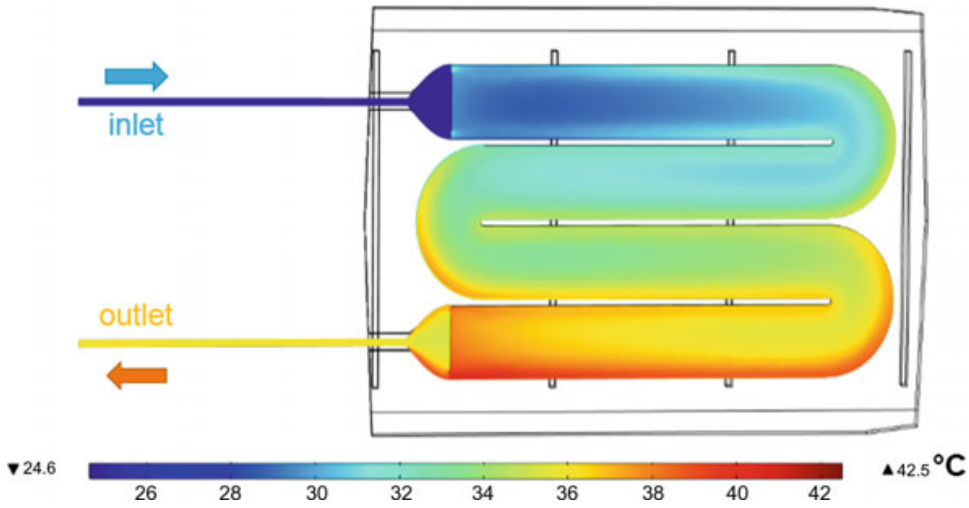


Fig. 81.4 Temperature distribution along the blood flow passage

There is a linear directly proportional relationship between the mass flow rate and heat energy required as depicted in the Fig. 81.3. The spatial thermal distribution of the blood flow path is depicted in the Fig. 81.4. The temperature of the fluid increases along the blood flow passage.

The temperature distribution is higher in the exit region of the blood warmer and lower at the entry region of the blood warmer. The spatial thermal variations on the heater plate are depicted in Fig. 81.5.

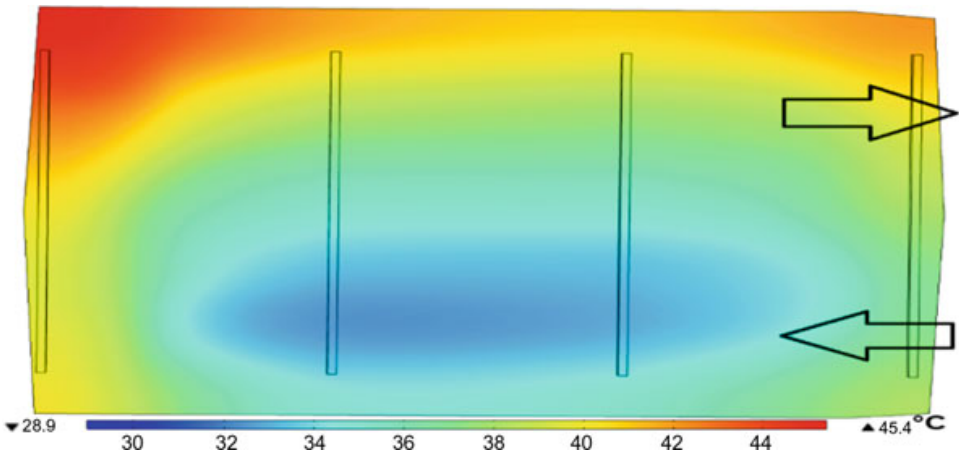


Fig. 81.5 Temperature distribution in the heater plate

81.3 Conclusion

1. Based on the simulation results, the heat energy required at different flow rates to heat the blood to 37 °C was obtained. This information was useful for further developments in the product design cycle. There is a linear relationship between the flow rate and heat energy required.
2. The temperature gradient is observed across flow passage as it gets heated up.
3. The correlation between testing and multiphysics simulation results was good. The multiphysics simulations were coupled with MATLAB for creating the algorithms, which was possible because of the virtual prototyping capability of COMSOL.
4. The product development process has been shortened in the concept and prototyping phase by effectively using multiphysics simulations.

References

1. Hendrarsakti, J., Ichsan, Y.: Experimental study of isothermal plate uniformity for blood warmer development using geothermal energy. IOP Conf. Seri. Earth Environ. Sci. **42** (2016)
2. http://e-pao.net/epSubPageExtractor.asp?src=education.Health_Issue.Dialysis_Everything_you_need_to_know_By_Sanjeev_Gulati

Chapter 82
Finite Element Analysis of a Disc Brake Mounted on the Axle of a Train

E. Madhusudhan Raju, Somayajula Kavya Geetika and Devireddy Jitesh Krishna

Abstract Axle-mounted disc brakes are generally employed in high-speed trains owing to the lesser weight and sustainable thermal properties. Most widely used material for manufacturing of the disc brake is grey cast iron (GCI), because of its availability. However, it being a brittle material has certain limitations. Hence, an aluminium composite material (AlSiC) is considered for analysis under different loading conditions. Three models of brake discs are developed on which both structural analysis and coupled analysis were done and results are compared.

Keywords Brake disc · Finite element analysis · Stress analysis

82.1 Introduction

In the olden days, the braking system of railway coaches consisted of wheel thread braking system. But they are not applicable to high-speed trains, as they damage the wheel thread. Hence, axle-mounted disc brakes are being employed in coaches. The braking system of a train involves a very complex process and is of great importance by making an essential contribution to the passenger's safety. This complexity results from numerous mechanical, thermal, pneumatic and electrical phenomena occurring during the process of braking. The actions of the above processes act on the different parts of the train with variable intensities. Also, the thermal deformation and overheating of the disc brake system during braking are turning out to be crucial. It is difficult to determine beforehand and analyse discs and linings separately, due to the amount of heat transferred to the discs, which may depend on their design and the train speed. In this paper, the effects of thermal and structural loads coupled together on the brake disc are analysed with different materials.

E. Madhusudhan Raju · S. Kavya Geetika · D. Jitesh Krishna (✉)
MED, UCE, OU, Hyderabad, India
e-mail: jiteshd.be19@uceou.edu

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_82

82.2 Literature Review

Reibenschuh et al. [1] used two brake discs (new and old) for both structural and thermal analysis. The material used for both the discs was spheroidal graphite cast iron. They found that the stresses developed in the new disc were more than worn disc, but stress values for both the discs were within the limits.

Maleque et al. [2] has developed a method of selection for selecting a better material in terms of optimization for the brake disc system, highlighting on substitution of cast iron by any lightweight material. Requirements for the material performance were scrutinized, and alternate solutions were assessed among different materials like cast iron, aluminium alloy, titanium alloy, ceramics and composites. The analysis led to a conclusion of usage of an aluminium metal matrix composite as the most appropriate and optimum material for brake disc system.

Yildiz and Duzgun et al. [3] had studied the stress analysis of ventilated brake discs by means of the finite element method. During this study, three different designs of ventilated brake discs were manufactured, and their braking system performances were examined experimentally together with a solid disc. These assessments indicate that the application of variable force along the brake pads leads to a reduction of the stress on ventilated discs from 8.8% to 19.1%.

Jung et al. [4] had studied shape optimization and the thermal characteristics of a ventilated disc. In this study, an analysis technique was developed to estimate the temperature rise and thermal deformation of the ventilated disc. Also, the braking condition and properties of the disc and pad could be estimated by this technique. Mathematical explanations for the analytical process of the braking power generation were developed.

82.3 Problem Definition

82.3.1 Aim of the Project

The primary objective of this paper is to present the various stresses and deformations induced in the brake disc models made up of grey cast iron (GCI) and aluminium composite materials (AlSiC) under static structural and coupled (structural and thermal) loading conditions. Three models were developed for the estimation of stresses, one of which is similar to the existing design, whereas the other two are new designs suggested to overcome the limitations of the existing one. The existing model is of the bolted joint type, and the other models are of push fit type. Model 2 is the push fit type of the existing model, and Model 3 is push fit type with cross-drills to reduce the weight. To determine the boundary conditions for the problem, an emergency braking situation of 160 kmph to stop on a flat horizontal track was utilized. The input loads, pressure, rotational velocity and heat flux induced were calculated and applied on the disc, and the output in terms of von

Mises stresses, total deformation and temperature induced in the brake disc was recorded and used for the finite element analysis.

82.3.2 2D Drawings of Models

See Figs. 82.1, 82.2 and 82.3 and Tables 81.1 and 82.2.

82.3.3 Software Used

1. Modelling—SOLIDWORKS 2016
2. Analysis—ANSYS WORKBENCH R16.

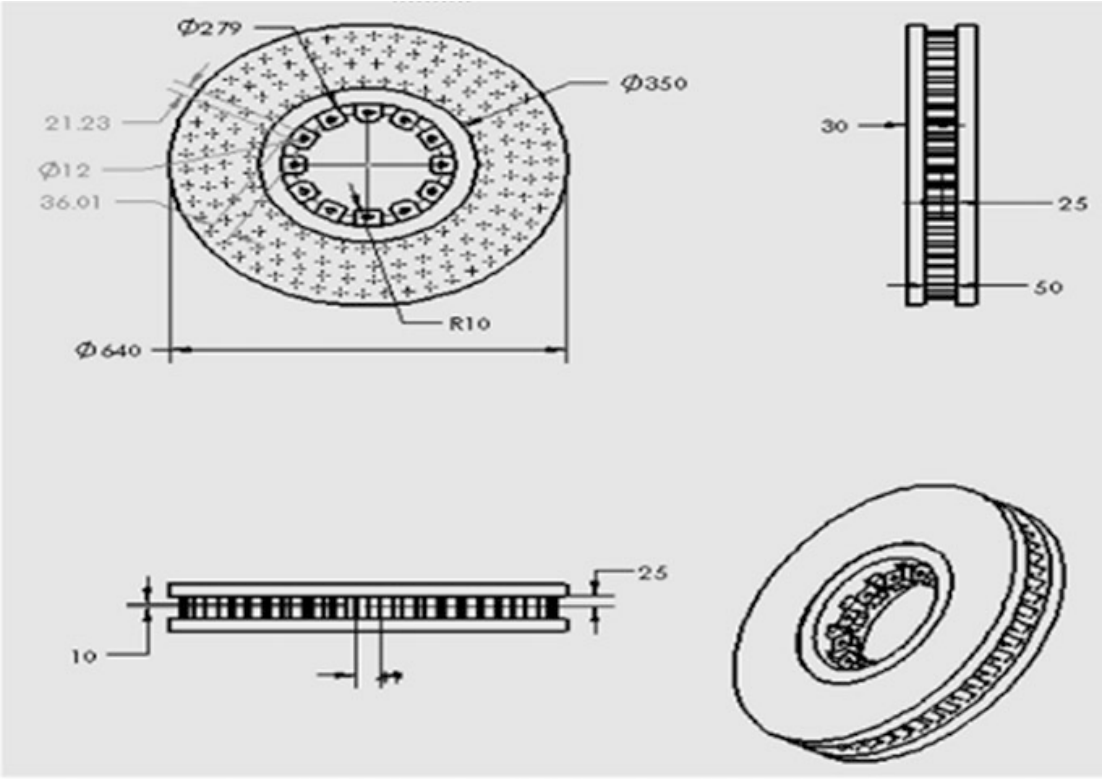


Fig. 82.1 2D model drawing of Model 1 (existing) [5]

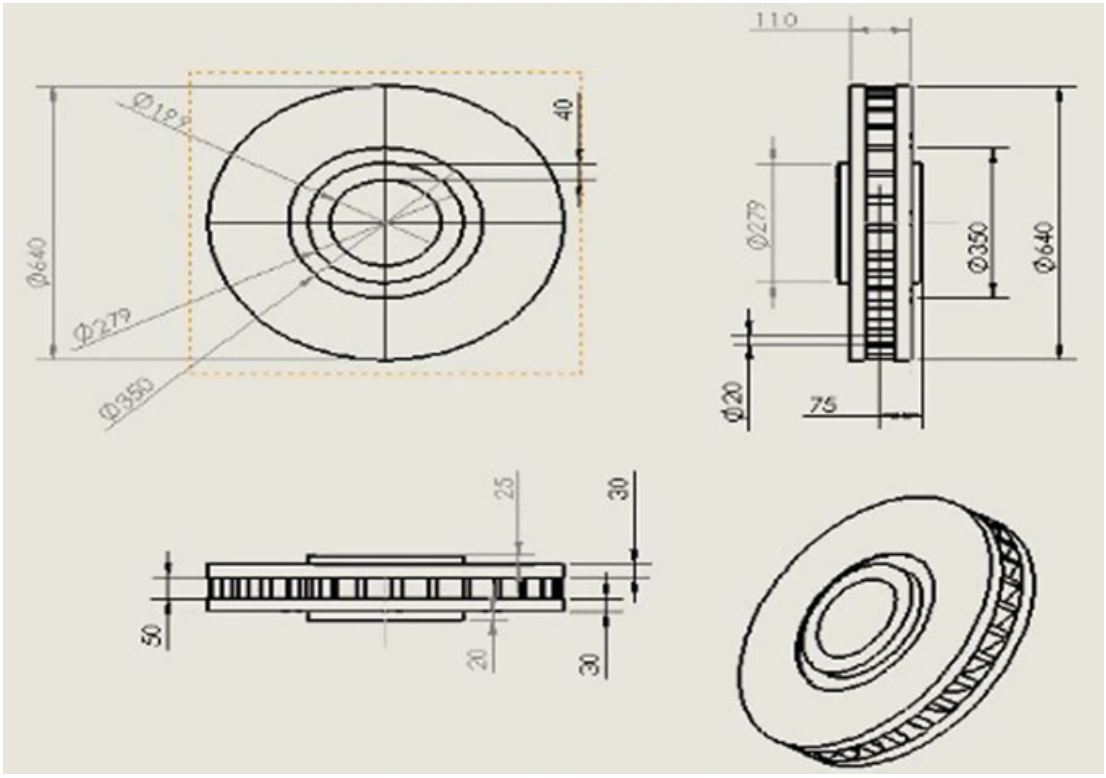


Fig. 82.2 2D model of Model 2 (push fit type)

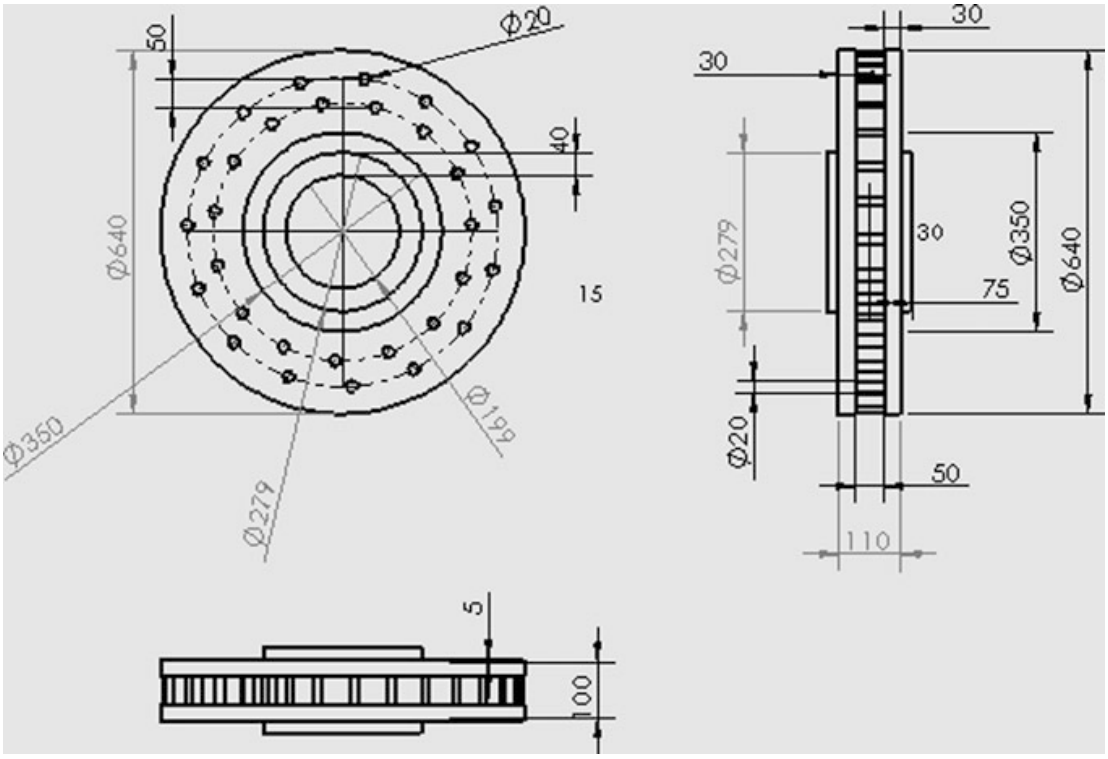


Fig. 82.3 2D model drawing of Model 2 (push fit type with cross-drills)

Table 82.1 Dimensions of disc brake [5]

S. No.	Description	Dimension
1	Dimensions	640 × 110 mm
2	Outer diameter of friction ring	$640^{+0/-1}$ mm
3	Inner diameter of friction ring	350 mm
4	Width of friction ring	$110^{+0/-0.3}$ mm
5	Inner bore	119 mm (H6)
6	Width of hub	150 mm

Table 82.2 Disc brake material properties

S. No.	Parameter	Units	GCI	AlSiC
1	Density	kg/m^3	7200	2950
2	Yield strength	MPa	–	2400
3	Ultimate strength	MPa	240	340
4	Young's modulus	GPa	43	93
5	Poisson's ratio	–	0.28	0.24
6	Shear modulus	GPa	43	93
7	Bulk modulus	GPa	83	147
8	Thermal conductivity	W/m°C	52	197
9	Specific heat	J/kg-C	447	808
10	Coefficient of thermal expansion	$*10^{-6}$	0.11	6.5

82.4 Research Methodology

82.4.1 Theoretical Background

Introduction. Due to the complex nature of the loading on the disc brake, the von Mises stresses induced due to the loading conditions are examined to determine the failure conditions of the model. The von Mises yield criterion (also known as the maximum distortion energy theory of failure) suggests that the yielding of a ductile material begins when the second deviatoric stress invariant reaches a critical value. This theory works well for most cases, especially when the material is ductile in nature behaviour.

Structural load calculations. Refer Appendix (Table 82.3).

82.4.2 Solution Methodology

The solution methodology to determine the better material and better geometry for railway brake disc is given by the following steps:

Table 82.3 Boundary conditions

S. No.	Analysis	Boundary conditions
1	Static structural	Pressure on disc—1.5 MPa Rotational velocity—97.12 rad/s
2	Coupled (static structural + thermal)	Temperature, as derived from transient thermal analysis for different materials and above-mentioned structural load

1. Dimensions, as given in Table 82.1, are taken from Railway Specification No. RDSO/2015/CG-03 'Brake disc for LHB-type coaches' [5] for modelling of brake discs.
2. Brake discs are modelled using SOLIDWORKS 2016 software and saved as an IGES file.
3. Import 3D model of brake discs into ANSYS WORKBENCH 16.
4. Analysis of brake discussing ANSYS WORKBENCH 16 by applying the boundary conditions as given above.
5. Static structural analysis is done by applying the pressure on each side of the disc surface and rotational velocity to investigate the von Mises and total deformation developed in the models.
6. The coupled analysis is carried out at all time points with respective temperatures by applying structural loads to obtain the stresses, deformation and strain under the coupled load of thermal and structural.
7. The obtained results are tabulated and compared. The material and geometry with the least values are considered as better material and better geometry.

82.5 Results and Discussions

82.5.1 Static Structural Analysis

The models developed in SOLIDWORKS 2016 were converted into IGES file and imported back into ANSYS WORKBENCH R16 for the finite element analysis. Boundary conditions as calculated in Sect. 82.3.1 were applied on the brake disc models. The constant pressure of 1.5 MPa was applied on both sides normal to the frictional surface of the discs. A rotational velocity of 97.12 rad/sec (ramped) was given at the axis of the disc. Fixed support was provided in the eye of fixing bolt. Both disc temperature and surrounding environment temperature are taken as 45 °C. The analysis was run for 36 s, i.e. the time taken by the vehicle to stop due to the application of the emergency brake. The output results of von Mises stress and total deformation developed in the model were recorded (Figs. 82.4, 82.5 and 82.6).

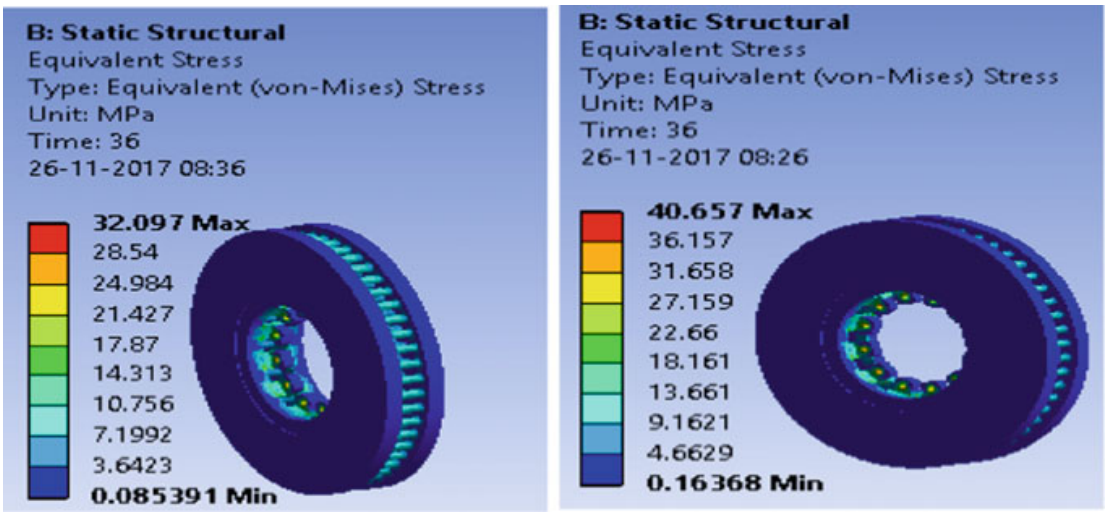


Fig. 82.4 SS GCI and AlSiC profile of von Mises stress (Model 1)

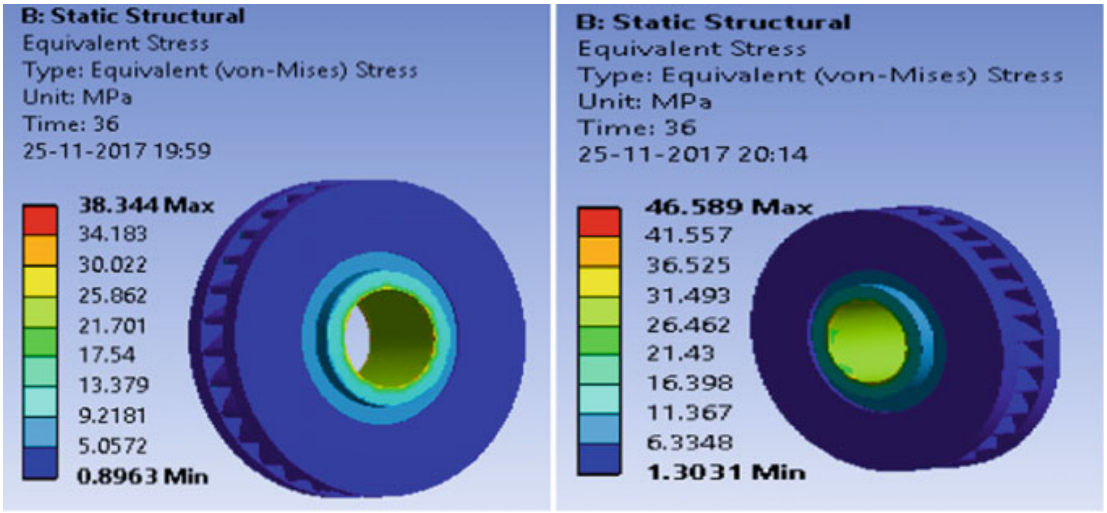


Fig. 82.5 SS GCI and AlSiC profile of Von Mises stress (Model 2)

82.5.2 Coupled Analysis

In the coupled analysis, the thermal load was coupled with the structural load to find out the combined effect on brake disc models. The temperature induced at various time points was imported into static structural, and then structural loads were applied. The analysis was run for 36 s, i.e. the time taken by the vehicle to stop due to the application of the emergency brake. The output results of von Mises stress and total deformation developed in the model were recorded (Figs. 82.7, 82.8 and 82.9).

Some important points that can be drawn from the analysis are:

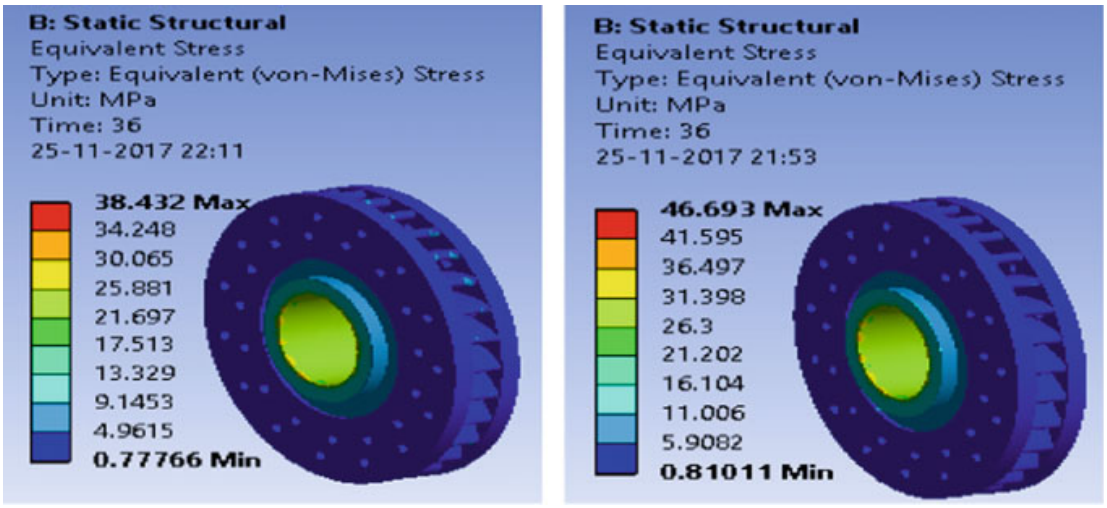


Fig. 82.6 SS GCI and AlSiC profile of von Mises stress (Model 3)

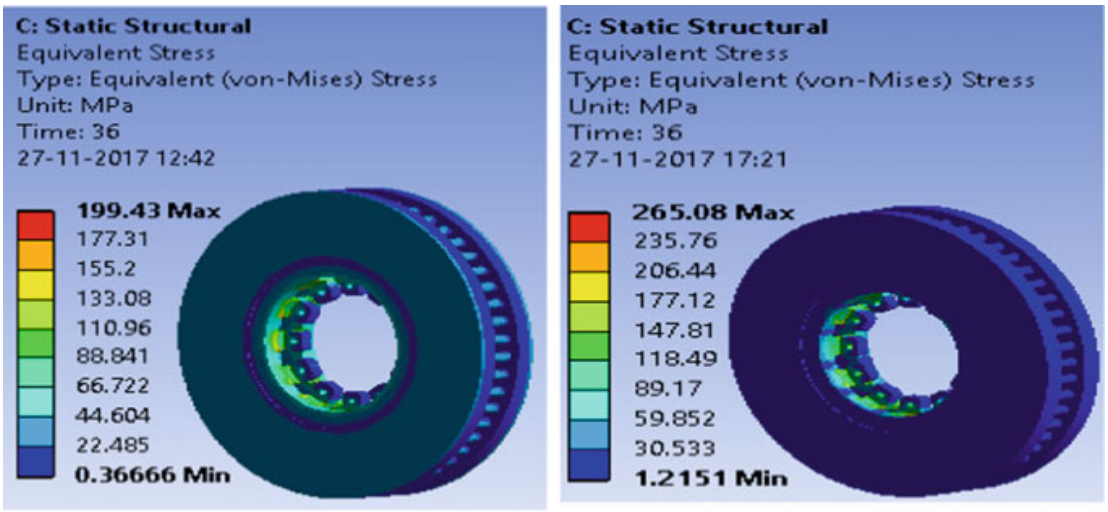


Fig. 82.7 CA GCI and AlSiC profile of von Mises stress (Model 1)

- When compared to the Factor of Safety offered by the GCI models, the AlSiC models offer higher Factor of Safety.
- For the same applied load, the AlSiC models have lower thermal stresses than the GCI models, as AlSiC material has greater thermal conductivity and heat dissipation capability.
- The weight of AlSiC models is lesser when compared to the GCI models (GCI model having a weight of about 134 kg gets reduced to 54 kg in case of AlSiC material).

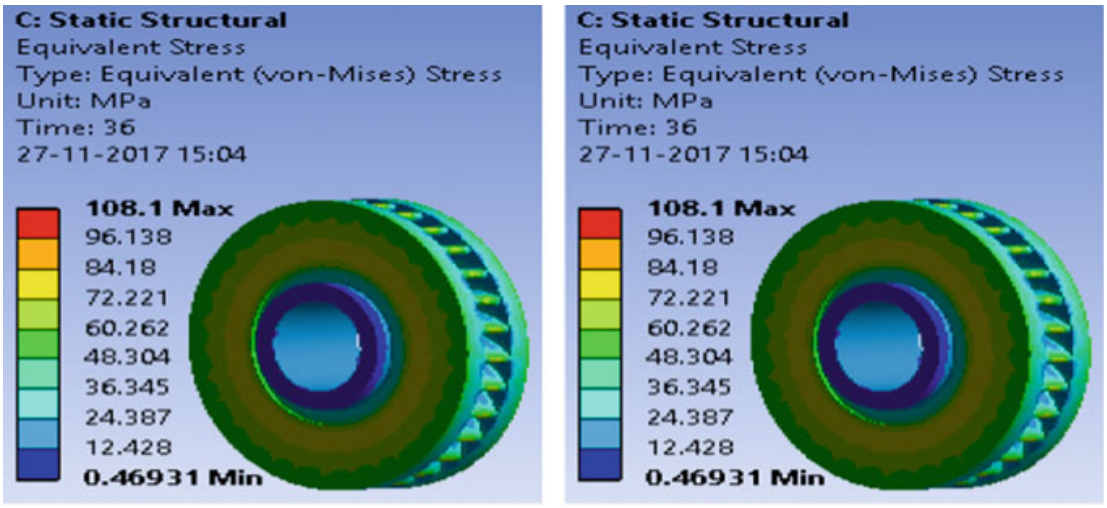


Fig. 82.8 CA GCI and AlSiC profile of von Mises stress (Model 2)

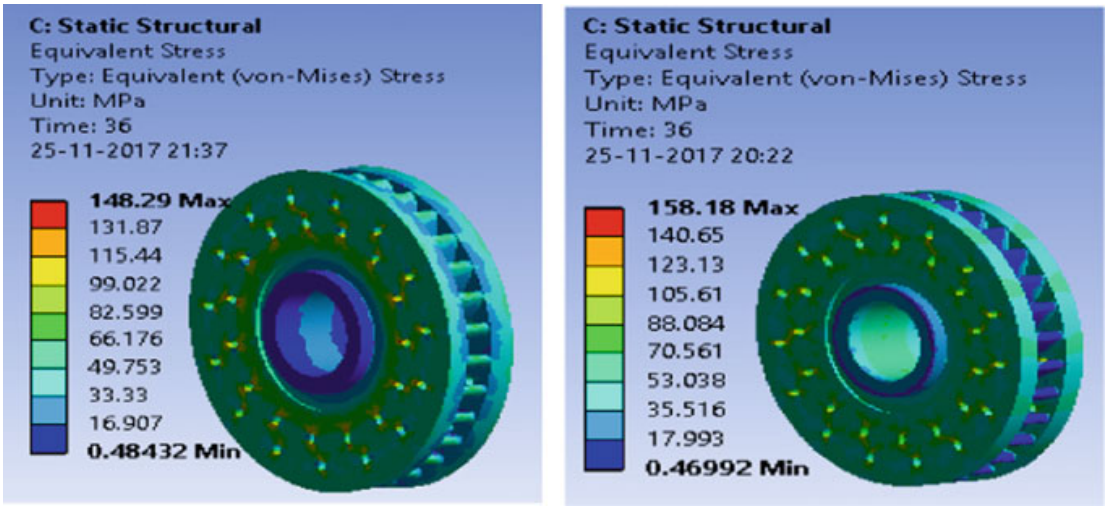


Fig. 82.9 CA GCI and AlSiC profile of von Mises stress (Model 3)

82.6 Future Scope

- Suggested models of brake discs are of a solid design which needs to be pressed on to the axle before pressing of the wheels. For replacement of these brake discs, wheels are necessary to be pressed out, which is not required if the models are of a split type. Hence, further experiments may be done with split-type brake disc models.
- Carbon matrix composites may be used to reduce the sound barrier and higher resistance to temperatures.
- Brake pad analysis with various materials suitable to the disc material may be carried out to optimize pad material.

Appendix

Structural load calculations: In this case, a railway vehicle travelling at a speed of 160 kmph on a horizontal track stops due to application of emergency brake was considered. Time of travel before stopping, deceleration, weight of the vehicle, clamping force on the brake disc, brake pad area, coefficient of friction, etc., for calculating the loads are taken from the railway specification. Value of Mass of railway vehicle—M = 64000 kg, No. of axles per vehicle = 4, Maximum load per axle = 16000 kg, no. of brake discs per axle = 2, Load on each wheel = 8000 kg, Start speed v_0 = 44.4 m/s, Deceleration a = 1.2 m/s^2, Braking time t_a = 36 s, Effective radius of the brake disc r_{disc} = 0.247 m, Radius of the wheel r_{wheel} = 0.458 m, Mean coefficient of friction brake pad μ = 0.35, Clamping force F_c = 42.1 kN, Surface area of brake pads A_c = 400 cm^2, Maximum temperature under sun = 70 °C, Maximum temperature under shade = 45 °C, Factor of Safety = 1.5 (Fig. 82.10).

Stopping distance

$$S = v_o t_s - \frac{1}{2} a t_s^2 = 822.24m \tag{82.1}$$

Determination of pressure on disc

Pressure acting on the brake disc,

$$P = \frac{F_c}{A_c \times \mu} = \frac{42.1 \times 1000}{800 \times (10)^{-4} \times 0.35} = 1.5\,\text{Mpa on each side} \tag{82.2}$$

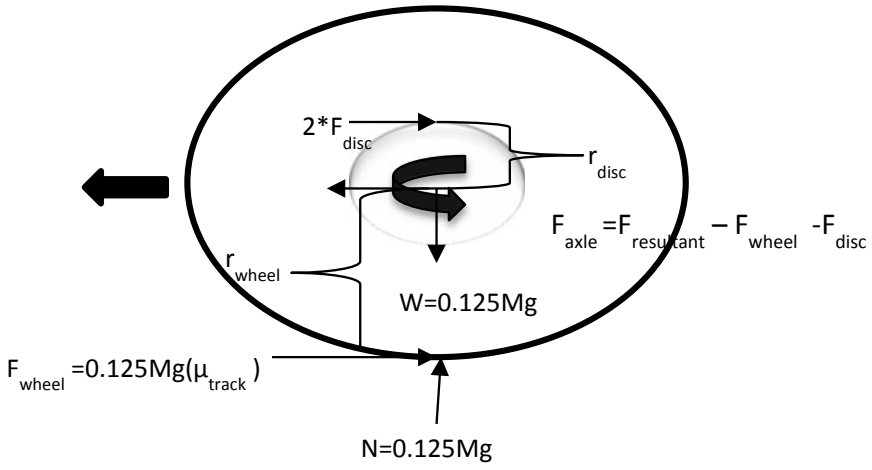


Fig. 82.10 Representation of forces acting on wheel and disc brake

where

F_c Clamping force (i.e. 42 kN)
A_c Contact area of brake pad on each side (i.e. 400 cm^2)
μ Coefficient of friction (i.e. 0.35).

Angular velocity

$$\omega = \frac{\text{Velocity}}{\text{radius}} = \frac{v_0}{r_{\text{wheel}}} = \frac{44.44}{0.458} = 97.12\,\text{rad/s} \tag{82.3}$$

Thermal load

The kinetic energy for one wheel (disc brake) is equivalent to the energy balance

$$\begin{aligned} 0.125 * \frac{1}{2} * M * v^2 &= \int_0^{t_s} P(t)\mathrm{d}t \\ &= 2 * F_{\text{disc}} \int_0^{t_s} v_{\text{disc}}(t)\mathrm{d}t \end{aligned} \tag{82.4}$$

The energy change at the moment is equal to the heat flux on the surface of the disc.

Equation (82.4) is valid in the case of constant braking deceleration. The braking force on the disc is equal to Eq. (82.7)

$$F_{\text{disc}} = \frac{0.125 * \frac{1}{2} * M * v_0^2}{2 * \frac{r_{\text{disc}}}{r_{\text{wheel}}}\left(v_0 * t_s - \frac{1}{2} * a * t_s^2\right)} = 8940\,\text{N} \tag{82.5}$$

The heat flux at the moment, which affects one half of the disc, is calculated according to the

$$\begin{aligned} Q(t) &= F_{\text{disc}} * v_{\text{disc}}(t) = F_{\text{disc}} * \frac{r_{\text{disc}}}{r_{\text{wheel}}}(v_0 - a * t) \\ &= 8940 * \frac{0.247}{0.458}(44.44 - 1.2 * t) = 214261 - 5786 * t\,\text{Watts} \end{aligned} \tag{82.6}$$

$$\text{Area of friction surface} = \frac{\pi}{4}\left(0.64^2 - 0.35^2\right) * 2 = 0.45\,\text{m}^2$$

$$Q(t) = \frac{214261 - 5786 * \mathrm{t}}{0.45} \ \mathrm{W/m^2} \tag{82.7}$$

For the case of emergency braking on horizontal track from 160 kmph to stop, the analysis was carried out in 36 steps, each step being 1 s long.

References

1. Reibenschuh, M., Oder, G., Čuš, F., Potrč, I.: Modelling and analysis of thermal and stress loads in train disc brakes—braking from 250 km/h to standstill. J. Mech. Eng. **55**(7–8), 494–502 (2009)
2. Maleque, M., Dyuti, S., Rahman, M.: Material selection method in design of automotive brake disc. In: Proceedings of the World Congress on Engineering, London (2010)
3. Yildiz, Y., Duzgun, M.: Stress analysis of ventilated brake discs using the finite-element-method. Int. J. Autom. Technol. **11**(1), 133–138 (2010)
4. Jung, S.P., Young, G.K., Park, T.W.: A study on thermal characteristic analysis and shape optimization of a ventilated disc. Int. J. Precis. Eng. Manuf. **13**(1), 57–63 (2012)
5. 'Brake disc for LHB type coaches', Indian Railways specification RDSO/2015/CG-03

Chapter 83
Modelling and Analysis of Composite Conical Space Adapter

Rajesh Addanki, Pranay Narinder Kumar, Ayush Mahajan and Manoj Kumar Buragohain

Abstract This paper is concerned with the development of model for the finite element analysis of the conical space adapter which is subjected to axial compression. It also provides the information about the parameters required for generating the input for modelling the structure. Analysis is performed using Ansys, i.e. software for finite element analysis. The results of analysis include the behaviour of the adapter under axial compression, e.g. stresses in the helical and circumferential ribs, displacement of the space adapter, critical buckling load and global buckling. Modal analysis has been carried out to study the modes of vibrations and obtain the frequency of vibration of the structure.

Keywords Composite lattice structures · Finite element analysis · Buckling analysis · Critical buckling loads · Modal analysis

Nomenclature

β Semi-cone angle
α_0 Geodesic angle at the start (at larger radius)
α_f Geodesic angle at the end (at smaller radius)
θ Angle of rotation of fibre
D Larger diameter of the adapter
d Smaller diameter of the adapter
b Width of the fibre
t Thickness of the fibre
H Height of the space adapter
W Width of the fibre

R. Addanki (✉) · P. N. Kumar · A. Mahajan · M. K. Buragohain
Composite Product Development Centre, Advanced Systems Laboratory, DRDO, Hyderabad 500069, India
e-mail: rajeshaddanki@rediffmail.com

P. N. Kumar
e-mail: pranaynkumar@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_83

83.1 History of Lattice Structures

The development of grid-stiffened structures dates back to WWII English bomber Wellington. The application of these structures is not limited to space applications. Earlier, grid-stiffened structures were made by machining metals, e.g. aluminium isogrid structures by machining aluminium slabs. But today the material has changed from metal to carbon fibre composites. Currently, the fabrication methods of a lattice structure are done by grooving the fibre path over the mandrel of required dimensions, and fibres coated in epoxy resin are wound over it. After curing it at a required temperature, the desired shape is obtained.

One of the earliest lattice structures was developed by Russian-based laboratory. Central Institute of Special Machine Building (CRISMB) was Moscow radio tower [1] designed by Shukov in 1921. It stands 148.3 m tall and marks the advent of the development of lattice structures for Russians. It looks quite similar to modern-day lattice structures.

In 1985, US navy developed a launch vehicle adapter fitting for their satellite named as [2] Petite Amateur Navy Satellite (PANSAT) for communication purposes. The study conducted on this adapter fitting provides a basis for structural design and analysis for the adapter. The thesis gives an in-depth knowledge about the finite element analysis of such structures.

Since the inception of space adapters, they were manufactured in various shapes and sizes [3–6] offering flexibility in design to the user, such as using a cylindrical and conical adapter. A conical configuration of the shaped adapter has been modelled out for easiness of mounting the payload with available space and thereby reducing the additional stage. Even though the manufacturing of such structures is limited to axisymmetric shapes, the use of these structures remains dynamic.

83.2 Introduction

A typical space adapter consists of helical and circumferential ribs as shown in Fig. 83.1 wherein the impregnated fibres are laid along the grooves for circumferential and longitudinal pattern to form a lattice structure. When the structure is subjected to axial compression, the helical ribs transfer the load to circumferential ribs as tensile force. Now, this tensile force is transferred to a member to the adjacent structure which stabilizes its shape and increases the load carrying capacity. The load-carrying capacity is greatly influenced by dimensions of lattice structure. Therefore, keeping the space constraints in mind, the roadmap of designing process is laid down.

The space adapter is used to install payload over the last stage of the composite rocket motor casing. The conical-shaped adapter is made of carbon fibre with epoxy resin which provides higher strength to weight ratio than conventional metallic

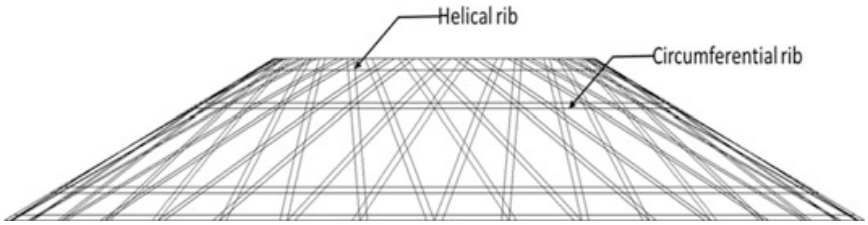


Fig. 83.1 Schematic of space adapter

structure made of steel or aluminium. Further, handling and installation of the adapter becomes easy due to its low weight.

In the present context, a space adapter is designed and modelled for finite element analysis [7]. The analysis consists of static analysis followed by buckling analysis. For calculating the natural frequency, modal analysis is carried out. The boundary conditions and load cases are explained in further sections. The results obtained from the analysis are also discussed in the paper. Further, prospects and research areas are also discussed which can be explored in the future for further research.

83.3 Configuration

The conical adapter is designed with two circumferential stiffeners in addition to metallic stiffeners at top and bottom and 20 crossed axial stiffeners with varying helical angle. A total of 40 layers with 0.5 mm thickness are considered for analysis. The helical angles with fibre path variation cater for a height of about 300 mm with outer diameter to inner diameter ratio of three. The helical ribs of cross section 8 × 8 to 20 × 20 in multiples of four have worked out to cater the axial payload of about 50 kg with 10 g axial force. A typical 2D view of space adapter is shown in Fig. 83.2.

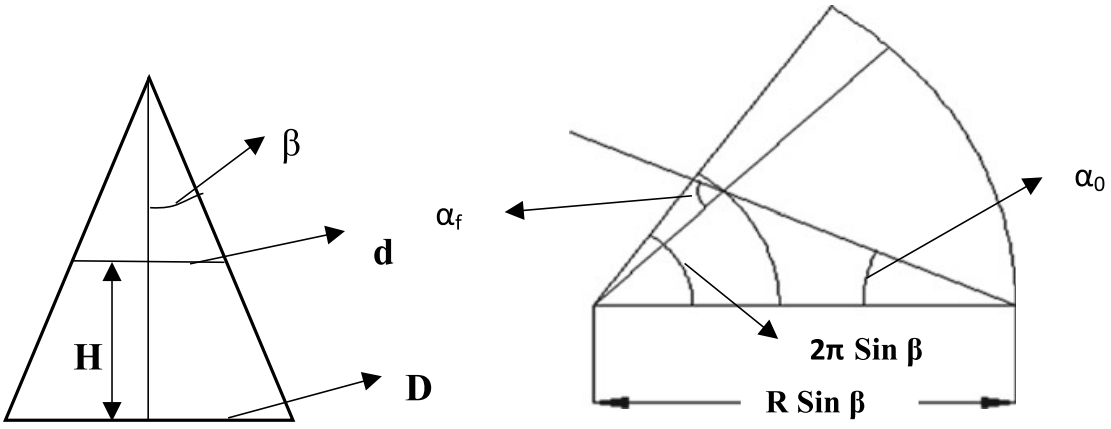


Fig. 83.2 Development of the cone

The empirical relations for calculating the path, angle and radii are as given below (from Eqs. 83.1 to 83.6).

$$\frac{r\varnothing}{\sin\beta} = r\theta \tag{83.1}$$

$$\theta = \frac{\varnothing}{\sin\beta} \tag{83.2}$$

$$\varnothing = \alpha - \alpha_0 \tag{83.3}$$

Also, as per Clairut's equation

$$R\ \sin\alpha_0 = r\ \sin\alpha \tag{83.4}$$

$$\alpha = \sin^{-1}\left(\frac{R\ \sin\alpha_0}{r}\right) \tag{83.5}$$

Substituting Eqs. 83.2 and 83.5 in Eq. 83.3, we get,

$$\theta = \frac{1}{\sin\beta}\left[\sin^{-1}\left(\frac{R\ \sin\alpha_0}{r}\right) - \alpha_0\right] \tag{83.6}$$

83.4 Input Generation

The geometric parameters used for the design and modelling the conical space adapters serve as the inputs. Equations which are used for modelling are stated in the previous section. Various cross sections were designed and analysed to arrive at a final configuration. The parameters used are as follows:

- Radius of the frustum of the cone (both smaller and larger)
- No. of helical ribs
- No. of circumferential ribs
- Winding angle at start and end
- Width of fibre
- Cross section of the rib.

Once the above parameters are finalized then model becomes ready for analysis. The values of above parameters are used in the governing equations refered in Sect. 83.2 to obtain the set of coordinates. The properties of composites, which were experimentally evaluated from the sample test as per ASTM standards, are used for the analysis by testing the small specimens constructed in the scale down model of the full-scale prototype. After the characterization of the high-strength carbon fibre with epoxy resin, the following properties are obtained which are given

Table 83.1 Carbon epoxy composite properties

Parameter	E_{xx} (MPa)	E_{xy} (MPa)	E_{yz} (MPa)	μ_{xx}	μ_{xy}	μ_{yz}	G_{xx} (MPa)	G_{xy} (MPa)	G_{yz} (MPa)
Value	117,000	6000	6000	0.3	0.29	0.29	3000	2000	2000

Note Where symbols have their usual meaning like elastic modulus, shear modulus and Poisson's ratio in respective directions of the laminate

Table 83.2 Properties of steel

Properties	Value (MPa)
E	201,000
μ	0.3

below in Table 83.1. Also, the material input properties for the steel are listed in Table 83.2.

83.5 Modelling

The helical path with circumferential ribs has been modelled in ANSYS which is software based on finite element analysis. The relations mentioned in the previous sections are used to create a set of coordinates of fibre path. With +θ coordinates and −θ coordinates have been developed the path is fed to software. In the similar lines, 90° path (circumferential ribs) was also modelled. The nodal points and the element generation were done by brick node 185. In total, 20 helical ribs with an angular difference of 18° between any two +θ helical ribs are considered. The nodes are merged to eliminate any discrepancy at the crossovers. While modelling, a special attention must be given to crossovers (where ribs are meeting each other) because crossover points act as single unit not different one, otherwise results obtained may be erroneous or not at all accurate. The adapter model with helical and circumferential ribs is shown in Fig. 83.3. Figure 83.4 shows the element detail depicting the nodal merging at the crossover points.

83.6 Boundary Conditions and Load Cases

The space adapter is modelled in global Cartesian system in such a way that the bottom surface lies in *X-Y* plane and positive *z*-axis is aligned with axis of the cylinder. Following are the boundary conditions applied on the mode for the analysis:

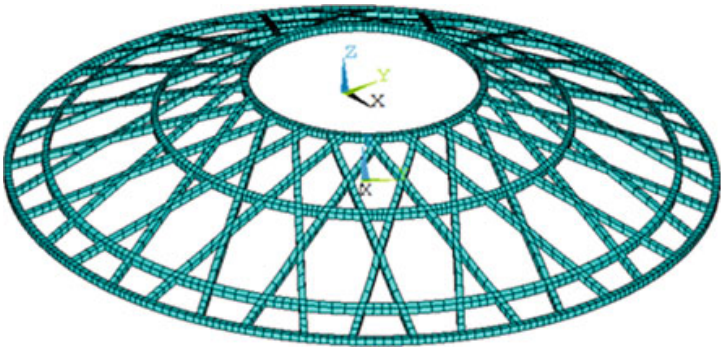


Fig. 83.3 FEA model of space adapter

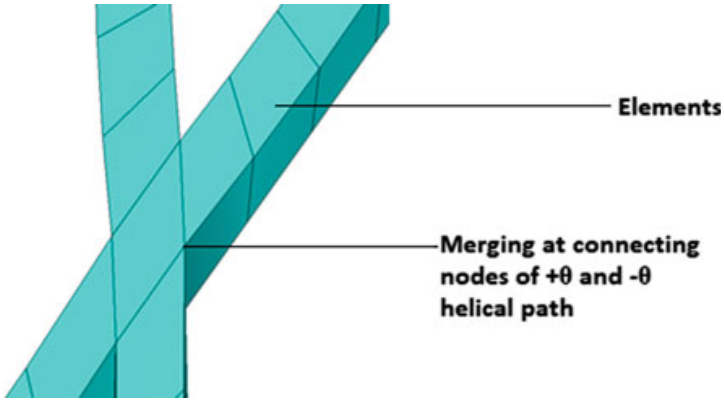


Fig. 83.4 Elements in the model

- Displacement of the bottom surface has been arrested and at the top of it is kept free where the load is applied, i.e. at $z = 0$ $U_x = 0$, $U_y = 0$, $U_z = 0$ and at $z = H$ U_x, U_y, $U_z \neq 0$.
- A uniform load is applied on the top nodes at $z = H$.
- To have the interface, a metallic bulkhead at top and bottom of the adapter (lower and higher Diameter) steel is considered with properties as follows.

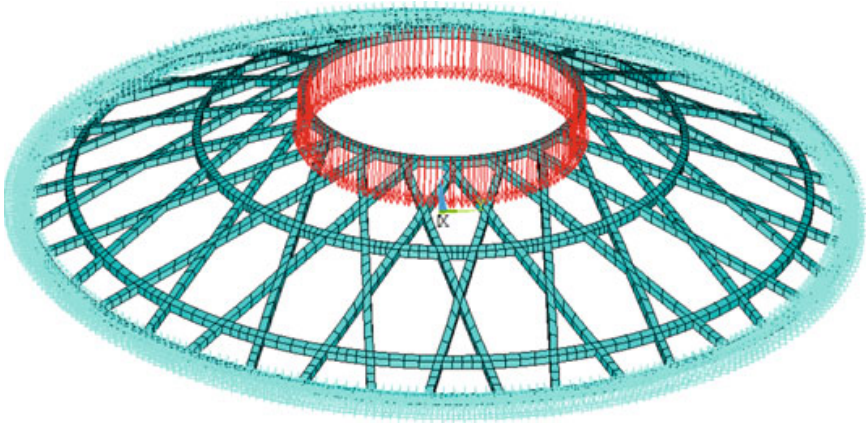


Fig. 83.5 Load application and boundary condition for the adapter

Design load, in this model, is applied for two types of analysis, i.e. static and buckling. Axial compressive load used for analysing the structure is 10 tons and the loading pattern is shown in Fig. 83.5.

The model is fixed at the bottom and load is applied at the top. A load of 10 tons is applied on 280 nodes at the top. Then load per node comes out to be 357.14 N.

In addition to static analysis, modal analysis was also conducted on the space adapter model (20 × 20) to study the vibrations and calculate frequency of the adapter.

Modal analysis is used to calculate the natural frequency and mode shapes of a continuous structure. A continuous structure has infinite degrees of freedom, but finite element approximates with a finite number of degrees of freedom.

Block Lanczos method is used for performing the modal analysis, which uses an algorithm for finding the null space of a matrix over a finite field. The same method is used in the present study for the matrix functions and in calculating buckling factor. The eigen buckling analysis is performed and three modes are extracted for the buckling factor.

The results of modal analysis serve as the basis for harmonic analysis. Harmonic analysis solves time-dependent equations of motion for linear structures undergoing steady state vibrations.

83.7 Results and Discussions

Several models were configured with different cross sections ($w \times t$) varying as 8 × 8, 16 × 16, 18 × 16, 20 × 20. Few results are represented to validate the strength and configuration of the space adapter, and the buckling factors variations are tabulated in Table 83.3.

A typical buckling mode is also represented in Fig. 83.6.

It is observed that, as the rib cross section increases, the load carrying capacity of the adapter also increases. The buckling factor of 5.62 indicates that critical load is 5.62 times of applied load. Higher buckling factor ensures that design will not fail under the action of axial compressive loads. The stresses are developed in the layers of the model.

The graphical plot in Fig. 83.7 demonstrates the behaviour of stress in 40 layers of the model at three different heights, i.e. top, middle and bottom of the adapter.

Table 83.3 Buckling factor variations

SI No.	Cross section of the ribs (w × t)	Buckling factor
1.	8 × 8	0.2295
2.	16 × 16	0.9122
3.	18 × 16	1.1011
4.	20 × 20	5.6200

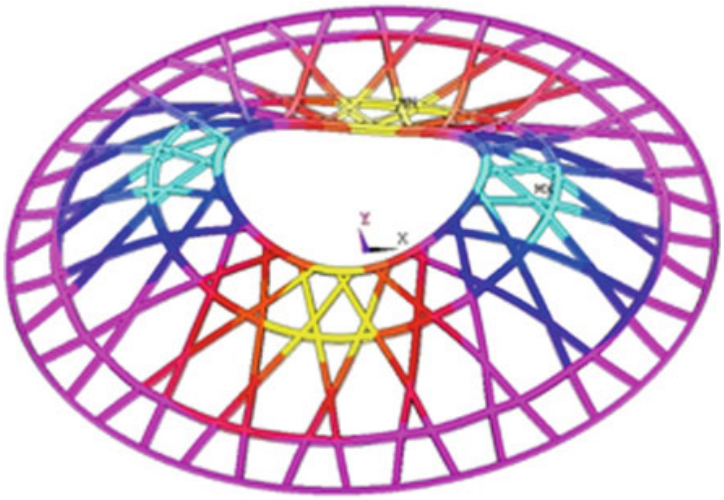


Fig. 83.6 Buckling mode of the adapter

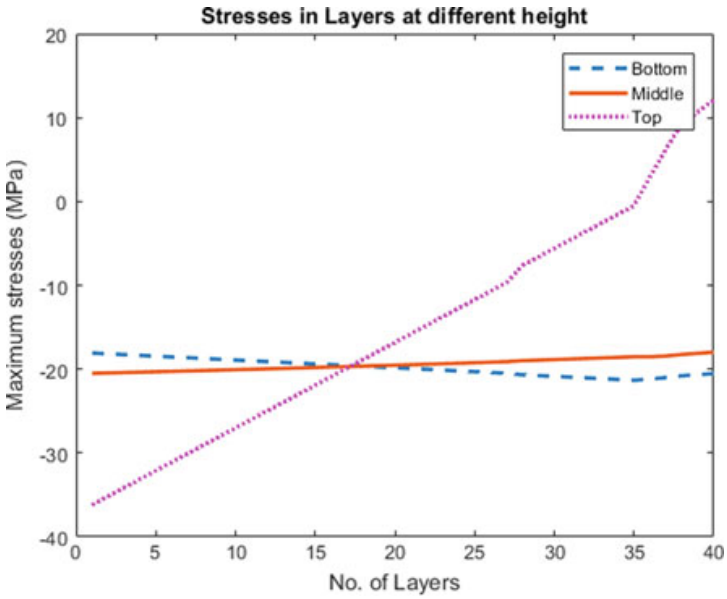


Fig. 83.7 Stress pattern in the helical plies at various height

The behaviour of stress shown in the graph indicates that the layer between 15 and 20 experiences same stress and behaves as the neutral layer.

The graphical plot in Fig. 83.8 depicts the stress in layers at the crossovers of the helical rib. The plot indicates that stresses start as compressive but as we move along the transverse direction to the layers, the nature of stress starts changing. The topmost layer, i.e. layer 40, experiences lesser compressive force compared to lowermost one. Up to layer no. 35, the reduction in compressive stress is gradual, but later there is drastic reduction in the compressive stress value.

The graphical plot in Fig. 83.9 shows that stresses in layers of circumferential rib are largely tensile attributing to the fact that helical ribs transfer compressive stress to the circumferential ribs which change it into tensile. There is gradual decrease in the stress in the transverse direction indicating that the load is distributed uniformly.

As we noted earlier, that nature of stress is different in helical and circumferential rib, the stress pattern changes from compressive to tensile is clearly seen in

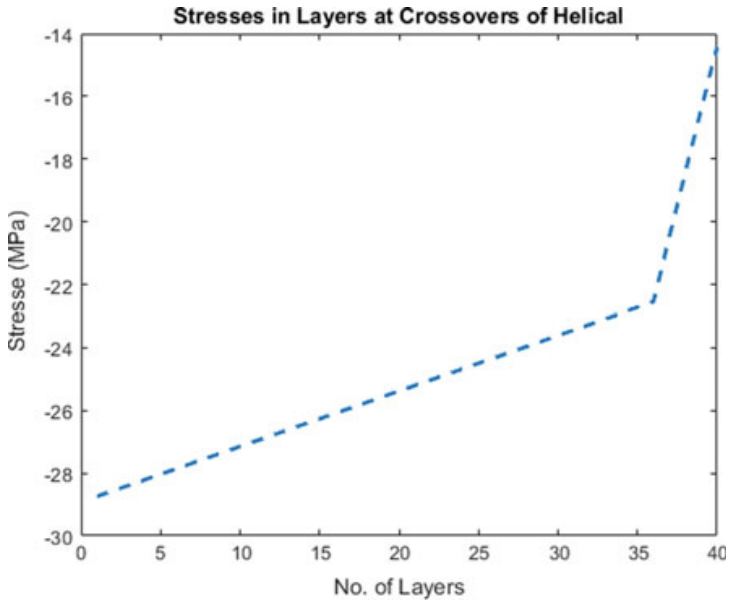


Fig. 83.8 Stress pattern in the helical plies at crossover

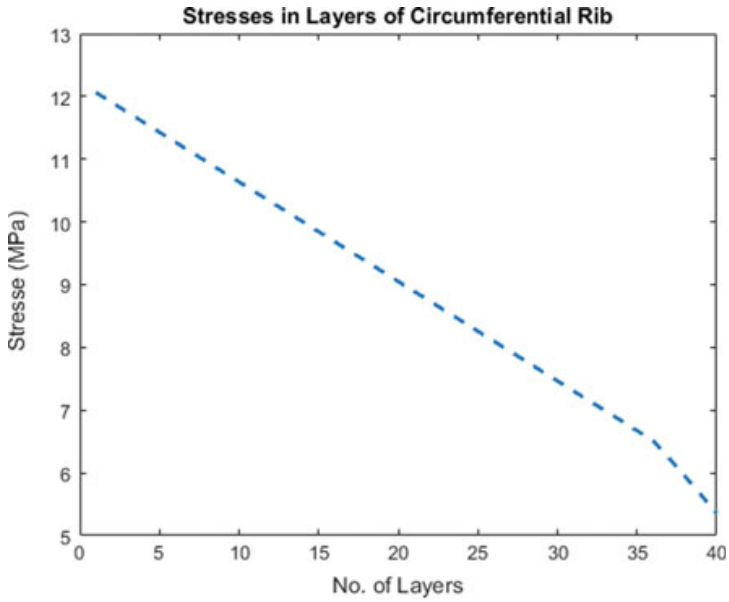


Fig. 83.9 Stress pattern in the circumferential ribs

Fig. 83.10. It can be seen, at the crossover, the stresses change from compressive to tensile indicating the transfer of the loading and is evident from the plot that the load is transferred from lowermost ply to topmost ply. Since, the value as well as the nature of the stress changes gradually up to layer 35 and the slope is high.

Figure 83.11. shows that as load is increased, the displacement also increases in the negative z-direction, i.e. along the height of the adapter and also follows a linear relationship. The deformed shape of the adapter superimposing onto the undeformed configuration is shown in Fig. 83.12.

The results obtained from static analysis have been saved and further utilized for extraction of mode shapes in modal analysis. In modal analysis, three modes were

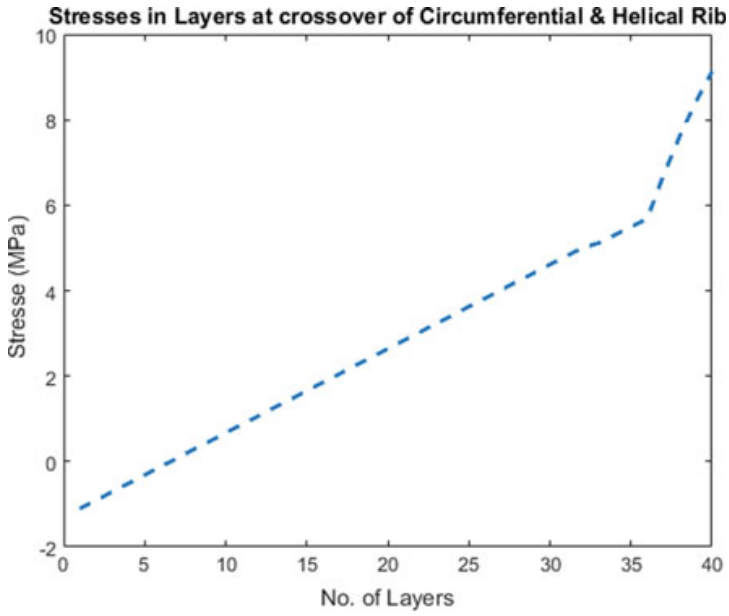


Fig. 83.10 Stress pattern in the circumferential and helical crossovers

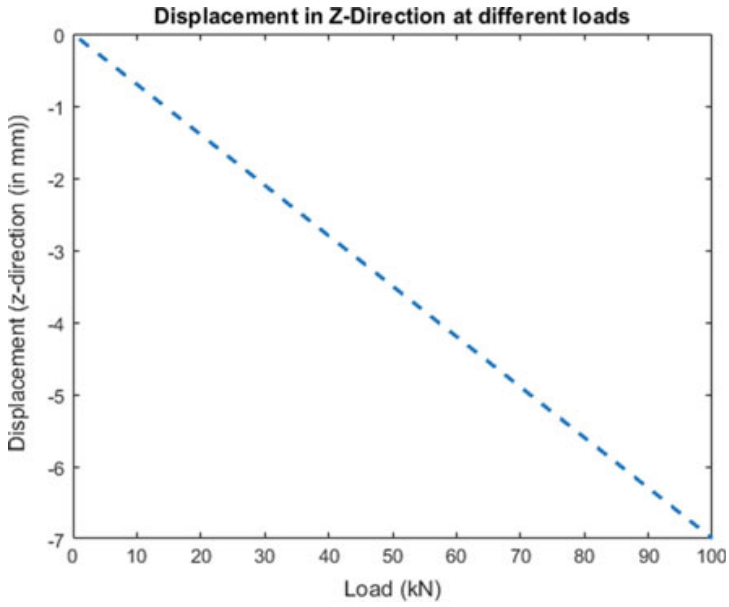


Fig. 83.11 Displacement in z-direction at different loads

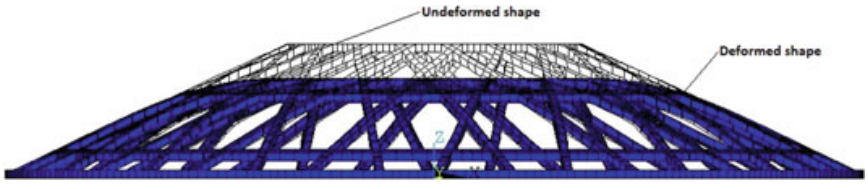


Fig. 83.12 Deformed shape versus original shape

extracted and the results obtained in modal analysis are as follows (refer Fig. 83.13):

The modal frequencies are listed in Table 83.4.

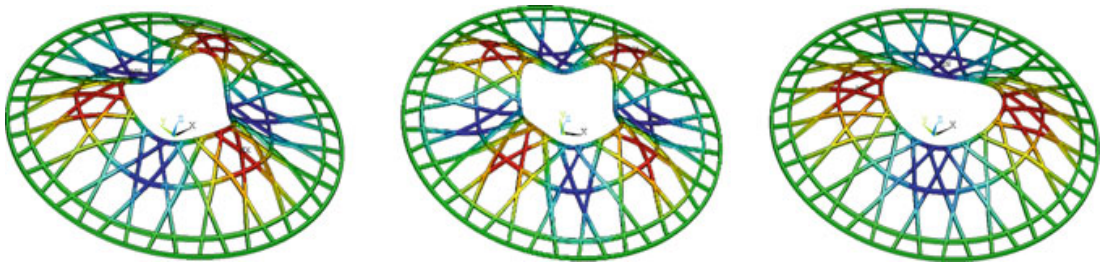


Fig. 83.13 Three mode shapes of the adapter

Table 83.4 Frequency of the mode shapes

Modes	First	Second	Third
Frequency (Hz)	5.09	5.15	6.56

83.8 Conclusion

The results from different models are studied and found that if there is increase in the cross section of the helical groove of carbon fibre, the buckling factors increase. Also, if two circumferential ribs are used, it contributes in increasing buckling factor. It can also be noted that, the bottommost layer in the ply sequence is the critical layer in the helical ply as it is experiencing the highest value of stress (in both the ribs circumferential and helical plies).

The modal analysis shows the structures that have three modes. Since the value of natural frequency is low it can be concluded that the stiffness of the structure is high enough to tolerate the load fluctuations during the actual application as stiffness is inversely proportional to frequency.

Further studies can be made considering the change in winding angle (helical rib angle), varying circumferential rib location, total no. of helical ribs and total no. of helical plies to find out buckling load capacity. The present structure designed is having a weight of approximately 15 kg and weight optimisation studies with higher modulus carbon fibre can be studied.

Acknowledgements A sincere thanks to the Director, ASL and CRMC team for their valuable inputs in the process.

References

1. Vasiliev, V.V., Barynin, V.A., Razin, A.F.: Anisogrid lattice structure-survey of development and application. Compos. Struct. **54**, 361–370 (2001)
2. Gannon, B.B.: Design and analysis of launch vehicle adapter fitting for petite amateur navy satellite (PANSAT)

3. Vasiliev, V.V., Razin, A.F.: Anisogrid conical adapter for commercial space applications. In: AIAA, p. 3440 (2005)
4. Marichela, D.: Advanced grid stiffened structure. A thesis Louisiana State University and Agricultural and Mechanical College
5. Vasiliev, V.V., Barynin, V.A., Razin, A.F.: Anisogrid composite lattice structures-development and aerospace applications
6. Vasiliev, V.V., Razin, A.F.: Anisogrid composite lattice structures for spacecraft and aircraft application. Compos. Struct. **76**, 182–189 (2006)
7. Ansys Products Ltd., version 17.0

Chapter 84
Parametric Study of Axisymmetric Circular Disk Subjected to Rotational Autofrettage

S. M. Kamal and R. Kulsum

Abstract Rotational autofrettage is a potential strengthening technique for thick-walled circular disks/cylinders used in many engineering applications. In particular, the rotating disks used as a rotor in an aircraft engine, rotating impeller wheel of a compressor and fasteners in high pressure pipelines are the important application areas, where rotational autofrettage is useful in mitigating the detrimental tensile stresses during operation. The process induces compressive residual stresses in the disk on unloading a previously applied plastically deforming centrifugal load. The amount of residual stresses generated during the process depends upon different influencing parameters. The geometrical dimension of the disk, rotational speed and material properties are the important parameters to be considered for assessing the effectiveness of rotational autofrettage of a disk. In this paper, a parametric study of the different influencing parameters during the rotational autofrettage of a thick-walled axisymmetric circular disk is carried out.

Keywords Rotational autofrettage · Hydraulic autofrettage · Rotational speed · Residual stress

84.1 Introduction

Usually, autofrettage is performed in thick-walled cylinders to induce compressive residual stresses in the vicinity of the bore surface by a process of loading and unloading of a plastically deforming load applied at the inner surface of the cylinder. The process may also be useful in strengthening thick-walled hollow circular disks such as fastener holes and rotating disks used in automobile and aerospace industries. During the process, the material at the neighborhood of the inner bore surface is subjected to plastic deformation and the material at the outer portion of the cylinder is in the elastic state. The early known autofrettage processes

S. M. Kamal (✉) · R. Kulsum
Department of Mechanical Engineering, Tezpur University, Tezpur 784028, Assam, India
e-mail: smkmech@tezu.ernet.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_84

are hydraulic [1], swage [2] and explosive autofrettage [3]. The hydraulic autofrettage employs a very high internal pressure, whilst the swage autofrettage employs an oversized solid mandrel to be pushed through the bore of the cylinder to achieve the desired plastic deformation in the cylinder wall. The explosive autofrettage utilizes explosive charge to achieve the beneficial effect of autofrettage. The hydraulic and swage autofrettage processes have received considerable attention of the researchers attempting theoretical and experimental studies on the processes [4–11]. The explosive autofrettage process has been less investigated. Recently, researchers have developed some new methods of achieving autofrettage, viz. thermal autofrettage [12] and rotational autofrettage [13] to circumvent certain disadvantages associated with the earlier processes. In thermal autofrettage, a high radial thermal gradient is used for achieving plastic deformation at the inner side of the vessel. Detailed theoretical and experimental studies of the process have been carried out in Refs. [12, 14, 15]. Some recent works on thermal autofrettage have been reported in [16, 17].

Rotational autofrettage is a process where the beneficial compressive residual stresses are generated in the vicinity of a thick-walled disk/cylinder by rotating the component at sufficiently high speed and then gradually reducing its angular velocity to zero [13, 18]. The high angular velocity initiated to the disk/cylinder produces a large centrifugal force plasticizing the inner bore region keeping the region near the outer surface in the elastic state. This phase is called the loading phase. In a later stage when the centrifugal force is made zero by ceasing the rotation of the disk/cylinder, the material near the bore surface is subjected to compressive residual stresses. This phase is called the unloading phase. When the disk/cylinder with compressive residual stresses is put in the actual working environment, it mitigates the magnitude of tensile stresses induced in the component due to service load. This eventually enhances the load-bearing capacity of the component. The rotational autofrettage was first conceptualized by Zare and Darijani [13] in 2016. They carried out a theoretical analysis of the process for thick-walled cylinders assuming plane strain condition. The analysis was based on Tresca yield criterion and its associated flow rule, elastic-perfectly plastic material model with the incorporation of Bauschinger effect. The authors further extended their analysis to incorporate linear kinematic hardening [19]. Most recently, the rotational autofrettage has been studied for thick-walled hollow circular disks by Kamal [18]. He assumed plane stress condition and Tresca yield criterion and its associated flow rule for the analysis. The effect of strain hardening was taken into account using Ludwik's hardening law. The author carried out the numerical simulation of the model considering aluminum and SS304 disks. It was shown that the rotationally autofrettaged disk could enhance its working internal pressure, temperature difference and centrifugal load individually during service. Kamal [18] also carried out a three-dimensional finite element method (FEM) analysis of the rotational autofrettage of a disk. The FEM results were compared with the analytical models, and it was found that the FEM stresses were in close agreement with the theoretical stresses.

It has been observed that there are only a few research papers [13, 18, 19] available in the open literature so far in the area of rotational autofrettage. All the

literatures deal with the theoretical analysis of the process revealing it as a potential process. In any process, the process parameters play an important role in satisfying certain design requirements. For example, in an autofrettage process, the ultimate design requirement is to achieve the compressive residual stress at the inner surface of the component and the corresponding increase in the load carrying capacity as per the actual working condition. In order to achieve the required design objectives in rotational autofrettage process, the process parameters such as the geometrical dimension of the component, rotational speed and material properties should be selected appropriately. No literature is available studying the effect of process parameters on rotational autofrettage. Thus, in this work, a parametric study of the various process parameters in the rotational autofrettage of a thick-walled disk is carried out. The effect of different parameters on the residual stress and on different load-carrying capacity is studied. The study is based on the work of Kamal [18]. Before carrying out detailed parametric study of the process, the plane stress model of Kamal [18] is briefly described in Sect. 84.2.

84.2 Plane Stress Model of Rotational Autofrettage Proposed by Kamal [18]

An axisymmetric disk with inner radius a and outer radius b is considered. Under plane stress condition, the wall of the circular hollow disk splits into two deformed zones, viz. the inner plastic zone and the outer elastic zone when subjected to a plastically deforming centrifugal force by rotating the disk at sufficiently high rotational speed. The details are described in [18]. The radial and hoop stresses induced in the disk due to rotation, when the centrifugal force is not sufficient to deform the disk plastically, are given by

$$\sigma_r = (3+\nu)\frac{\rho\omega^2}{8}\left(a^2+b^2-\frac{a^2b^2}{r^2}-r^2\right) \tag{84.1}$$

$$\sigma_\theta = \frac{\rho\omega^2}{8}\left\{(3+\nu)\left(a^2+b^2+\frac{a^2b^2}{r^2}\right)-(1+3\nu)r^2\right\}, \tag{84.2}$$

where ρ is the density of the disk material, ω is the rotational speed, ν is the Poisson's ratio, a is the inner radius and b is the outer radius of the disk. The yield onset speed of the disk is given by

$$\omega_i = \left[\frac{4\sigma_Y}{\rho\{(1-\nu)a^2+(3+\nu)b^2\}}\right]^{1/2} \tag{84.3}$$

where σ_Y is the yield strength of the material. On increasing the rotational speed beyond ω_i, the material at the inner surface and some portion beneath it deforms

plastically, whereas the outer region remains in the elastic state. Thus, an inner plastic zone, $a \leq r \leq c$ and an outer elastic zone, $c \leq r \leq b$ are formed within the wall of the disk. The radial position, c is known as the radius of elastic–plastic interface. The stresses in the plastic and elastic zones are given by the following equations:

Plastic zone, $a \leq r \leq c$

$$\sigma_r = \frac{c}{r}\left(\frac{c^2-b^2}{b^2+c^2}\right)\left\{\sigma_Y + (1+3v)\frac{\rho\omega^2c^2}{8}\right\} + \sigma_Y\left(1-\frac{c}{r}\right) + (3+v)\left(\frac{\rho\omega^2b^2}{8}\right)\left(\frac{c}{r}\right)\left(\frac{2b^2}{b^2+c^2}-\frac{c^2}{b^2}\right) + \frac{\rho\omega^2}{3}\left(\frac{c^3}{r}-r^2\right) - \frac{K}{r}\int_r^c \left(\varepsilon_{eq}^p\right)^n \mathrm{d}r_1 \tag{84.4}$$

$$\sigma_\theta = \sigma_{eq} = \sigma_Y + K\left(\varepsilon_{eq}^p\right)^n \tag{84.5}$$

Elastic zone, $c \leq r \leq b$

$$\sigma_r = \left(\frac{1+3v}{b^2+c^2}\right)\frac{\rho\omega^2c^2}{8}\left(c^2-\frac{b^2c^2}{r^2}\right) + (3+v)\frac{\rho\omega^2b^2}{8}\left\{1-\frac{c^2}{b^2+c^2}+\frac{b^2c^2}{(b^2+c^2)r^2}-\frac{r^2}{b^2}\right\} + \left(\frac{c^2}{b^2+c^2}\right)\sigma_Y\left(1-\frac{b^2}{r^2}\right) \tag{84.6}$$

$$\sigma_\theta = \left(\frac{1+3v}{b^2+c^2}\right)\frac{\rho\omega^2c^2}{8}\left\{c^2+\frac{b^2c^2}{r^2}-\frac{(b^2+c^2)r^2}{c^2}\right\} + (3+v)\frac{\rho\omega^2b^2}{8}\left\{1-\frac{c^2}{b^2+c^2}-\frac{b^2c^2}{(b^2+c^2)r^2}\right\} + \left(\frac{c^2}{b^2+c^2}\right)\sigma_Y\left(1+\frac{b^2}{r^2}\right) \tag{84.7}$$

The residual stresses induced in the disk are obtained by subtracting the elastic stresses (Eqs. 84.1 and 84.2) from the respective stresses in the plastic and elastic zones. It is assumed that the unloading process is purely elastic devoid of Bauschinger effect. The unknown radius of elastic–plastic interface c can be obtained from Eq. (84.4) by using the boundary condition of vanishing radial stress at the inner radius. A numerical solution procedure such as bisection method may be employed for this. The present parametric study is based on the equations presented in this section.

84.3 Parametric Study

In this section, a parametric study of different process parameters influencing rotational autofrettage of thick-walled disk is carried out. The geometrical parameter such as wall thickness ratio, rotational speed, density and yield stress of the material influences the performance of rotational autofrettage of a thick-walled disk. In the following subsections, the influence of these parameters on the rotational autofrettage of typical thick-walled disks is studied in detail.

84.3.1 Influence of the Wall Thickness Ratio (b/a)

In any autofrettage process, the geometric parameter, wall thickness ratio (b/a) has significant influence on the load requirement, residual stresses and increase in service load-carrying capacity. In the following subsections, the effect of wall thickness ratio on different factors is discussed in detail.

Effect of Wall Thickness Ratio on Rotational Speed In rotational autofrettage, the disk is subjected to a sufficiently high rotational speed in the loading phase. Yielding onsets at a rotational speed ω_{i} at the inner radius of the disk according to Eq. (84.3). Expressing Eq. (84.3) in non-dimensional form, one obtains

$$\Omega_i = \frac{2}{\left\{(1-\nu)+(3+\nu)\left(\frac{b}{a}\right)^2\right\}^{1/2}} \tag{84.8}$$

where Ω_i is the non-dimensional yield onset rotational speed given by

$$\Omega_i = \omega_{\mathrm{i}} a \left(\frac{\rho}{\sigma_Y}\right)^{1/2} \tag{84.9}$$

Figure 84.1 shows the variation of normalized yield onset rotational speed Ω_i with the wall thickness ratio (b/a). It is observed that the initial normalized rotational speed required to cause yielding decreases with the increase of wall thickness ratio. This indicates that for higher b/a ratios, the disk requires comparatively lesser rotational speed for partially plasticizing the wall of the disk during loading phase for attaining the maximum residual stress at the inner side after unloading. For wall thickness ratios up to five, there is a significant reduction in the autofrettage rotational speed requirement. The subsequent increase of wall thickness ratio beyond five provides less than 5% reduction in speed.

Effect of Wall Thickness Ratio on Residual Stress The wall thickness ratio also influences the residual stress induced at the inner surface of the disk. In order to assess the effect of wall thickness ratio on the residual stress, a typical SS304 disk is

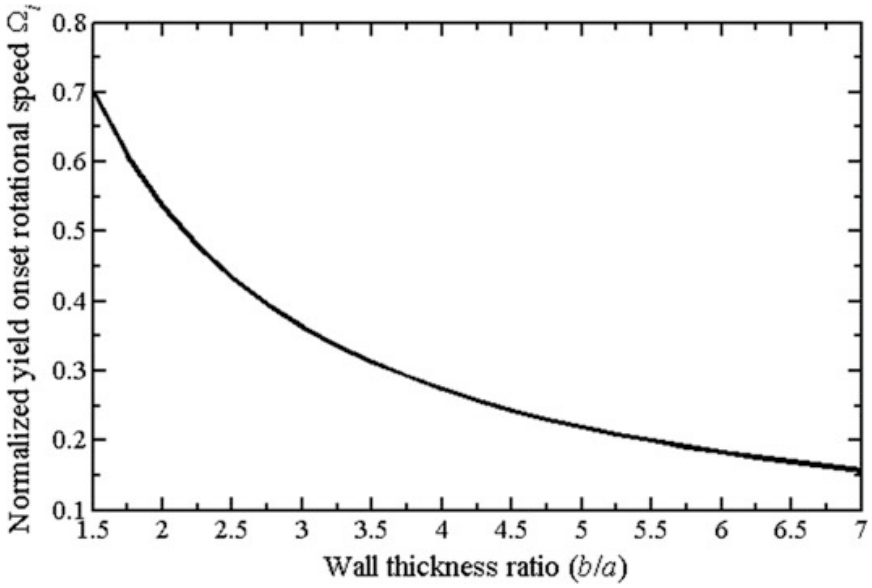


Fig. 84.1 Variation of normalized yield onset speed with wall thickness ratio

considered with the material properties given as follows: Young's modulus of elasticity, E = 193 GPa, yield stress, σ_Y = 205 MPa, Poisson's ratio, ν = 0.3, density, ρ = 8000 kg/m^3. The numerical simulation of the residual stresses is carried out for the disk as described in Sect. 84.2. The detailed expressions are available in Ref. [18]. In Ref. [18], the author has discussed the effect of strain hardening and found that its effect is not very significant. Hence, in the present study, the effect of strain hardening is neglected. Consequently, for numerical simulation, the hardening parameter K is taken as zero in Eqs. (84.4) and (84.5) in Sect. 84.2. The simulated results are presented in Table 84.1 showing the variation of normalized residual hoop stress at the inner and outer radii with the wall thickness ratio (b/a). The residual hoop stress distribution along the positive radial direction of the disk for all (b/a) ratios is also shown in Fig. 84.2.

It is observed from Table 84.1 that as the wall thickness ratio is increased, the normalized compressive residual hoop stress increases at the inner radius. In all cases, the maximum rotational speed of the disk during loading phase is also shown. The maximum rotational speed is considered in such a way that it causes maximum overstrain level (ε)[1] in the disk, and at the same time the residual stresses should not initiate reyielding after unloading. The maximum speed in the loading phase follows an increasing trend with the wall thickness ratio as expected. The normalized tensile residual hoop stress induced at the outer surface of the disk is also presented in Table 84.1 for different wall thickness ratios. It is observed that the normalized tensile residual stress at the outer surface increases up to b/a = 4.5. This indicates that the sensitivity to stress corrosion cracking of the outer surface of

[1]The overstrain level, ε is defined as the percentage of disk's wall thickness that undergoes plastic deformation.

Table 84.1 Variation of normalized residual hoop stress at the inner and outer radii with wall thickness ratio

Wall thickness ratio (b/a)	Normalized maximum rotational speed in the loading phase $\Omega = \omega a\left(\frac{\rho}{\sigma_Y}\right)^{1/2}$	Normalized residual hoop stress $\left(\frac{\sigma_\theta^{Res}}{\sigma_Y}\right)$	
		At the inner surface	At the outer surface
1.5	0.794	−0.2806	0.1906
2	0.654	−0.4863	0.2935
2.5	0.554	−0.6362	0.3432
3	0.480	−0.7510	0.3891
3.5	0.423	−0.8414	0.4212
4	0.377	−0.9101	0.4532
4.5	0.341	−0.9630	0.4219
4.8	0.322	−0.9900	0.4047
5	0.309	−0.9969	0.3218

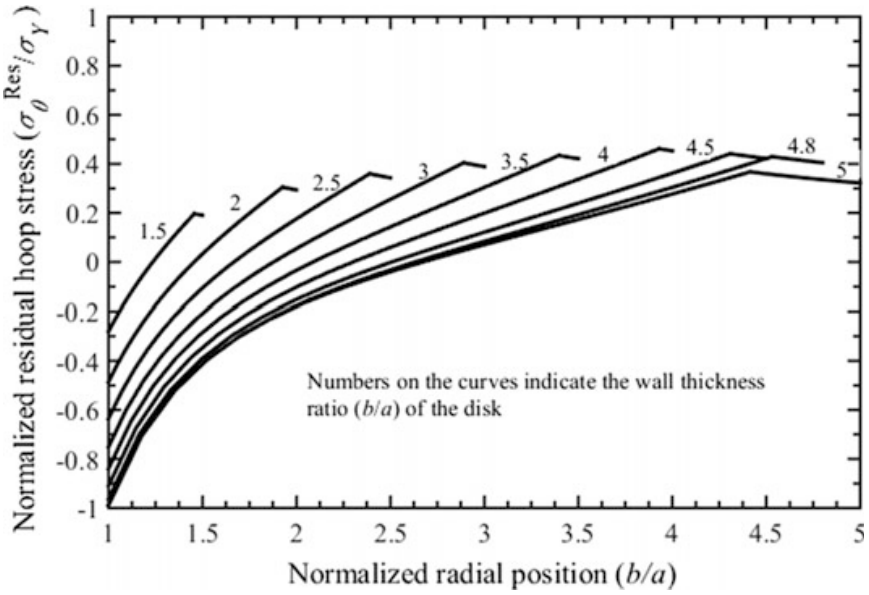


Fig. 84.2 Normalized residual hoop stress distribution in the disk for different wall thickness ratios corresponding to maximum rotational speed

the disk increases with the wall thickness ratio of the disk up to 4.5 for the maximum rotational speed. However, from b/a = 4.8, the maximum normalized tensile hoop stress at the outer surface of the disk starts decreasing. This may be due to the decrease of relative percentage overstrain in the disk. Thus, for the disks with b/a = 4.8 and above, it appears that the maximum compressive residual stress at the inner surface is almost the order of the yield stress of the material. It is observed that the tensile residual hoop stress at the outer surface increases gradually up to

b/a = 4, and thereafter it follows a decreasing trend. The lowering of the tensile hoop stress at the outer surface may be considered as relatively advantageous along with the maximum possible compressive residual stress at the inner surface. The residual hoop stress distribution in Fig. 84.2 shows that the maximum magnitude of the compressive residual hoop stress appears at the inner surface of the disk. It is to be noted that the maximum tensile residual hoop stress occurs at the radius of elastic–plastic interface for all (b/a) ratios. Figure 84.2 shows that the variation of the maximum tensile residual stress follows an increasing trend up to b/a = 4.5, and thereafter it starts decreasing with the increase of b/a.

Effect of Wall Thickness Ratio on Different Load Carrying Capacities The rotationally autofrettaged disk can bear significantly higher loads in service when it is subjected to internal pressure, temperature difference and rotational speed individually [18]. The effect of b/a on the different load-carrying capacities of the disk is also studied corresponding to the cases presented in Table 84.1. In Fig. 84.3, the maximum percentage increase in the service load for different wall thickness ratios is shown when the disk is subjected to internal pressure, radial temperature difference and rotational speed individually. The percentage increase in all service loads is estimated with reference to a corresponding non-autofrettaged disk. The maximum load-carrying capacity of the non-autofrettaged disk is decided based on the respective yield onset load following Tresca yield criterion. For internal pressure and radial temperature difference, the yield onset loads in non-dimensional form are given by [20].

$$\bar{P}_{\mathrm{Y}} = \frac{P_{\mathrm{Y}}}{\sigma_{\mathrm{Y}}} = \frac{1}{2}\left(1 - \frac{a^2}{b^2}\right) \tag{84.10}$$

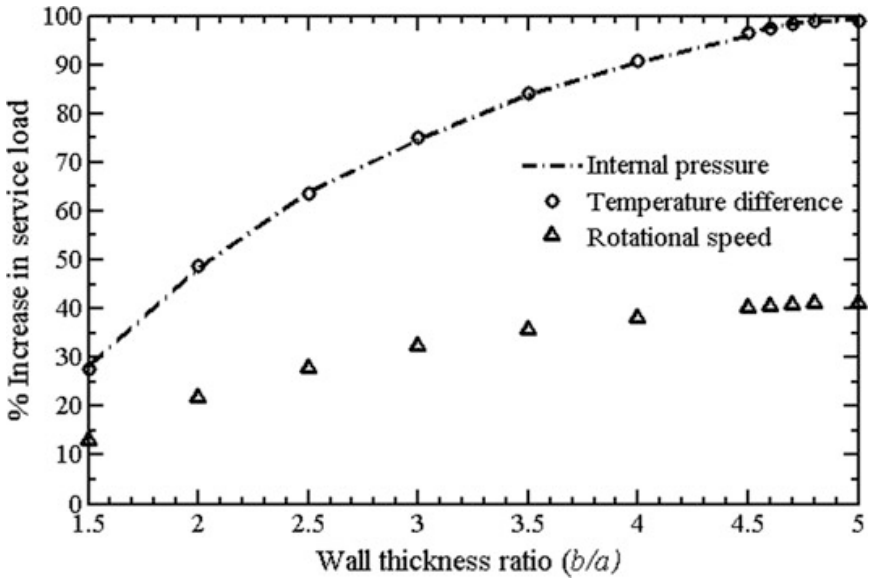


Fig. 84.3 Variation of percentage increase in the maximum different service loads as a function of wall thickness ratio

$$\theta_{\mathrm{Y}} = \frac{E\alpha\Delta T_{\mathrm{Y}}}{\sigma_{\mathrm{Y}}} = \frac{2}{\left\{\frac{\ln\left(\frac{b}{a}\right)-1}{\ln\left(\frac{b}{a}\right)} + \frac{1}{\left\{\left(\frac{b}{a}\right)^2-1\right\}}\left\{\left(\frac{b}{a}\right)^2+1\right\}\right\}} \tag{84.11}$$

where P_{Y} is the yield onset pressure, ΔT_{Y} is the yield onset temperature difference and α is the coefficient of thermal expansion. The yield onset rotational speed is given by Eq. (84.8). The maximum load during service in the autofrettaged disk is obtained by superposing the loading stresses with the rotational residual stresses as discussed in Ref [18] such that the autofrettaged disk does not yield.

It is observed from Fig. 84.3 that the percentage increase in the internal pressure loading and temperature difference of the autofrettaged disk compared to the non-autofrettaged one is of the same order for all wall thickness ratios. The percentage increase in the rotational speed of the disk in service is comparatively lesser than the increase in the other two loads. The percentage increase in all the loads is increased significantly up to the wall thickness ratio 4.6. There is no significant change in the percentage increase of the relative load-carrying capacities as the wall thickness ratio is increased beyond $b/a = 4.6$. Thus, for attaining the maximum load carrying capacity in a disk, the maximum wall thickness ratio can be considered as $b/a = 4.6$. The maximum increase in the internal pressure and the radial temperature difference is about 97% and that in the rotational speed is about 41% during operation corresponding to the wall thickness ratio $b/a = 4.6$.

84.3.2 Influence of Autofrettage Rotational Speed

In rotational autofrettage, the plastically deforming load is in the form of a centrifugal force due to a sufficiently high rotational speed. Thus, the rotational speed plays an important role during rotational autofrettage of a disk. In Table 84.1, it is shown that the maximum rotational speed requirement for rotational autofrettage increases with the increase of wall thickness ratio. Consequently, the residual hoop stress generated at the inner and outer surfaces of the disk is also estimated for all wall thickness ratios up to five. Here, in this section, keeping the wall thickness ratio constant, the effect of rotational speed on the residual stress and the corresponding load-carrying capacity when subjected to different individual loads are assessed. For instance, the wall thickness ratio is taken as three for a typical SS304 disk. The normalized rotational speed for initiation of yielding at the inner side for $b/a = 3$ is 0.363, and the normalized rotational speed for completely plasticizing the disk wall is 0.481. Thus, for all simulations the rotational speed is taken in the range, $0.363 \leq \Omega \leq 0.481$. Within this range, the residual stress at the inner and outer surfaces of the disk is presented in Table 84.2 along with the percentage overstrain (ε) and the percentage increase in different individual loads during operation.

Table 84.2 Effect of rotational speed on percentage overstrain, residual stress and different load-carrying capacities

Normalized rotational speed (Ω)	Percentage overstrain $\varepsilon = \frac{c-a}{b-a} \times 100\%$	Normalized residual stress $\left(\frac{\sigma_\theta^{Res}}{\sigma_Y}\right)$		% increase in the maximum service load		
		At the inner surface	At the outer surface	Pressure	Temperature difference	Rotational speed
0.381	5.50	−0.1036	0.001453	10.30	10.86	5.05
0.399	12.79	−0.2148	0.007219	21.28	21.74	10.22
0.418	22.37	−0.3314	0.02043	33.13	33.69	15.38
0.437	35.07	−0.4533	0.04743	45.32	45.65	20.55
0.456	52.11	−0.5805	0.1033	58.05	58.69	25.72
0.475	78.19	−0.7131	0.2488	71.22	71.73	30.88
0.480	98.79	−0.7537	0.4367	74.51	75	32.43

It is observed from Table 84.2 that as the rotational speed in the loading phase is increased for a constant wall thickness ratio, the percentage overstrain, level of normalized residual stresses at the inner and outer surfaces and the percentage increase in the individual pressure loading, temperature difference across the wall thickness and the rotational speed during working increases. The maximum percentage increase in the respective individual service load is achieved for $\Omega = 0.480$. In order to attain a high rotational speed of the disk, high power/torque motor will be needed in practice. This may increase the equipment cost. However, one can choose the autofrettage speed depending upon the strength requirement of the disk in the actual working environment and accordingly select a motor of required capacity. It is also observed from Table 84.2 that at lower rotational speed, the outer surface of the disk is subjected to very low tensile residual stresses. Thus, at lower rotational speed, the outer surface of the disk is less sensitive to the stress corrosion cracking.

84.3.3 Influence of Material Properties

In rotational autofrettage of thick-walled disk, the properties of the disk material plays a very important role on the rotational speed during loading phase and residual stresses. The yield stress (σ_Y) of the material and density (ρ) are the two dominating material properties which influences the rotational autofrettage of a disk. It follows from Eq. (84.3) that for materials with high density and low yield stress predicts yielding at low rotational speed. In order to assess the effect of ρ and σ_Y, disks of different materials with different combinations of ρ and σ_Y are considered as listed in Table 84.3. The radial dimensions for all disks are kept constant

Table 84.3 Materials with different combinations of ρ and σ_Y

Materials	ρ (kg/m^3)	σ_Y (MPa)
Aluminum	2700	50.3
SS304	8000	205
Ti-0.3Mo-0.8Ni	4510	380
Titanium unalloyed	4510	170
ASTM-A723	7850	1050

and is taken as $a = 0.1$ m and $b = 0.3$ m for instance. The Poisson's ratio for all materials is taken as 0.3.

Now, the rotational autofrettage for all disks is numerically simulated such that the inner surface is subjected to the maximum possible compressive residual stress. The results are presented in Table 84.4. It is observed that in order to generate the maximum possible compressive residual stress ($-0.7537\sigma_Y$ in all disks), the rotational speed increases proportionately for higher yield stress. For the low values of yield stress and density, e.g., aluminum, the rotational speed required to achieve the maximum possible compressive residual stress is lower. Thus, for such materials the speed required for rotational autofrettage can be attained using a low power/low torque electrical motor. This makes the process comparatively inexpensive. However, for high values of yield stress and density, e.g., ASTM-A723, makes inconvenience in achieving the rotational autofrettage in terms of the high speed requirement. For ASTM-A723 disk, the rotational speed required to achieve the maximum possible compressive residual stress is 1756.86 rad/s (16776.78 rpm), which is very high. This requires a high power motor to induce high speed in the disk. Also, in such cases, extra care needs to be taken for dynamic balancing of the rotating system. Materials possessing same density (e.g., Ti-0.3Mo-0.8Ni and titanium unalloyed) require higher rotational speed with the higher yield stress. The materials with medium density and lower yield stress, e.g., SS304 requires moderate speed, which is also convenient for rotational autofrettage.

Table 84.4 Effect of material properties on rotational autofrettage

Materials	ω_i (rad/s)	ω (rad/s)	Compressive residual stress at the inner surface
Aluminum	495.1027	655.66	$-0.7537\sigma_Y$
SS304	580.6643	768.96	$-0.7537\sigma_Y$
Ti-0.3Mo-0.8Ni	1052.9	1394.39	$-0.7538\sigma_Y$
Titanium unalloyed	704.2541	932.64	$-0.7538\sigma_Y$
ASTM-A723	1326.6	1756.86	$-0.7538\sigma_Y$

84.4 Conclusions

In this work, a parametric study of a new method of rotational autofrettage for a thick-walled axisymmetric circular disk is carried out. The study is based on the theoretical analysis of Kamal [18]. The influence of different process parameters on the performance of rotational autofrettage of thick-walled disk is carried out. First, the effect of wall thickness ratio on the rotational speed is studied followed by its effect on the residual stresses and different load-carrying capacities. Then, the effect of rotational speed, material density and yield stress on the residual stress and load-carrying capacities is considered. It is found that as the wall thickness ratio is increased the rotational speed required to cause plastic deformation for achieving rotational autofrettage in the disk is decreased. The rotational speed requirement is the minimum for wall thickness ratio of five with less than 5% reduction in speed upon subsequent increase of wall thickness ratio. It is also found that with the increase of wall thickness ratio, the compressive residual stress induced at the inner surface increases corresponding to the maximum rotational speed. For $b/a = 4.8$ and above, the maximum compressive residual stress induced in the disk is $-0.99\ \sigma_Y$. The study of the influence of wall thickness ratios on the different load-carrying capacities shows that with the increase of wall thickness ratio, the percentage increase in the individual pressure loading, temperature difference and rotational speed of the disk in service increases. It is found that the maximum load-carrying capacities in all the three individual loads is achieved for $b/a = 4.6$ with an error of less than 1%. The percentage in the working pressure loading and temperature difference are of the same order for all wall thickness ratios. However, the percentage increase in the centrifugal load (rotational speed) is lesser compared to the increase in working pressure and temperature difference for wall thickness ratios. The effect of rotational speed for constant wall thickness ratio of the disk shows that on increasing the rotational speed, the residual stresses at the inner and outer surfaces and the percentage increase in the different individual loads during working increases for a typical disk. The yield stress of the material and density also influence the rotational autofrettage of a disk. It is observed that the lower the yield stress of the material, the lower the rotational speed necessary to achieve a certain level of autofrettage. Materials with high yield stress and high density require very high rotational speed to achieve rotational speed. The necessary high speed may be produced by utilizing a high power/high torque motor in such cases. However, this will make the process comparatively expensive. High rotational speed may also cause dynamic imbalance of the system. Thus, rotational autofrettage is convenient for low yield stress and low density materials. Finally, it may be concluded that for the applications of the disks as a rotating component in aircraft and automobile industries or as fastener holes subjected to high pressure/temperature, rotational autofrettage is a very useful strengthening method. For this purpose, the process parameters in the rotational autofrettage of the disk should be selected as per the strength requirement of the disk against the working load during operation.

References

1. Jacob, L.: La Résistance et L'équilibre Elastique des Tubes Frettés. Mémoire de L'artillerie Navale **1**(5), 43–155 (1907). (in French)
2. Davidson, T.E., Barton, C.S., Reiner, A.N., Kendall, D.P.: New approach to the autofrettage of high-strength cylinders. Exp. Mech. **2**, 33–40 (1962)
3. Mote, J.D., Ching, L.K.W., Knight, R.E., Fay, R.J., Kaplan, M.A.: Explosive autofrettage of cannon barrels. Army Materials and Mechanics Research Center, Watertown, Massachusetts, p.02172 Report No. AMMRC CR 70-25. (1971)
4. Rees, D.W.A.: Autofrettage theory and fatigue life of open-ended cylinders. J. Strain Anal. Eng. Des. **25**(2), 109–121 (1990)
5. Gao, X.L.: An exact elasto-plastic solution for an open-ended thick-walled cylinder of a strain-hardening material. Int. J. Press. Vessel Pip. **52**, 129–144 (1992)
6. Parker, A.P., Underwood, J.H., Kendall, D.P.: Bauschinger effect design procedures for autofrettaged tubes including material removal and Sachs' method. ASME J. Press. Vessel Technol. **121**(4), 430–437 (1999)
7. Huang, X.P., Moan, T.: Residual stress in an autofrettaged tube taking Bauschinger effect as a function of the prior plastic strain. ASME J. Press. Vessel Technol **131**(2), 021207-1–021207-7 (2009)
8. Alexandrov, S., Jeong, W., Chung, K.: Descriptions of reversed yielding in internally pressurized tubes. ASME J. Press. Vessel Technol. **138**(1), 011204-1–011204-10 (2016)
9. Iremonger, M.J., Kalsi, G.S.: A numerical study of swage autofrettage. ASME J. Press. Vessel Technol. **125**, 347–351 (2003)
10. Bihamta, R., Movahhedy, M.R., Mashreghi, A.R.: A numerical study of swage autofrettage of thick-walled tubes. Mater. Des. **28**, 804–815 (2007)
11. Rees, D.W.A.: A theory for swaging of discs and lugs. Meccanica **46**, 1213–1237 (2011)
12. Kamal, S.M., Dixit, U.S.: Feasibility study of thermal autofrettage of thick-walled cylinders. ASME J. Press. Vessel Technol. **137**(6), 061207-1–061207-18 (2015)
13. Zare, H.R., Darijani, H.: A novel autofrettage method for strengthening and design of thick-walled cylinders. Mater. Des. **105**, 366–374 (2016)
14. Kamal, S.M., Dixit, U.S.: Feasibility study of thermal autofrettage process. In: Narayanan, R. G., Dixit, U.S. (eds.) Advances in Material Forming and Joining, pp. 81–107. Springer, New Delhi (2015)
15. Kamal, S.M., Borsaikia, A.C., Dixit, U.S.: Experimental assessment of residual stresses induced by the thermal autofrettage of thick-walled cylinders. J. Strain Anal. Eng. Des. **51**(2), 144–160 (2016)
16. Shufen, R., Dixit, U.S.: A finite element method study of combined hydraulic and thermal autofrettage process. ASME J. Press. Vessel Technol. **139**(4), 041204-1–041204-9 (2017)
17. Shufen, R., Dixit, U.S.: An analysis of thermal autofrettage process with heat treatment. Int. J. Mech. Sci. **144**, 134–145 (2018)
18. Kamal, S.M.: Analysis of residual stress in the rotational autofrettage of thick-walled disks. ASME J. Press. Vessel Technol. **140**(6), 061402-1–061402-10 (2018)
19. Zare, H.R., Darijani, H.: Strengthening and design of the linear hardening thick-walled cylinders using the new method of rotational autofrettage. Int. J. Mech. Sci. **124–125**, 1–8 (2017)
20. Kamal, S.M., Dixit, U.S.: A comparative study of thermal and hydraulic autofrettage. J. Mech. Sci. Technol. **30**(6), 2483–2496 (2016)

Chapter 85
Influence of Varying Fully Constrained Circular Boundary Condition Area on Damage Mechanism of GFRP Laminate Under Low Velocity Impact Loading

Mahesh and K. K. Singh

Abstract Modern aircrafts, automobiles, space crafts, etc., need light weight and high-performing structures or materials to reduce the total weight of the structure without compromising in strength which highly influences the efficiency of that structure and fuel cost associated with it. Thus, fiber-reinforced polymer (FRP) laminates are the most suitable materials for these kinds of applications due to their high specific strength to weight ratio which is better than many other conventional metal or alloy materials. But these materials are highly prone to catastrophic and sudden failure because of their unpredictable and undetectable crack growth and formation in the laminate structure. These undetectable cracks, also known as barely visible impact damage (BVID), which are induced into the laminate under low velocity impact. Example includes dropping of a tool in an assembly line. Low velocity impact (LVI) is a highly complex phenomenon as it consists of subsurface micro cracks (fiber and matrix cracking), debonding, delamination, and fiber breakage. Due to anisotropic nature of the material, prediction of damage induced becomes difficult. Hence, in this numerical analysis, LS-DYNA is used to carry out the LVI loading onto a bidirectional plain woven glass fiber-reinforced polymer (GFRP) laminate using a hemispherical striker of mass 20 kg with a nose radius of 5 mm impacted at a velocity of 2 m/s. During the analysis, fully constrained circular boundary condition (FCCBC) is used and the radius of the circular boundary condition (CBC) is varied from 15 to 35 mm with an increment of 10 mm. Effect of variation in boundary condition diameter on damage area is examined using von Mises stresses, energy versus time, force versus time, and displacement versus time graphs.

Mahesh (✉) · K. K. Singh
Indian Institute of Technology (ISM), Dhanbad 826004, Jharkhand, India
e-mail: itsshindhe@gmail.com

K. K. Singh
e-mail: kksingh@iitism.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_85

Keywords GFRP · Low velocity impact (LVI) · Boundary condition · Damage mechanism

85.1 Introduction

FRP composites gained widespread application during the World War-II due to their high strength to weight ratio. Since then, it is used in various industries. Impact is one of the critical loading conditions for FRP composites particularly low velocity impact is fatal because it creates subsurface or barely visible impact damages or nonvisible cracks to the naked eyes. These cracks, over a prolonged period of time grows throughout the laminate and reduces strength of the laminate or structure which leads to sudden or catastrophic failure of the structure. Even though experimental investigation can be carried out to study the damage mechanism, understanding the damage behavior and evolution becomes complex because these materials are highly anisotropic in nature. Moreover, damage mechanism involves various failure mechanisms such as fiber breakage, matrix cracking, delamination, etc. Thus, numerical analysis helps in better understanding of damage mechanism. Damage mechanism and its pattern are influenced by various aspects such as impactor shape, impacting angle, etc., which are all considered as impact loading parameters. One among them is boundary condition because this determines the material space available for absorbing energy from the impactor to the laminate particularly if the boundary condition (BC) is fully constrained clamped type. Although wave propagation in FRP laminate influences the damage pattern and deformation of the material but only at high velocity impact, and this effect is less observed for low velocity impact (LVI).

An extensive literature review was conducted by [1–3] on low velocity impact behavior of FRP laminates and covered various aspects of laminates as well as impact parameters. Moreover, these review works covered the various finite element analysis tools used particularly LS-DYNA. Li et al. [4] studied the GFRP laminates behavior under LVI loading using drop-weight impact machine and simulated the experimental results using ABAQUS FEA tool with edge constrained boundary condition. Both experimental and numerical results showed that central deflection and extent of damage at front as well as back surface behaved linearly with respect to the impact energy. Both numerical and experimental results were reasonably in good agreement. Rawat et al. [5] used LS-DYNA to study the behavior of symmetric and asymmetric GFRP laminates with and without pre-cracks under low velocity impact. From results, it was clear that without pre-crack symmetric laminate performed better than symmetric with pre-crack and asymmetric laminate with and without pre-crack. Author's used circular boundary condition for both symmetric and asymmetric laminate design. Finally, concluded that energy absorption is influenced by symmetric nature of the plies arranged in the laminate about the mid-plane. Rawat et al. [6] investigated the influence of impactor shape on damage area and damage shape under LVI. Eight ply symmetric GFRP

laminate was impacted with flat, hemispherical, spherical, and oval-shaped impactor. LS-DYNA numerical simulation showed that flat-headed impactor caused maximum damage area while oval-shaped impactor caused minimum damage area but it fractured maximum fibers in the laminate. It was concluded that contact surface area with the laminate primarily influenced the damage extent and pattern. Ansari and Chakrabarti [7] examined the influence of boundary condition and impactor nose angle on GFRP laminates under impact loading. Experimental work is carried out on pneumatic gun with flat-and conical-headed impactor and ANSYS was used for numerical simulation. Clamped and simply supported boundary conditions were considered during the analysis. From experimental and numerical analysis, it was concluded that the blunt impactor showed more resistance to penetration under simply supported boundary condition. Liu et al. [8] investigated the satin weave and unidirectional hybrid carbon fiber laminates damage resistance under LVI. Drop-weight impact tester was used for experimentation while ABAQUS was used for numerical validation. From experimental and numerical analysis, it was concluded that damage resistance was affected by lay-up sequence. Delamination was found to be the major failure mode and observed a good correlation between numerical and experimental results. Hashin failure criteria was used by Kim et al. [9] to develop a three-dimensional progressive damage model for GFRP laminate. Developed model was simulated using LS-DYNA and experimental validation was done by conducting impact and flexural testing. Author's concluded that to predict the behavior of impact, interlaminar delamination, and damage progression, three-dimensional progressive failure can be used. Constant Menna et al. [10] investigated GFRP laminates made of different thickness under low velocity impact using LS-DYNA numerical analysis tool. From analysis, it was found that delamination was the major damage mode of failure and majorly observed at mid-plane. Both experimental and numerical analysis found to be in good agreement.

Many available literature is related to behavior of FRP laminate under different boundary conditions, impactor nose shapes, laminate ply orientation, symmetric nature of laminate about the mid-plane, etc. But there is no work that correlates the damage area or energy absorption by the laminate with the boundary condition area. Thus, in this simulation work bidirectional woven glass fibers-reinforced with epoxy is considered. Laminates are made of eight plies with symmetric configuration [11] and impacted with a hemispherical impactor made of mild steel. Circular boundary condition with varying area is considered and LS-DYNA is used to carry out the numerical simulation. To author's knowledge more work is done on different boundary conditions and impactor shapes but a very less work is done on influence of varying boundary condition area and its effect on damage mechanism of FRP laminates under LVI.

Table 85.1 Properties of material used [5]

Young's modulus (GPa)			Poisson's ratio			Tensile strength (GPa)			Compressive strength (GPa)			Shear strength (GPa)		
Laminate properties														
E_{11}	E_{22}	E_{33}	v_{21}	v_{31}	v_{32}	T_X	T_Y	T_Z	C_X	C_Y	C_Z	S_{ba}	S_{ca}	S_{cb}
26	26	8	0.1	0.25	0.25	0.85	0.85	0.12	0.72	0.72	0.5	0.105	0.065	0.065
Impactor properties														
210			0.3											

1, 2 and 3 indicates the material properties with respect to that direction or plane
X, *Y* and *Z* indicates the material properties with respect to that direction or plane
a, b and c indicates the material properties with respect to that direction or plane

85.2 Numerical Simulation

85.2.1 Materials Used

Bidirectional plain woven glass fiber and epoxy laminate are considered for laminates. All the materials properties are provided in Table 85.1 and orientation of the plies is $[(0/90)/(+45/-45)/(+45/-45)/(0/90)]_S$. A mild steel hemispherical-shaped impactor of diameter 10 mm is considered.

85.2.2 Numerical Modeling

MAT_59_SOLID_COMPOSITE_FAILURE_SOLID_MODEL is used to model the GFRP laminate. This card considers the material as orthotropic material and is modeled for enhanced damage modeling. MAT_20_Rigid is used to model the mild steel impactor. Contact between plies and impactor is modeled using AUTOMATIC_SURFACE_TO_SURFACE while TIED_SURFACE_TO_SURFAC- E is used to model the contact between the plies. BOUNDARY_ CONDITION_ SPC is used to model the circular boundary condition with varying radius of 15, 25, and 35 mm. Area of boundary conditions is 706.85, 1963.49, and 3848.45 mm^2. These boundary condition from here onwards are referred as boundary condition 15 (BC 15), boundary condition 25 (BC 25), and boundary condition 35 (BC 35), respectively.

85.3 Results and Discussion

As the boundary condition area is varied, the response time and von Mises stresses generated during the impact are also varied. Effective von Mises stresses (Fig. 85.1) are high for boundary condition 15 at midplies when compared to other two boundary conditions because less material space or region is available to absorb the

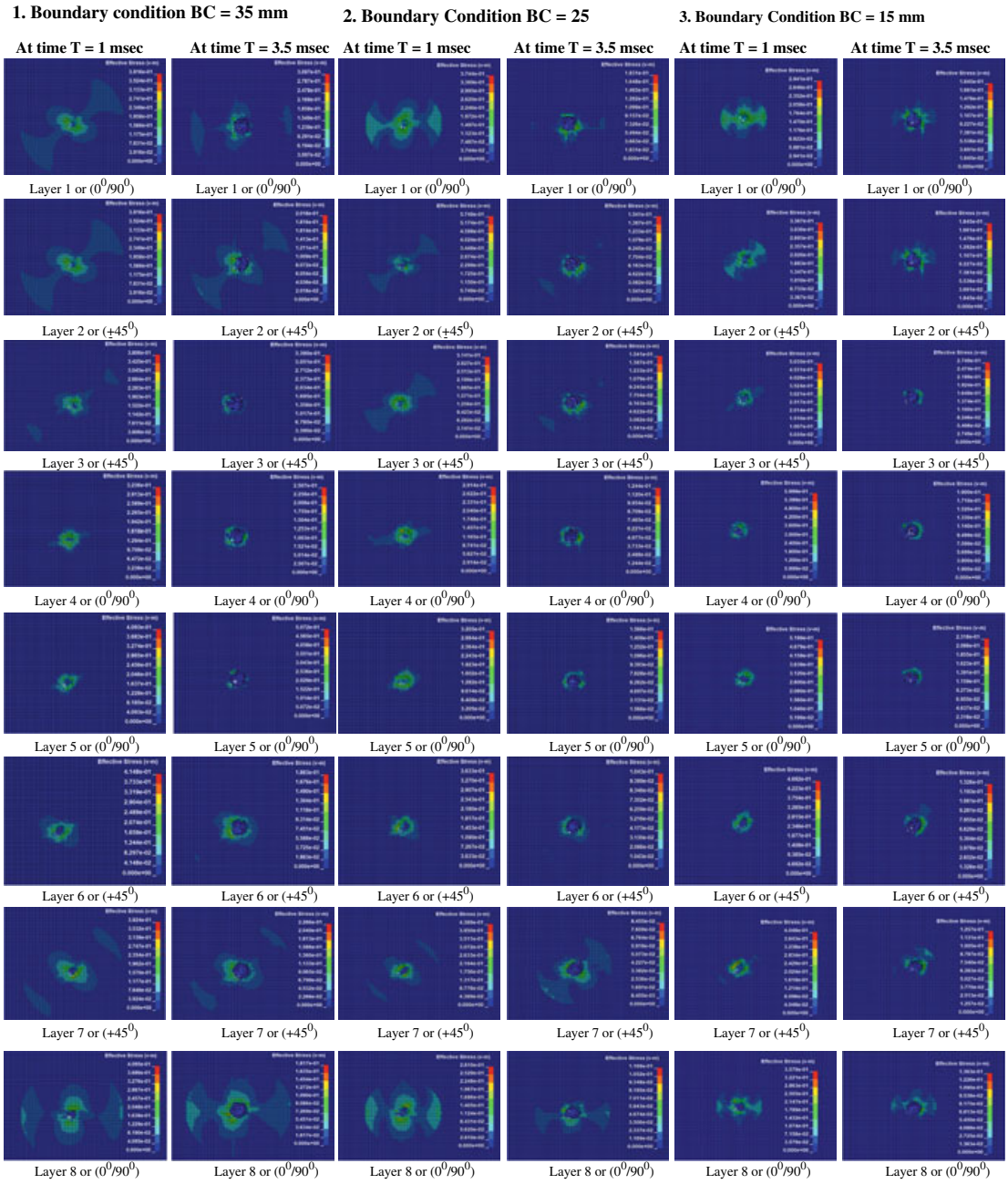


Fig. 85.1 Effective von Mises stresses at time T = 1 ms and T = 3.5 ms

impacted energy. Whereas, at boundary condition 35, the increase in von Mises stresses is gradual because sufficient material space is available for the material to react with the impacted energy. Table 85.2 provides the maximum effective von misses stresses generated in each plies at time T = 1 ms and it is observed that at boundary condition area 35, there is gradual drop in the effective von Misses stresses from top ply to the mid ply and then gradual increase from midply to the bottom ply while exactly opposite is observed in case of boundary condition 25, while a sudden increase in effective von Mises stresses and reaches maximum value at mid ply and then starts declining gradually from mid ply to the bottom ply

Table 85.2 Maximum von Mises stresses under varying boundary condition at time $T = 1$ ms

Ply position in laminate from top to bottom	Boundary condition = 15	Boundary condition = 25	Boundary condition = 35
1	2.941e−1	3.744e−1	4.732e−1
2	3.367e−1	5.749e−1	3.916e−1
3	5.035e−1	3.141e−1	3.806e−1
4	5.999e−1	2.914e−1	3.236e−1
5	5.199e−1	3.205e−1	4.093e−1
6	4.692e−1	3.633e−1	4.148e−1
7	4.048e−1	4.389e−1	3.924e−1
8	3.579e−1	2.810e−1	4.095e−1

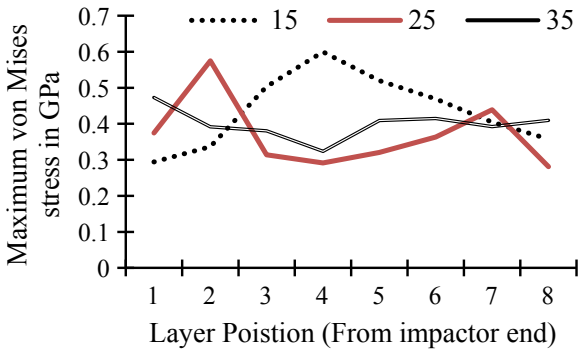


Fig. 85.2 Effective von Mises stresses generated at time $T = 1$ ms

(Fig. 85.2). This kind of behavior is mainly influenced by the boundary condition area and this transaction in behavior is observed from boundary condition 25 which bridges between boundary condition 15 and 35.

Complete fiber fracture is observed in case of boundary condition 15 at time $T = 3.5$ ms, while no such fracture is observed for boundary condition 35 whereas, boundary condition 25 showed mixed behavior as shown in Fig. 85.1. Effective von Mises stresses is high for boundary condition 35 in each ply when compared to boundary condition 15 (Table 85.3). Here, gradual increase and then decrease trend is observed for boundary condition 15, while exactly opposite characteristic is exhibited by boundary condition 35 as seen in Fig. 85.3.

Almost square damage is observed in all boundary conditions. Initially, at time $T = 1$ ms, damage area for boundary condition 15 is high when compared to other boundary conditions, while at time $T = 3.5$ ms damage area is high for boundary condition 35 then boundary condition 25 and boundary condition 15 (Fig. 85.4) because boundary condition 15 at initial stage reacted to impactor with all available material space. But over a period of time, damage area gets saturated for boundary condition 15 while for BC 35, there is enough material space available to react with the impactor to absorb the energy over a period of time through deformation. But boundary condition 15, after saturation absorbed energy is dissipated in terms of damage height, and maximum damage height is also observed for boundary condition 15 (Fig. 85.5 and Tables 85.4, 85.5).

Table 85.3 Maximum von Mises stresses under varying boundary condition at time $T = 3.5$ ms

Ply position in laminate from top to bottom	Boundary condition = 15	Boundary condition = 25	Boundary condition = 35
1	1.845e−1	1.831e−1	3.097e−1
2	2.674e−1	1.541e−1	2.018e−1
3	2.749e−1	1.183e−1	3.390e−1
4	1.900e−1	1.244e−4	2.507e−1
5	2.318e−1	1.566e−1	5.072e−1
6	1.326e−1	1.043e−1	1.863e−1
7	1.257e−1	8.455e−1	2.266e−1
8	1.363e−1	1.169e−1	1.817e−1

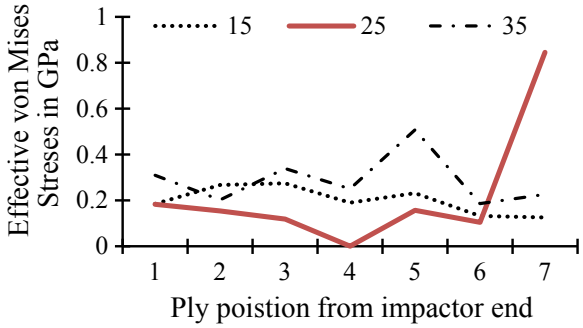


Fig. 85.3 Effective von Mises stresses at time $T = 3.5$ ms

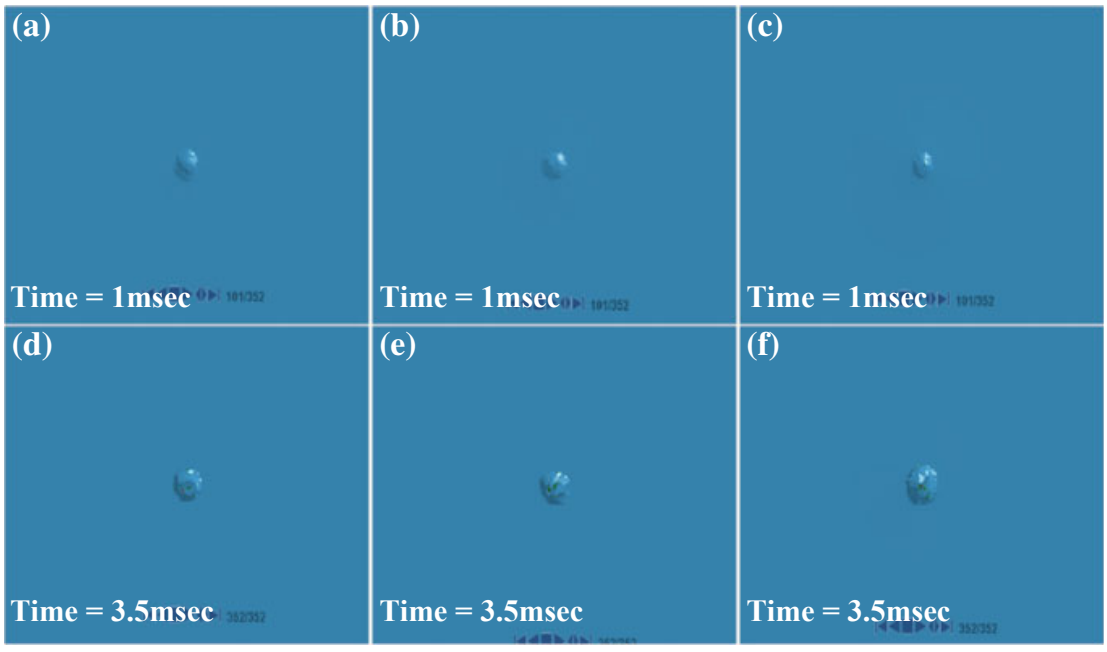


Fig. 85.4 Damage along X-and Y-direction at time $T = 1$ ms **a** and **d** boundary condition 15, **b** and **e** boundary condition 25, and **c** and **f** boundary condition 35

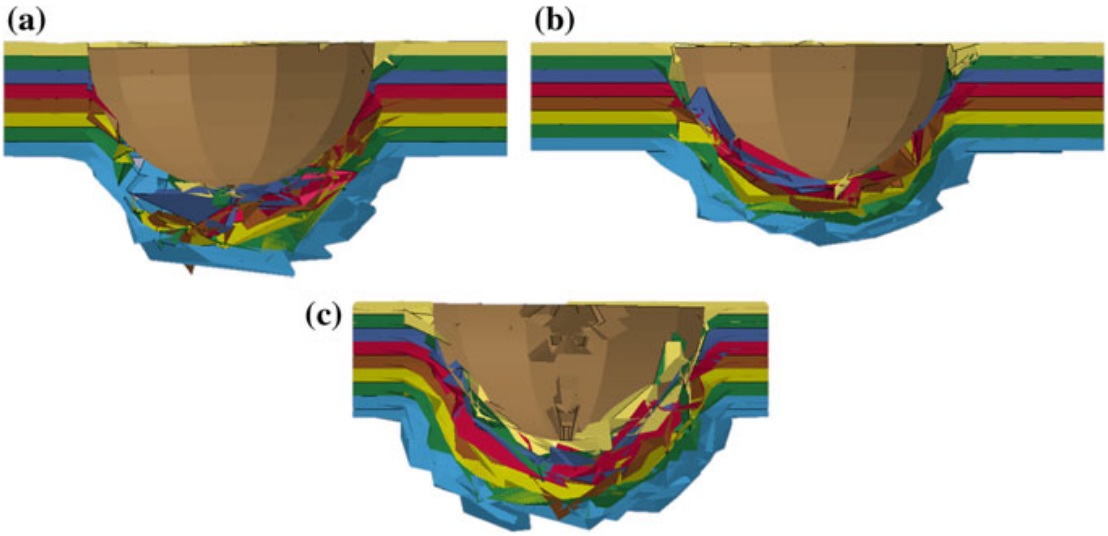


Fig. 85.5 Damage height at time T = 3.5 ms **a** BC 15 **b** BC 25 and **c** BC 35

Table 85.4 Approximate damage area

Time = 1 ms	Damage length in mm		Approx. damage area at base mm^2
Boundary condition	*X*-direction	*Y*-direction	
15	10.135	10.0464	101.8202
25	10.0703	8.2277	82.8554
35	8.2384	8.1574	67.2039

Table 85.5 Approximate damage area and damage height

Time = 3.5 ms	Damage length in mm		Approx. damage area at base in mm^2	Approx. damage height in mm^2
Boundary condition	*X*-direction	*Y*-direction		
15	11.1415	10.167	113.2746	8.3503
25	11.9445	10.2682	122.6485	7.4180
35	13.0267	12.663	164.9571	8.3008

Energy absorption is more for boundary condition 35 than 15 and 25 (Fig. 85.6) but this difference is very small and causes for this variation is material space availability during the impact. Maximum resultant displacement is shown by boundary condition 35 and is more compared to BC 15 and 25. Moreover, BC 15 showed linear displacement because there is no deformation in material instead involves fiber breakage. But BC 25 indicates the transaction form linear displacement to more elongate or bell-shaped deformation (Fig. 85.7). More fiber breakage

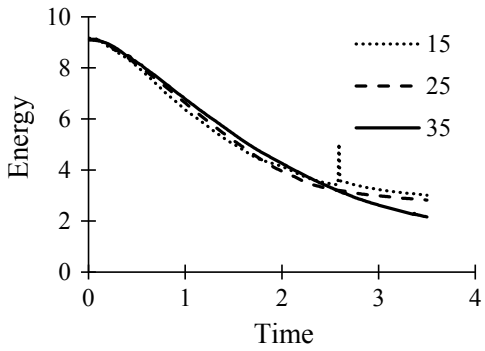


Fig. 85.6 Impact energy (J) versus time (ms)

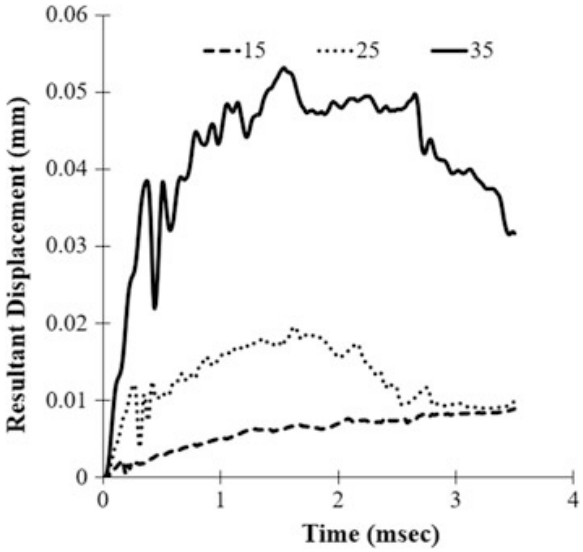


Fig. 85.7 Resultant force (N) versus time (ms)

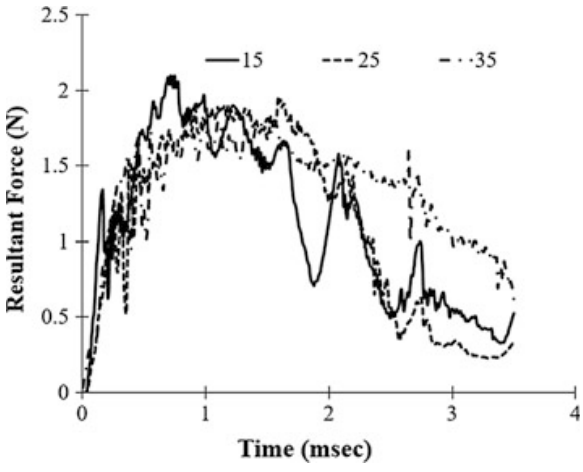


Fig. 85.8 Resultant displacement (mm) versus time (ms)

is observed in case of boundary condition 15 than other boundary conditions (Fig. 85.8). because there is more space available at boundary condition 35 for material to undergo sufficient deformation.

85.4 Conclusion

Influence of boundary condition area is studied for GFRP laminate under LVI using LS-DYNA. Three circular boundary conditions 15, 25, and 35 mm in radius are considered for simulation. Damage length in *X*-and *Y*-direction, damage height, effective von Mises stresses, resultant displacement, energy absorption, and resultant force are evaluated. From numerical simulation results and observation following points were concluded.

- As the boundary condition area reduces, the damage area reduces by 31% at 3.5 ms for boundary condition 15 when compared to boundary condition 35 but fiber breakage and damage height increases in case of boundary condition 15.
- As the boundary condition area reduces, the deformation associated with the material also reduces. Reduction in deformation for boundary condition 15 is found to be 69% less when compared to boundary condition 35 at time $T = 3.5$ ms.
- As the boundary condition reduces, the energy absorption capacity reduces but this reduction is less when compared to higher boundary condition. In case of boundary condition 35 absorbed 9% more energy compared to boundary condition 15.
- As the boundary condition area is reduced from 35 to 15, stress generation is also increased because in case of boundary condition area 15, the material space availability is less compared to 35. Moreover, maximum stress generation is also due to the fiber breakage which is maximum in case of boundary condition area 15 than 35.
- Stresses generated in the material show drastic and sudden increase for small boundary condition area.

References

1. Agarwal, S., Singh, K.K., Sarkar, P.K.: Impact damage on fibre-reinforced polymer matrix composite—a review. J. Compos. Mater. **48**, 317–332 (2014)
2. Singh, N.K., Singh, K.K.: Review on impact analysis of FRP composites validated by LS-DYNA. Polym. Compos. **36**, 1786–1798 (2015)
3. Rawat, P., Singh, K.K.: An impact behavior analysis of CNT-Based fiber reinforced composites validated by LS-DYNA: a review. Polym. Compos **38**, 175–184 (2017)
4. Li, Z., Khennane, A., Hazell, P.J., Brown, A.D.: Impact behavior of pultruded GFRP composites under low-velocity impact loading. Compos. Struct. **168**, 360–371 (2017)
5. Rawat, P., Singh, K.K., Singh, N.K.: Numerical investigation of low velocity impact in symmetric and asymmetric GFRP laminate with and without precrack. Adv. Mater. Proc. **2** (3), 152–155 (2017)
6. Rawat, P., Singh, K.K., Singh, N.K.: Numerical investigation of damage area due to different shape of impactors at low velocity impact of GFRP laminate. Mater. Today Proc. **4**, 8731–8738 (2017)

7. Ansari, M.M., Chakrabarti, A.: Effect of boundary condition and impactor nose angle on impact behavior of FRP composite: experimental and FE analyses. Mater. Proc. **4**(9), 9645–9649 (2017)
8. Liu, H., Falzona, B.G., Tan, W.: Experimental and numerical studies on the impact response of damage tolerant hybrid unidirectional/woven carbon-fiber reinforced composite laminates. Compos. B **136**, 101–118 (2018)
9. Kim, D.H., Jung, K.H., Lee, I.G., Kim, H.J., Kim, H.S.: Three-dimensional progressive failure modeling of glass fiber reinforced thermoplastic composites for impact simulation. Compos. Struct. **176**, 757–767 (2017)
10. Menna, C., Asprone, D., Caprino, G., Lopresto, V., Prota, A.: Numerical simulation of impact test on GFRP composites laminates. Int. J. Impact. Eng. **38**, 677–685 (2017)
11. Singh, K.K., Singh, N.K., Jha, R.: Analysis of symmetric and asymmetric glass fiber reinforced plastic laminates subjected to low-velocity impact. J. od Compos. Mater. **50**(14), 1853–1863 (2015)

Chapter 86
Process Capability Index for Geometrically Distributed Quality Characteristics

N. K. Singh

Abstract Process capability indices play a very important role in the control and the improvement of quality characteristics of products during the production process. Various process capability indices have been developed for quality characteristics which follow normal, Weibull, exponential and other distributions. Many researchers have also developed process capability indices for quality characteristics which follow bivariate normal and exponential distributions, respectively. Some of these indices are being used by industries for the quality improvement of products during the production process for competing in global markets. Some quality characteristics of the products should be treated as discrete random variables following distributions like geometric distribution rather than widely known continuous distribution like normal. Such situations are encountered while dealing with quality characteristics like number of operations before the failure of electrical, electronic or mechanical switches and components. The control and the improvement of such quality characteristics of such products during the production process, in order to ensure the conformity to the specification (usually one-sided), are quite important. No much work has been done so far to develop process capability index applicable to such quality characteristics following the geometric distribution. In this paper, the author proposes a process capability index for a number of successes or operations before the first failure of a component or product assumed to follow a geometric distribution and derives the expectation and variance of the estimated index. The optimal choice of process capability interval related to the said index has also been discussed. The said process capability index may be low for a highly capable process and for the minimum variation of the index, the percentage area lying outside natural tolerance limit may be low as 0.02, and the probability of successes or operations before the first failure may be high as 0.99.

Keywords Process capability index · Geometric distribution · Quality characteristics · Sampling properties

N. K. Singh (✉)
Indian Institute of Technology (Indian School of Mines), Dhanbad 826004, India
e-mail: nirmal@iitism.ac.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_86

86.1 Introduction

Process capability index C_p is a simple relative number or numerical value comparing the value of required process variability (required tolerance interval) to natural process variability (natural tolerance interval). Kane [1] introduced the process capability index C_p, also called process potential index, which considers spread only.

$$C_p = \frac{(U - L)}{6\sigma} \tag{86.1}$$

$$C_p \geq 1 \quad \text{For a capable process}$$

where U and L are upper and lower specification limits, respectively.

However, C_p focuses on the dispersion of the studied process and does not take into account the centering of the process. To overcome this problem, Kane [1] introduced the process performance index, C_{pk} as,

$$C_{pk} = \min\left\{\frac{\text{USL} - \mu}{3\sigma}, \frac{\mu - \text{LSL}}{3\sigma}\right\} \tag{86.2}$$

where μ is the process mean and σ is the process standard deviation of the in-control process.

The reason for six in the denominator in Eq. (86.1) is that under the assumption of normality, 6σ spread covers 99.73% of the data, leaving only 0.27% in non-conforming range. Larger values of C_p indicate high process capability while smaller values (less than 1.0) indicate lower process capability and incapable process.

The indices C_p and C_{pk} do not take into account the target value. Chan et al [2] point out that the index while ensuring that the process spread is within the specification limits, fails to take into account the target value, T of the process. To correct this, the authors suggested a modified version of C_p, Thus, this index is defined as

$$C_{pm} = \frac{U - L}{6\tau} = \frac{C_p}{\sqrt{1 + \left(\frac{\mu - T}{\sigma}\right)^2}} \tag{86.3}$$

where $\tau^2 = \sigma^2 + (\mu - T)^2$

It is obvious that C_{pm} is smaller than C_p.

The index C_{pm} was also proposed independently by Taguchi and Hsiang [3]. C_{pm} is, sometimes, called Taguchi capability index.

A capability index, which is more sensitive than C_{pk} and C_{pm} with regards to departures from the process mean to the target value, was introduced by Pearn et al. [4] as

$$C_{pmk} = \min\left\{\frac{\text{USL} - \mu}{3\sqrt{\sigma^2 + (\mu - T)^2}}, \frac{\mu - \text{LSL}}{3\sqrt{\sigma^2 + (\mu - T)^2}}\right\} \tag{86.4}$$

A superstructure or family of capability indices, containing C_p, C_{pk}, C_{pm} and C_{pmk} involving two parameters, is introduced by Vannman [5] as,

$$C_p(u, v) = \frac{d - u|\mu - m|}{3\sqrt{\{\sigma^2 + v(\mu - T)^2\}}} \tag{86.5}$$

where $d = \frac{U-L}{2}$ and $m = \frac{U+L}{2}$

By letting $u = 0$ or 1 and $v = 0$ or 1 in the above equation, we can obtain four basic indices, i.e., $C_p(0,0) = C_p$, $C_p(1,0) = C_{pk}$, $C_p(0,1) = C_{pm}$ and $C_p(1,1) = C_{pmk}$.

Wang and Hubele [6] discussed geometric distance approach to reduce the dimensionality of the multivariate (normal and non-normal) data to univariate data and used established univariate process capability index (PCI) for process performance analysis.

Castagliola [7] extended the paper of Castagliola and Castellanos [8] for the estimation of bivariate PCIs in the case of non-normal bivariate distributions. The proposed method is based on the use of Johnson's distribution/transformation in order to transform the bivariate non-normal distribution into an approximate bivariate normal distribution. Numerical examples have been presented and some criteria are given in order to choose the appropriate Johnson's distribution.

Raut and Attar [9] provided a case study based on process capability in the vehicle industry. In this case study, area of analysis was the measurement section for which they selected the body shop production. The main highlights of the case study were to go through a comprehensive explanation of the entire body shop and the departments covered in it.

Senvar and Sennaroglu [10] examined Clements' Approach (CA), Box-Cox transformation (BCT) and Johnson transformation (JT) methods for process capability evaluations through Weibull-distributed data with dissimilar parameters to figure out the effects of the tail behaviors on process capability and matched their performances in terms of accuracy and precision.

In Sect. 86.2 of this paper, the author proposes a process capability measure for the number of success or operations before the failure assumed to follow a geometric distribution and derives the expectation and sampling variance of the estimated index. The optimum choice of process capability interval has been discussed in Sect. 86.3.

86.2 Process Capability Index for Geometric Distribution

*86.2.1 Process Capability Index (*Ig*)*

The probability density function (pdf) for geometric distribution is given by

$$f(x) = qp^x \tag{86.6}$$

$$\text{Let } \sum_{x=0}^{y} f(x) = 1 - p^{y+1} = \alpha \tag{86.7}$$

where x is the number of operations/successes before the first failure of the equipment or component, and p and q are probabilities of success and failures, respectively. α is the percentage area lying outside natural tolerance limit.

The natural process interval may be taken as (y, α) with the probability content $1 - \alpha$.

From Eq. (86.7), we get

$$y = \frac{\log\left(\frac{1-\alpha}{p}\right)}{\log p}$$

Then the process capability index (Ig) can be proposed as

$$Ig = \frac{L}{y} = \frac{L.\log p}{\log(\frac{1-\alpha}{p})} \tag{86.8}$$

Assuming y, the number of successes/operations follows a geometric distribution, L is the lower specification limit and U, the upper specification limit is ignored. It is also assumed that there is no measurement error.

It may be noted that process capability indices, e.g., C_p, C_{pk}, C_{pm} *and* C_{pmk}, mentioned in Sect. 86.1 should always be greater than unity for a capable process, but in Eq. (86.8) Ig may be taken less than unity for a capable process.

The sample estimate of Ig comes out to be

$$\hat{I}g = \frac{L}{\widehat{y}} = \frac{L.\log \hat{p}}{\log\left(\frac{1-\alpha}{\hat{p}}\right)} = LG(\hat{p}), \quad \text{say} \tag{86.9}$$

where $\widehat{y}$ = Sample quantile of order α

86.2.2 Estimation of $\hat{\mathbf{I}}\mathbf{g}$

We have (from Johnson and Kotz [11]).

$$E(x) = \frac{p}{q}$$

From Eq. (86.6)
Maximum likely estimate $m.1.e$ of $\hat{p}$ comes out to be,

$$\hat{p} = \frac{\bar{x}}{\bar{x}+1},$$

where $\bar{x}$ = Average number of successes before the first failure.
Then putting this value of $\hat{p}$ in Eq. (86.9), it is found that,

$$\hat{I}g = \frac{L.\log(\frac{\bar{x}}{\bar{x}+1})}{\log\left\{\frac{(1-\alpha)(\bar{x}+1)}{\bar{x}}\right\}} = L\psi(\bar{x}), \quad \text{say} \tag{86.10}$$

86.2.3 Sampling Properties of $\hat{\mathbf{I}}\mathbf{g}$

Bias of $\hat{Ig}$
From standard results in large sample theory (see, e.g., Kendall and Stuart [12]).

$$E(\hat{I}g) = L.E\{\psi(\hat{p})\} \approx L.\psi(p) = Ig$$

Thus, $\hat{I}g$ is asymptotically unbiased for Ig.

Sampling variance of $\hat{Ig}$
It is known that

$$\mathrm{Var}(\hat{I}g) = L\left(\frac{\delta\psi}{\delta\bar{x}}\right)^2 \mathrm{Var}(\bar{x}) \tag{86.11}$$

$$\bar{x} = \frac{p}{q}$$

where $\mathrm{Var}\,(\bar{x}) = \frac{\mathrm{Var}_x}{n} = \frac{1}{n}\cdot\frac{p}{q^2}$

On differentiation of $\psi(\bar{x})$ in Eq. (86.10) with respect to $\bar{x}$, it is found that

$$\frac{\delta\psi}{\delta\bar{x}} = \frac{L.\log(1-\alpha)}{\bar{x}(\bar{x}+1)\left[\log\left\{\frac{(1-\alpha)(\bar{x}+1)}{\bar{x}}\right\}\right]^2}$$

On putting this value in Eq. (86.11) and on simplification, it is found that

$$\mathrm{Var}(\hat{I}g) = \frac{L^2}{n}\frac{(1-p)^2}{p}\frac{\{\log(1-\alpha)\}^2}{\left\{\log\left(\frac{1-\alpha}{p}\right)\right\}^4} \tag{86.12}$$

86.3 Optimum Choice of α

Usually, the percentage area α lying outside natural tolerance limit will be given from assurance (confidence) requirement. However, one may, even for the sake of the theoretical interest, try to choose the value of α such that, for a given process parameter p, will result in the most stable estimate of process capability index *Ig* (in repeated sampling).

On differentiation of Var $(\hat{I}g)$ in Eq. (86.12) with respect to α and equating it to zero, it is found that for minimum variation of Var $(\hat{I}g)$.

α = 0 or 1 which are not feasible solutions.

So, the variance minimizing choice of α depends on *p* and vice versa.

From Eq. (86.12) and Table 86.1, on numerical analysis, it is observed that when α increases, Var $(\hat{I}g)$ decreases, *p* remaining constant. It means that the variation in the process capability may be less when the percentage area lying outside natural tolerance is more resulting in less Nos. of successes or operations of products. It is also observed from Table 86.2 that when *p* increases Var $(\hat{I}g)$ increases, α remaining constant. It means that the variation in the process capability may be more in order to achieve a high probability of successes or operations. More efforts may be needed for the improvement of the process to achieve high Nos. of successes or operations of products.

So, for the minimum variation of process capability index *Ig*, α may be chosen as low as possible, corresponding to a fixed *p*. For example, choice of α may be chosen as 0.02 if *p* (or its estimate) comes out to be 0.99.

Table 86.1 Values of Var $(\hat{I}g)$ corresponding to α at constant *p*

Values of α at $p = 0.99$	(n/L^2) x Var $(\hat{I}g)$
0.02	19.787
0.03	2.864
0.04	0.983
0.05	0.487

Table 86.2 Values of variance Var ($\hat{Ig}$) corresponding to p at constant α

Values of p at $\alpha = 0.02$	$(n/L^2) \times$ Var ($\hat{Ig}$)
0.99	19.787
0.985	736.95
0.97	181.43
0.96	19.95
0.95	6.099
0.94	2.747

86.4 Typical Applications

The process capability index (Ig) may be applied in finding out capability or variability of process or system having Nos. of flights or Nos. of landings before the first failure after planned repair and preventive maintenance of an aeroplane. This may also prove to be helpful in taking decision on the variability of the process for the production of electrical, electronics and computer components and hardware in order to increase their probability of success (p) resulting in their reliability as required in reliability-centered maintenance in special case.

86.5 Conclusion

The process capability index (Ig) may be less than unity and positive for a capable process for the quality characteristics following the geometric distribution and for the minimum variation of Ig, the percentage area (α) lying outside natural tolerance limit may be low as 0.02 and probability of successes or operations before the first failure may high as 0.99.

Acknowledgements I am very much thankful to Prof. (Retd.) S. P. Mukherjee former Centenary Professor of Statistics Department of Statistics and former Dean Faculty of Science Calcutta University Kolkata and presently Mentor of Indian Association for Productivity, Quality and Reliability Kolkata for constant inspiration for research.

References

1. Kane, V.: Process capability indices. J. Qual. Technol. **18**, 41–52 (1986)
2. Chan, L.K., Cheng, S.W., Spiring, F.A.: A new measure of process capability C_{pm}. J. Qual. Technol. **30**, 162–175 (1988)
3. Taguchi, G., Hsiang, T.C.: A tutorial on quality control and assurance—the Taguchi methods. Presentation given at the Annual Meeting of the Americal Statistical Association, Las Vegas, LV (1985)
4. Pearn, L.J., Kotz, S., Johnson N.L.: Distributional and inferential properties of process capability indices. J. Qual. Technol. **24**, 216–223 (1992)

5. Vannman, K.: A unified approach to capability indices. Research report, division of quality technology and statistics. Luleo University, Sweden. Stat. Sin. **3**(2) (1993)
6. Wang, F., Hubele, N.: Quality evaluation using geometric distance approach. Int. J. Reliab. Qual. Saf. Eng. **6**, 139–153 (1999)
7. Castagliola, P.: Evaluation of non-normal process capability indices using Burr's distributions. Qual. Eng. **8**, 587–593 (1996)
8. Castagliola, P., Castellanos, J.: Capability indices dedicated to the two quality characteristics case. Qual. Technol. Quant. Manage. **2**(2), 201–220 (2005)
9. Raut, R., Attar, P.R.: Optimization of process capability in an automobile industry: a case study. Int. Res. J. Eng. Technol. **3**(7) (2016)
10. Senvar, O., Sennaroglu, B.: Comparing performances of Clements, box-cox, Johnson methods with Weibull distributions for assessing process capability. J. Ind. Eng. Manag. **9**(3), 634–656 (2016)
11. Johnson, N. L., Kotz, S.: Distributions in Statistics, vol. I (Discrete Distribution). Wiley, New York, NY (1969)
12. Kendall, M.G., Stuart, A.: The Advanced Theory of Statistics, vol. 1, pp. 228–245. Charles Griffin, London (1973)

Chapter 87
Thermomechanical Analyses of Single Sided Single Pass Submerged Arc Welding of AISI 304 Austenitic Stainless Steel

P. V. S. S. Sridhar, Vishnu Nair, Pankaj Biswas and Pinakeswar Mahanta

Abstract The submerged arc welding (SAW) is an automated process that is capable to weld thick and thin plates of steel and non-ferrous materials due to its ease of operation, less spatter and high deposition rates. SAW is extensively used to fabricate bridges, pipe lines, pressure vessel construction, ship building and surface cladding. In present study numerical simulation on effect of welding speed on residual stresses for SAW with square butt geometry of austenitic stainless steel is explored. A moving heat source model considering heat generation is used in this present study. The element birth and death technique is used to incorporate the effects of filler metal deposition. The peak temperatures observed in experimentation were correlated with that of numerical results and the variation was observed to be 4.77%. At higher welding speeds the low peak temperatures were observed which is due to less heat penetrating into the joint. It was observed in the present study at higher welding speeds the residual stress and von Mises stress were in increasing trend whereas normal and transverse stresses are negligible along the weld line.

Keywords Submerged arc welding · Residual stress · Austenitic stainless steel

P. V. S. S. Sridhar (✉) · P. Biswas · P. Mahanta
Department of Mechanical Engineering, Indian Institute of Technology Guwahati, Assam 781039, India
e-mail: p.sridhar@iitg.ac.in

P. Biswas
e-mail: pankaj.biswas@iitg.ac.in

P. Mahanta
e-mail: pinak@iitg.ac.in

V. Nair
Engineering Research Centre, TATA Motors Ltd., Jamshedpur 831004, Jharkhand, India
e-mail: vsnunair@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_87

87.1 Introduction

The submerged arc welding is a popular production process due to its ease of automation, high deposition, deeper penetration, no spatter and fumes and ability to produce sound quality welds in thin and thicker section of materials. The process is extensively used to weld ferrous metals including all varieties of stainless steels and nonferrous alloys like copper and titanium [1]. In fusion welding heat is applied introduced through a heat source which is responsible for fusion between the plates to be joined, in addition to that several phenomena like distortion, residual stresses and metallurgical structural changes are inevitable as a result of applied heat. In order to get a sound weld, it is necessary to understand the effect of heat flow in a welding process. The residual stresses, distortion, microstructure and HAZ hardness introduced into the weld have direct bearing on the applied welding power. The local heating and cooling occurring during the welding process affects the structural changes in the material. Hence, the knowledge of heat flow and its resultant effects need to be predicted for establishing optimum welding parameters.

The heat source from which heat flows into the material is different for various welding processes. The analytical model and its solution for heat flow during welding was first, introduced by Rosenthal. In his famous paper, he assumed that conduction is the predominant mechanism of heat transfer, while assuming the shape of the weld pool in 2D and 3D [2, 3]. However, the Rosenthal model's yields poor predictions. The model developed by Ohring and Lugt [4] for gas metal arc welding (GMAW) considered all possible physical phenomena like evaporation, heat transfer by convection, radiation and viscous deformable free-surface conditions. The main drawback of his model is they assumed constant material properties. Brown and Song [5] considered 3D model of heat source for better understanding of the underlying phenomena. The essential parameters that are to be incorporated in the simulation, a moving heat source, welding parameters like speed, voltage and current, deposition of filler material into the joint. Goldak et al. [6] proposed double ellipsoidal heat source and incorporated the size and shape of heat source for shallow and deeper penetrations welds. Biswas and Mandal [7] formulated a numerical model and macros to simulate a distributed moving heat source, and also considered for deposition of filler metal in submerged arc welding for predicting weld distortions and residual stresses. Podder et al. [8] developed a generalized solution for a heat source model for submerged arc welding of low carbon and M.S. plates. The physical phenomena like forces of buoyancy, Lorentz and Marangoni were taken into account while formulating the model. The principal objective of modeling of welding heat source is to represent the true source of heat, which is the welding arc. The heat is dissipated from the weld pool by conduction, convection and radiation. In the present work effect of process variable i.e., welding speed on residual stress in the austenitic stainless steel is explored.

87.2 Experimental Details

The temperature measurement during was carried out using *K*-type thermocouples. This method was used during the submerged arc welding of butt welded single pass single sided joints from 304 austenitic stainless steel and the results were compared with those of finite element method. The thermocouples were located on the top surface of the plate using a thermocouple attachment setup. Temperatures were measured at two distances from the weld melt line on one side of the plates. The location of thermocouple was 17 mm from weld center line. Figure 87.1 Schematic of test specimen with location of thermocouple. The dimensions of weld sample are 200 mm × 100 mm × 8 mm. A data acquisition and controller (Make: AGILENT 34972A) is used to collect the signals, which was interfaced to a PC. It was set to collect temperature data for interval of 500 ms. Figure 87.1 shows Schematic of test specimen and actual SA welding.

The work pieces were welded in square butt condition, in a single pass with reverse polarity (DCEP). After conducting several trial runs, the welding parameters were decided. The voltage (*V*), current (*I*) and travel speed (*S*) in each joint were given in Table 87.1. The arc efficiency (η) was taken as 0.85 [9, 10] and was used to calculate the heat input per mm Q_w.

$$Q_w = \frac{\eta VI}{S} \tag{87.1}$$

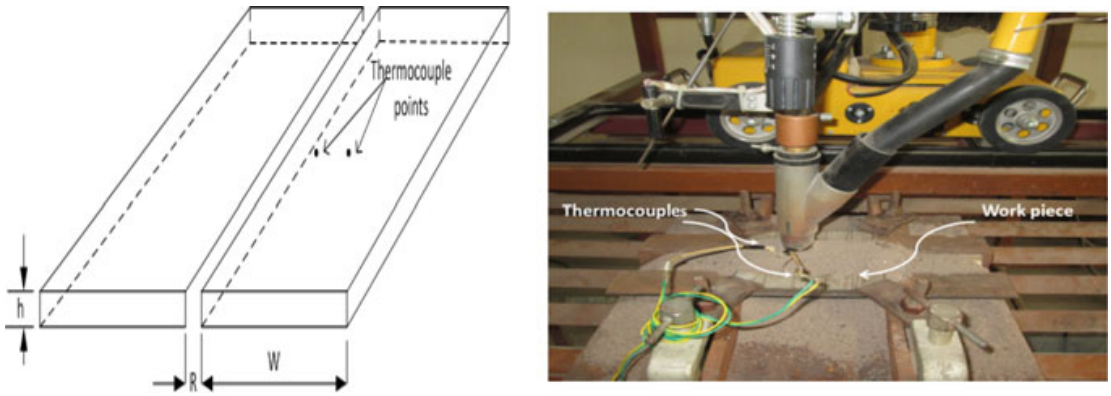


Fig. 87.1 Schematic of test specimen with location of thermocouple and SAW with thermocouples attached to the test specimen

Table 87.1 Parameters considered for analysis of variation of speed

Exp.	Welding arc power (W)	Thickness (mm)	Speed (m/hr)
1.	24 V × 430 A = 10,320	8	20
2.	24 V × 430 A = 10,320	8	22
3.	24 V × 430 A = 10,320	8	24

87.3 Thermomechanical Analysis

The present study involves transient thermal analysis to predict the temperature distribution in a welded sample and mechanical analysis to study structural impacts on the weld caused due to thermal load. A three dimensional finite element model was developed to study thermos-mechanical behavior of weld.

87.3.1 Assumptions for Thermo-Mechanical Analysis [7]

Following are the assumptions made to develop the model for welding process:

- All the thermal properties are considered as a function of temperature.
- No forced convection is considered unless mentioned.
- Linear Newtonian convection cooling is considered on all the surfaces except the weld zone.
- Radiation loss being small, the same is neglected.
- Arc efficiency is incorporated to account for other losses.
- Phase change is not considered.

87.3.2 3D Finite Element Model

A three dimensional finite element model was developed using ANSYS 18.0 to study the temperature distribution and heat transfer in a SAW process. In the thermal model actual conditions were incorporated as far as possible. The governing differential equation for heat conduction for a homogenous, isotropic solid in the rectangular coordinate system (x, y, z) is:

$$\frac{\partial}{\partial x}\left[K\frac{\partial T}{\partial x}\right]+\frac{\partial}{\partial y}\left[K\frac{\partial T}{\partial y}\right]+\frac{\partial}{\partial z}\left[K\frac{\partial T}{\partial z}\right]=\rho C\frac{\partial T}{\partial t} \tag{87.2}$$

where ρ = density of plate material (kg/m^3), C = specific heat (J/kg °C) and K = thermal conductivity (W/m °C). The initial condition and boundary conditions applied on the model are mention below.

Initial Condition
An initial temperature was assumed for the entire plate equivalent to ambient temperature.

$$T = T_\infty \quad \text{for } t = 0$$

where T_∞ is the ambient temperature.

Boundary Condition
As the heat flows over the weld surface, the same is conducted to the rest of the plate. The following equation represents the same:

$$q_n = -q_{\text{sup}} \tag{87.3}$$

Here q_{sup} represents the heat supplied from an external welding arc. And q_n represents the component of the conduction heat flux vector normal to the work surface.

As the heat losses due to convection from the plate surface. Considering h_f as convection coefficient, the heat loss q_{conv} is evaluated as:

$$q_{\text{conv}} = h_f(T - T_\infty) \tag{87.4}$$

Structural Analysis
The residual stress and plastic strain were evaluated post heat transfer analysis where in temperature distribution in heat affected zones was studied. Tack weld condition was applied and using the results of thermal analysis the structural analysis was carried out. The following are the formulations for stress-strain evaluation used in the current study:

$$\nabla\sigma + F = 0 \tag{87.5}$$

where

∇ Differential operator
σ Stress
F Body force

The stress strain relationship is as follows:

$$\{\sigma\} = [D]\{\varepsilon\} \tag{87.6}$$

where

$\{\sigma\}$ Stress tensor
$[D]$ Stiffness matrix
$\{\varepsilon\}$ Strain tensor

The von Mises yield criterion is used to analyze the stress and strain behavior post thermal analysis.

87.4 Modelling Details

The model was developed in ANSYS 18.0. For transient thermal analysis an eight-node SOLID70 element was considered. As the results of thermal analysis was further used to determine the residual stress and strain induced in the plate, thus

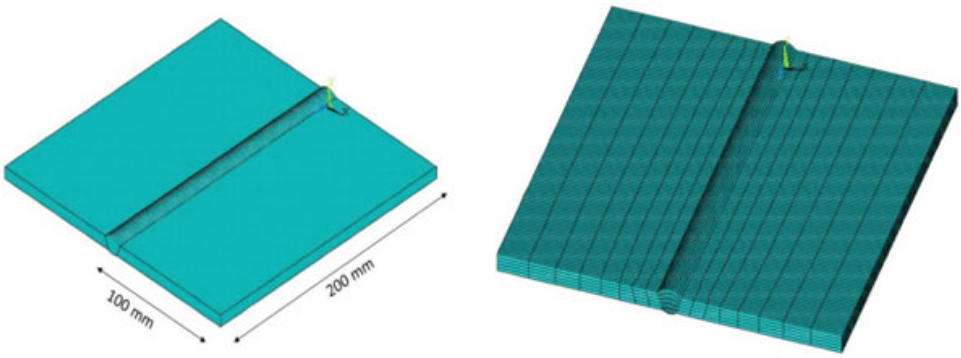


Fig. 87.2 3D solid model and FE Meshed model

the eight- node SOLID70 element was later replaced to eight-node SOLID185 element to perform structural analysis. The weld beads obtained from experiments were analyzed using Nikon SMZ25 microscope with magnification of 0.5× and 30 W scanning power. The bead geometry thus obtained was used for developing FE model. The modelling details single sided SAW is shown in Fig. 87.2.

The solid model was built based on actual experimental bead cross section, further the model was meshed using ANSYS 18.0. Fine mesh was carried out along the weld region to obtain better temperature gradient along weld zone and to reduce computation time coarse mesh was used away from the bead zone. The 3D solid model and FE mesh model was shown in Fig. 87.2. The temperature dependent material properties [11] of austenitic steel is mentioned in Figs. 87.3 and 87.4.

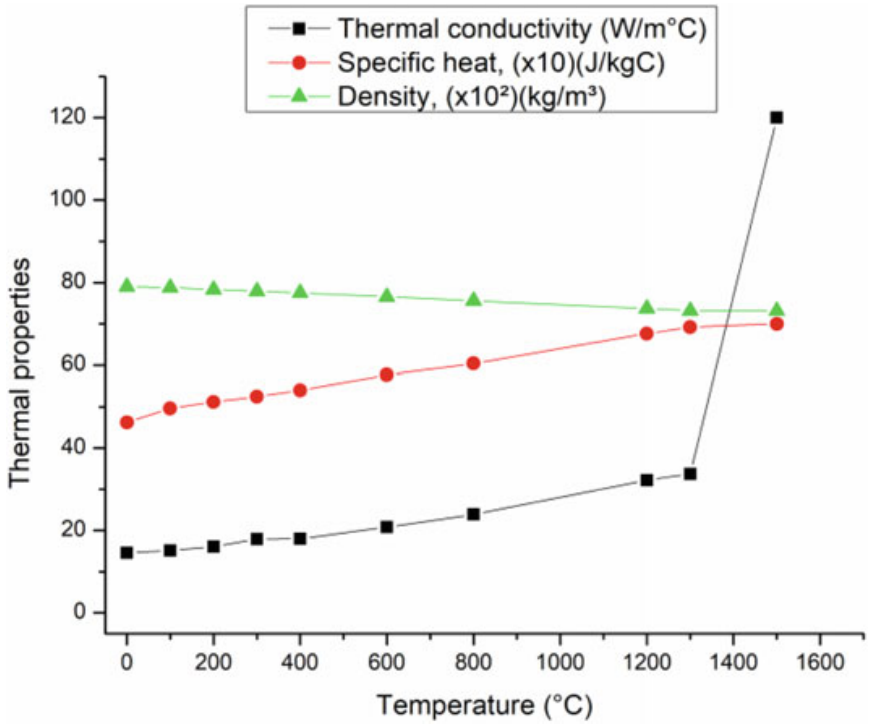


Fig. 87.3 Transient thermal properties of 304 stainless steel [11]

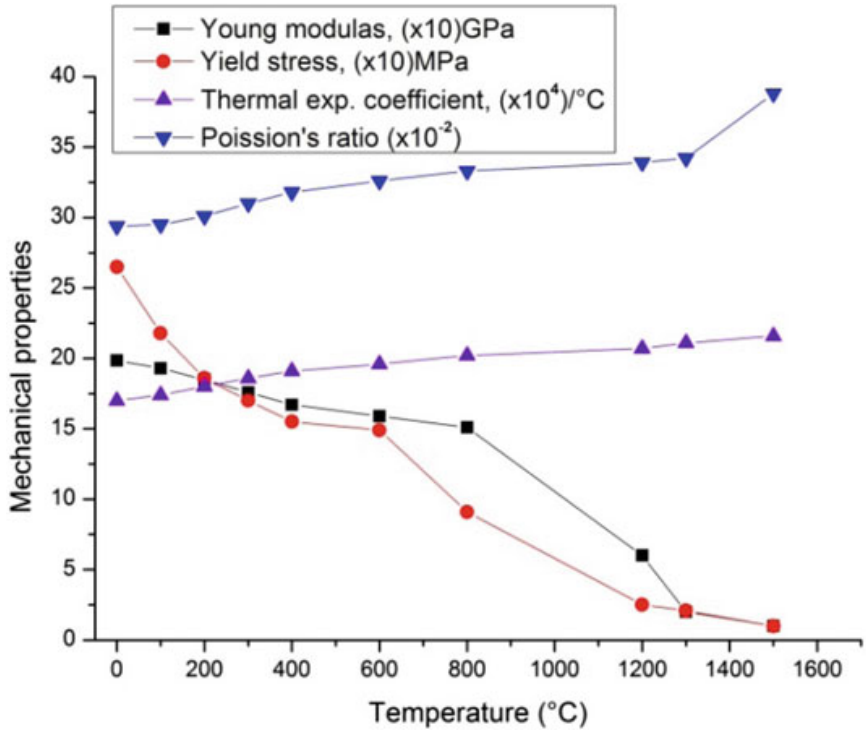


Fig. 87.4 Transient mechanical properties of 304 stainless steel [11]

87.4.1 Modelling the Moving Heat Source

The modelling of a transient thermal analysis with arc moving from one location to another with respect to time in order to generate uniform weld bead, the element birth and death technique was used. The entire weld pool was divided into multiple elemental volumes and subsequently with respect to time, heat was added and removed consecutively from one bead to another, thus attaining a moving heat source model. The quantity of heat added and removed was determined by following equation:

$$Q = \frac{\eta W}{v} \tag{87.7}$$

where Q is the quantity of heat added and removed, η is the efficiency of welding process, v is the volume of the individual weld bead considered, W is the welding power.

87.5 Results and Discussions

The heat applied to create a weld joint is dissipated by conduction to the base metal and by convection to the environment. Thus incorporating all possible conditions, and by varying the process parameter the numerical analysis was carried out in

order to study the variation in thermal and mechanical behavior of the weld. The same process parameters used to develop weld during experiments were considered for numerical analysis. The model is cooled till 500 secs after application of thermal load.

87.5.1 Moving Heat Source and Thermo-Mechanical Analysis

The simulated moving heat source model obtained numerically is shown in Fig. 87.5 for the welding variable voltage of 24 V, current of 430 A and speed of 20 m/h.

The tack weld boundary condition was applied and the results from thermal analysis were imposed to analyze the mechanical behavior of the model. Figures 87.6 and 87.7 shows the contour of transverse residual stress and residual stress along thickness; longitudinal residual stress and von Mises stress respectively.

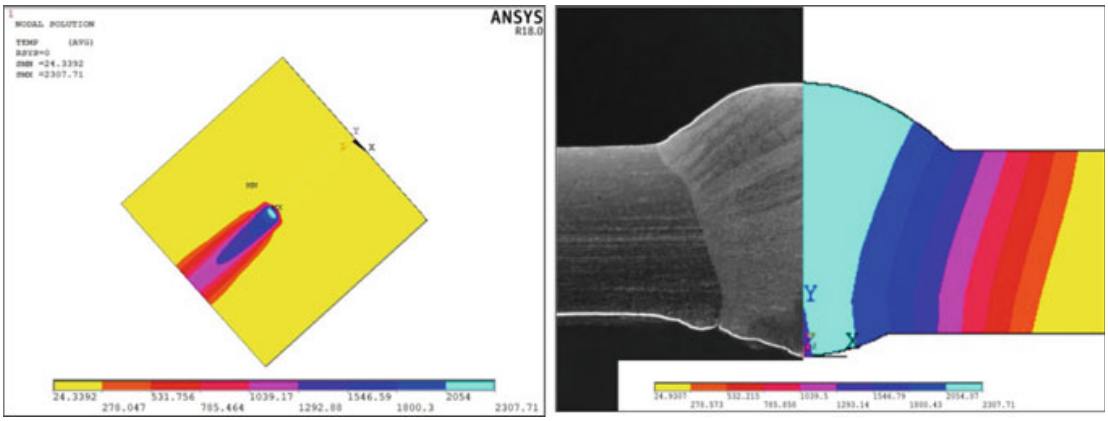


Fig. 87.5 FE model moving heat source and Experimental and numerical bead cross section

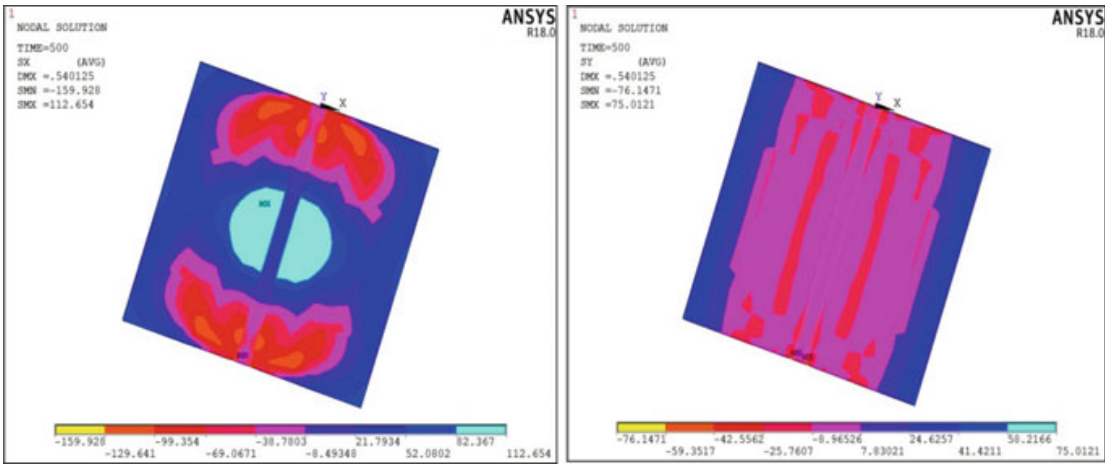


Fig. 87.6 Contour plots of transverse residual stress and residual stress across thickness

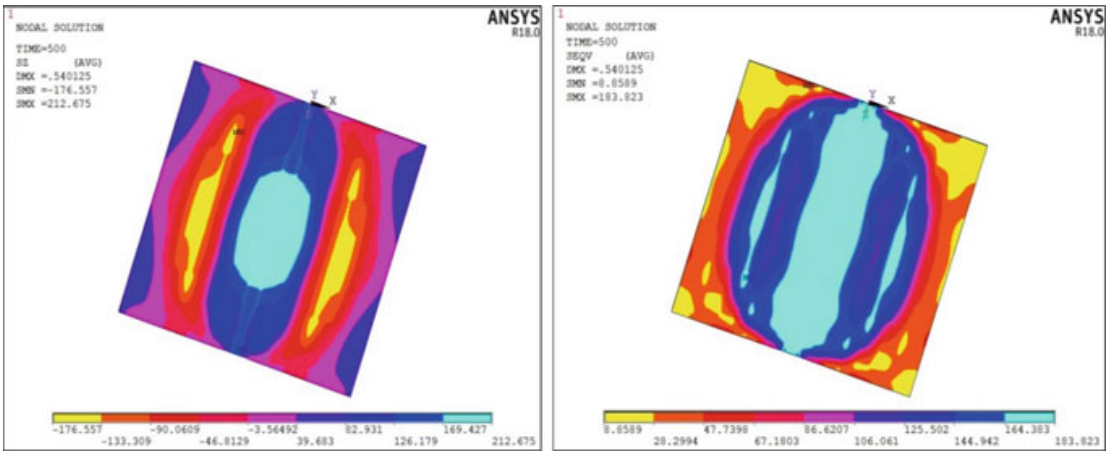


Fig. 87.7 Contour plots of longitudinal residual stress and von Mises stress

Due to variation in temperature gradient along the width direction (away from the weld), there are variation in shrinkage pattern which lead to development of tensile stress contour in the weld zone and its vicinity, and to balance the same compressive stresses are generated in rest of the portion of plate [12]. The stress reversal takes place between molten zone and adjoining zone due to drastic change in physical and mechanical properties of molten material and solid material [13]. Also at the start and end of a weld, large shrinkage pattern causes the variation in stress contour, as compressive stresses are generated in the ends and tensile stress are generated along the weld. This variation is evident in transversal and longitudinal stress contour [13]. As, the model is cooled after thermal loading cycle, phase transformation occurs in the welded plate establishing plastic deformation which is measured in terms of plastic strain in this study [16]. The material closer to weld path faces tensile strain, as a result of contraction due thermal loading, which is later balanced by compressive strain away from the weld zone [14]. As material is deformed plastically, significant intergranular residual strains develop in the loading direction owing to strain redistribution due to slip along the preferred systems. The repeated expansion and contraction during a thermal loading cycle along with consequent cooling effect increases the dislocation density along the weld [15]. Since heat transfer is more turbulent along the thickness direction than along the width, this affects the plastic strain contour along the thickness direction [16].

87.5.2 Effect of Welding Speed

The major process parameters that influenced the shape and quality of weld bead are arc power and welding speed. In the present study effect of welding speed on residual stress is considered for analysis. Three samples were considered to analyse effect of residual stresses on varying of welding speed. The results were analysed based on observations from peak temperature distribution, residual stress on the

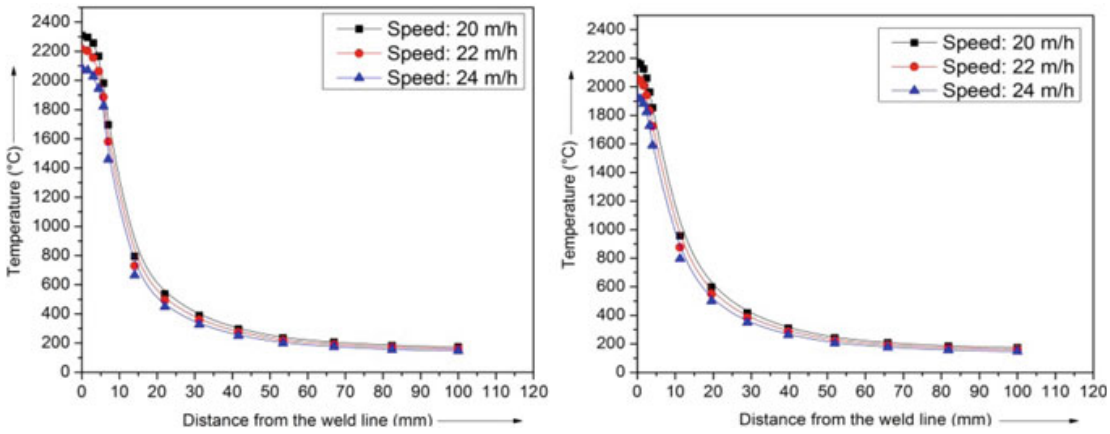


Fig. 87.8 Peak temperature perpendicular to weld line top and bottom surface

plate. The residual stress and plastic strain were observed along the weld centre line and perpendicular to the weld centre line. Residual stresses were considered for all three directions: transversal, longitudinal and along thickness direction of the plate. Also, von Mises stresses were calculated.

The various parameters used to study the effect of variation of welding speed are tabulated in Table 87.1.

The variation of peak temperature perpendicular to the weld line for three welding speeds for top surface and bottom surface of the plate is plotted in Fig. 87.8.

It was observed from the above temperature pattern that peak temperature reduces with increase in welding speed on both top and bottom side of the plate. Further analysis is carried out to determine the pattern of the residual stresses for top and bottom portion of the welded plates along and perpendicular the weld line.

Figure 87.9 shows the variation in residual stress for both top and bottom side of the plate along the weld line. Transversal residual stress (S_x), residual stress along

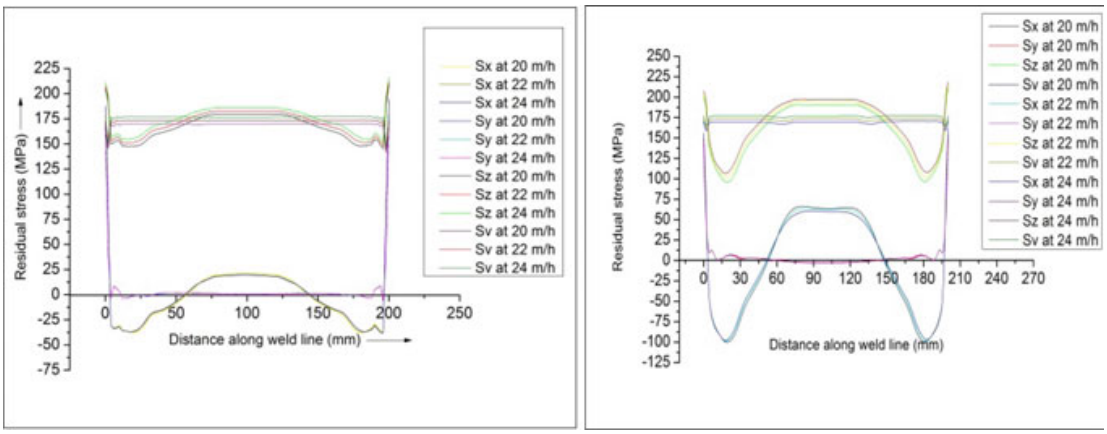


Fig. 87.9 Residual stress in *X*, *Y*, *Z* direction and von Mises stress along weld line for top side of plate and for bottom side of plate respectively

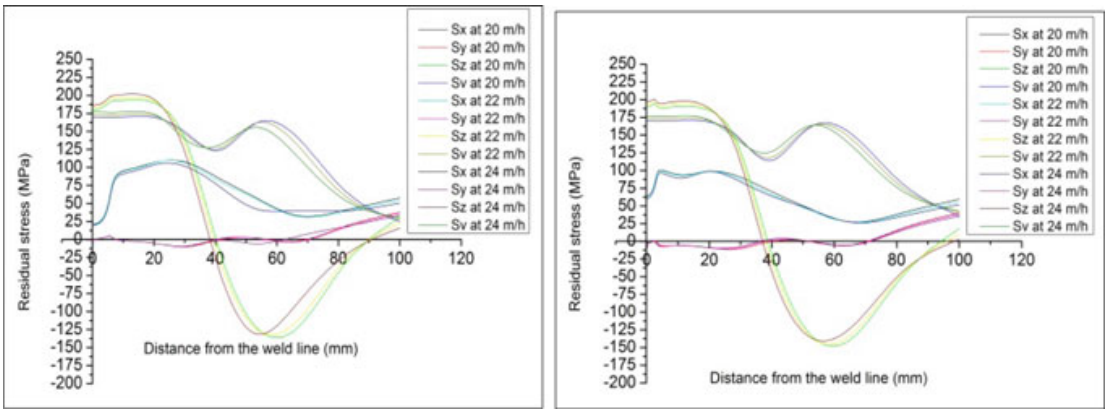


Fig. 87.10 Residual stress in *X*, *Y*, *Z* direction and von Mises stress perpendicular to weld line on top surface on bottom surface respectively

thickness direction (S_y), longitudinal residual stress (S_z) and von Mises stress (S_v) are considered for study. Figure 87.10 shows the residual stress in *X*, *Y*, *Z* direction and von Mises stress perpendicular to weld line on top surface on bottom surface respectively.

87.5.3 Analysis of Residual Stress on Top Surface of Plate

With increase in speed, longitudinal tensile stress and von Mises tensile stress increases, whereas, the variation in normal stress and transverse stress were negligible along the weld line. Both longitudinal tensile stress and von Mises tensile stress increases with increase in welding speed in the weld zone and in its near vicinity but reduces with increase in welding speed in the portion away from weld indicating increase in longitudinal compressive stress, but maximum value of longitudinal compressive stress was observed to decrease with increase in welding speed. Whereas transversal tensile stress reduces with increase in speed away from the weld center line, and negligible variation was noticed in normal stress away from the weld center line on the top surface. It is also noticed that the magnitude of normal stress both along the weld and away from the weld is not as evident as that of other components of stresses.

87.5.4 Analysis of Residual Stress on Bottom Surface of Plate

With increase in speed, longitudinal tensile stress and von Mises tensile stress increases, whereas, the variation in normal stress was negligible along the weld line. And transversal stress (both tensile and compressive) decreases with increase in

welding speed along the weld center line. Both longitudinal tensile stress and von Mises tensile stress increases with increase in welding speed in the weld zone and in its near vicinity but reduces with increase in welding speed in the portion away from weld. The maximum longitudinal compressive stress decreases with increase in welding speed which is located away from the weld vicinity. The Transverse residual tensile stress diminishes away from the weld center line with increase in welding speed, and negligible variation in normal stress was noticed away from the weld center line. It is also noticed that the magnitude of normal stress both along the weld and away from the weld is not as evident as that of other components of stresses.

87.6 Validation of the Model

The temperature profile was recorded for all welding parameters. For purpose of validating the numerical model, Experiment No. 1 thermal couple profile was matched with numerically obtained result for distance of 17 mm from the weld centre line. Figure 87.11 depicts that close agreement of experimental observation with FE model.

The peak temperature obtained at 17 mm from weld line was co-related with numerical values and is in good agreement with a maximum deviation of 4.77%. Hence the developed heat source model is validated with the experimental result.

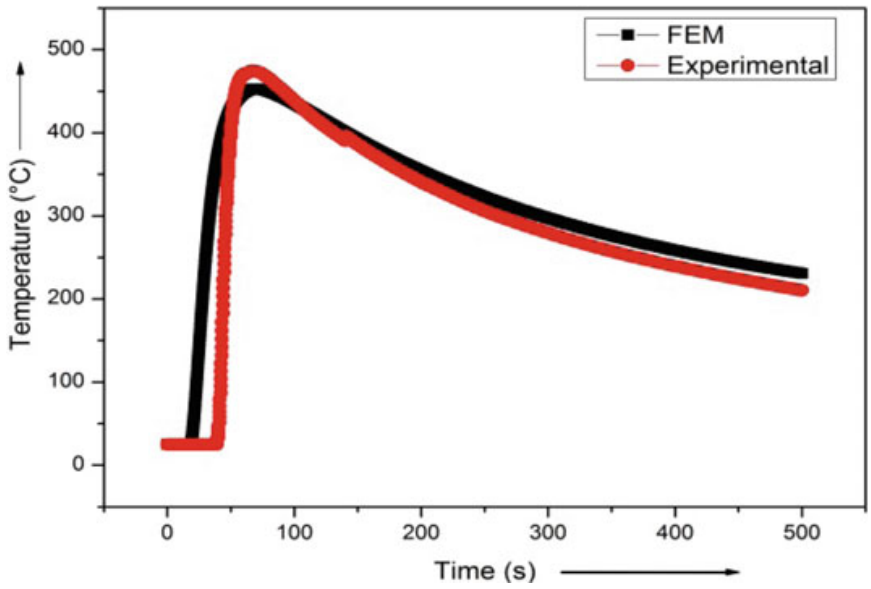


Fig. 87.11 Validation of experimental and numerical temperatures at 17 mm from weld line

87.7 Conclusions

1. The SA welding successfully conducted on AISI 304 stainless steel with a square butt configuration.
2. The temperature distribution obtained from the experiments is matched with that of developed heat source model and variation in peak temperatures is around 4.77%.
3. It was observed that with increase in welding speed the peak temperatures observed were low as higher speeds reduce the heat input into the material.
4. Thermomechanical analysis was done to determine the residual stress in *X*, *Y*, *Z* direction and von Mises stress for top and bottom of the plate.
5. With increase in speed, longitudinal tensile stress and von Mises tensile stress increases, whereas, the variation in normal stress and transversal stress were negligible along the weld line.

References

1. Houldcorft, P.T.: Submerged arc welding, 2nd edn. Abington Publishing, Cambridge UK (1989)
2. Rosenthal, D.: Mathematical theory of heat distribution during welding and cutting. Weld. J. **21**(5), 220-s–234-s (1941)
3. Rosenthal, D.: Theory of moving sources of heat and its application to metal treatments. Trans. AIME **43**(11), 849–866 (1946)
4. Ohring, S., Lugt, H J.: Numerical simulation of a time dependent 3-D GMA weld pool due to a moving arc. Weld. J. **79**(12), 416-s–424-s (1999)
5. Brown, S., Song, H.: Implication of three-dimensional numerical simulation of welding of large structures. Weld. J. 55-s–62-s (1992)
6. Goldak, J., Bibby, M., Moore, J., House, R., Patel, B.: Computer modeling of heat flow in welds. Metall. Trans. **17B**, 587–600 (1986)
7. Biswas, P., Mandal, N.R. Thermomechanical finite element analysis and experimental investigation of single pass single sided submerged arc welding of C-Mn steel plates. Proc. IMechE Part B J. Eng. Manuf. (224), 627–639 (2009)
8. Podder, D., Mandal, N R., Das, S.: Heat source modelling and analysis of submerged arc welding. Weld. J. (93), 183-s–192-s (2014)
9. Pathak, A.K., Datta, G.L.: Three dimensional finite element analysis to predict the different zones of microstructures in submerged arc welding. Proc. IMechE Part B **218**, 269–280 (2004)
10. Friedman, E.: Thermomechanical analysis of welding process using finite element method. Trans. ASME J. Press. Vessel Technol. **97**(3), 206–213 (1975)
11. Farid, V.H., Asl, A.Z.: Numerical and experimental investigation of T-shape fillet welding of AISI 304 stainless steel plates. Mater. Des. **47**, 615–623 (2013)
12. Velaga, S.K., Ravisankar, A.: Finite element based parametric study on the characterization of weld process moving heat source parameters in austenitic stainless steel. Int. J. Press. Vessels Pip. **157**, 63–73 (2017)

13. Ravisankar, A., Velaga, S.K., Rajput, G., Venugopal, S.: Influence of welding speed and power on residual stress during gas tungsten arc welding (GTAW) of thin sections with constant heat input: A study using numerical simulation and experimental validation. J. Manuf. Process. **16**, 200–211 (2014)
14. Dean, D., Wei, L., Hidekazu, M.: Determination of welding deformation in fillet-welded joint by means of numerical simulation and comparison with experimental measurements. J. Mater. Process. Technol. **183**, 219–225 (2007)
15. Murat, O.A., Fitzpatrick, M.E.: Determination of plasticity following deformation and welding of austenitic stainless steel. Mater. Sci. Eng., A **701**, 203–213 (2017)
16. Yadaiah, N., Bag, S.: Effect of heat source parameters in thermal and mechanical analysis of linear GTA welding process. ISIJ Int. **52**, 2069–2075 (2012)

Chapter 88
Stochastic Longitudinal Vibration of Single Walled Carbon Nanorods—A Non Local Elasticity Approach

Tanmoy Loha and Sudip Dey

Abstract The present study deals with stochastic nonlocal non-dimensional natural frequencies of single walled carbon nanorods and rotating cantilever beam. In this study, the chordwise bending–vibration behavior of nanocantilever is analyzed. The nonlocal parameter focused on the small-size effects when dealing with nanosize structures such as single-walled carbon nanotubes (SWCNTs). The Eringen's nonlocal elasticity theory in conjunction to the governing differential equations are utilized to solve the problem. The stochastic nonlocal natural frequencies for rotating nano cantilever are observed by employing the differential quadrature method (DQM). The effects of the scale is computationally analyzed based on Monte Carlo simulation (MCS). In another case free longitudinal frequency of a nanorods considering two types of boundary condition is observed, namely, clamped-clamped and clamped-free. The longitudinal vibration of the system are described by a set of partial differential equations, derived by using D'Alembert's principle and using by employing Fourier infinite series in conjunction to separation of variables. The nanoscale effects implement a important role on the frequency response of nanorods subjected to rotation. The statistical analysis are carried out based on stochastic input parameters such as material properties (elastic modulus, density) and geometric properties (length).

Keywords Nanorods · Stochasticity · Non local effect · Fundamental frequency · Scale effect

88.1 Introduction

Due to introduction of graphene and carbon nanotubes [1], nanostructure have gained immerse popularity in the field of the scientific community. Such materials can be employed to obtain the nanodevices made of complex structures. Due to

T. Loha (✉) · S. Dey
Mechanical Engineering Department, National Institute of Technology Silchar, Silchar, India
e-mail: tanmoyloha5@gmail.com

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_88

eminent effectiveness of usability in nanodevices also its eminent characteristics draws attention on the improvement of nanodevices and nanomachines. These nanomaterials contains superior mechanical, thermal, electrical, chemical and electronic properties. In classical continuum mechanics strain field of distant point to reference point is eliminated but whereas in non local elasticity approach, nonlocal elasticity and the differential quadrature method (DQM) application in order to determine flapwise bending vibration of a rotating nanocantilever consider these distance effect between surroundings atoms. Therefore stress field at a reference point depends not only that particular point but also on the strain at every other points at that corresponding domain. An experimental approach at nano scale is quite difficult and time consuming because handling all influencing parameters of a nanostructure simultaneously and obtained result by considering all parameters together is critical task and trouble worthy task. Furthermore employing molecular dynamics (MD) simulation needs large time and high computational efficacy.

In the present work the size effect and surroundings effect of molecule, strain field effect of nearby atoms on a structural element familiar as nonlocal theory is taken into consideration. The nonlocal elasticity theory was first introduced by Eringen's [2] to study nanomaterials behaviour and utilization of nanomaterials. Vibrational response of various nano or micro level structures like as single and multiple beam, rod, nanoplate and nanodisc in conjunction with elastic, visco elastic or other varieties of sheet thickness has been explored by various authors [3–5]. The actual significance of nonlocal elasticity theory stimulated us to study various features of nano/micro structure and implements that specific properties into nanodevices. The viability of nonlocal continuum theory in the field of nanotechnology was first reported by Peddieson et al. [6]. Various efforts related to nonlocal elasticity theory have been reported on several well established literature e.g. [7–9]. Recently Lee and Chang [10] have analysed vibrational response of single-walled carbon nanotube (SWCNTs) filled with fluid embedded in an elastic medium using winkler-type model. Further Murmu and Pradhan [11] established a single non local beam model to evaluate flapwise bending vibration behaviours of a rotating cantilever. They observed that vibration responses of a rotating nano cantilever is strongly effected in small scale. Recently Pradhan [12] has carried out finite element analysis for getting various mechanical response buckling, bending and vibration for non local beam under various boundary conditions like hinged-hinged, clamped-hinged, and clamped-free are extended their results to taper beams. According to Wang [13] a modified non local beam model was developed and they have found that nano scale non local parameter and critical flow velocity has a great impact on vibration and stability response of a nanotube carrying fluid. Danesh et al. illustrated the scale effect on tapered nanorod subjected to longitudinal vibration using the nonlocal theory. The authors have used the differential quadrature method to solve motion equations for three different boundary conditions [14]. Aydogdu, was modelled Carbon nanotube immersed into elastic medium i.e. a nanorod encompassed by elastic layers using the Eringen's nonlocal theory. The author distinguished the longitudinal frequencies between the nonlocal and classical continuum models [15]. Abdollahi [16] implement the EBFs on bounded domains, the effects of the

boundary conditions are handled by truncating the kernel/attenuation function in the constitutive relations. This helps to found residuals in the governing equations, also to minimize the residuals near the boundary region by using weighted residual approach. Danilo Karličić et al. applied nonlocal Kirchhoff-Love plate theory to illustrate mechanical behaviours of single-layered and multi layered graphene sheets used as an orthotropic material. Here they have derived the governing equation of motion of a simply supported orthotropic nanoplate. They also had showed that magnetic field could be successfully implemented to enhance sensitivity of nanomechanical sensors [17]. Bahaadini [18] defined the instability behaviour of cantilever nanotubes on dynamic structural fluid and they have incorporated end concentrated follower force and distributed tangential load to improve accuracy of the results like surface elasticity and residual stresses. The frequency response function is an important structural parameter which is defined as the ratio of output response of a structure due to an applied force. Such modal response can be measured in the form of displacement, velocity, and acceleration. In spite of recent development the model which considered in published work [11], stochasticity effect based on parametric study has not yet been considered. Uncertainty quantification on various properties of the above said model has not yet been considered because geometric and material uncertainties leading to unpredictable system response on natural frequencies. Reliability based failure analysis of single walled carbon nanorods structure still not observed. Based on carbon nanostructure various design configuration of nanomachines are found but for that proper understanding of mechanical behaviour is needed. One such nano machines is nano turbine whose performance depends upon dynamics properties of beam like nanocantilever elements. From design point of view study of bending vibration on nanocantilever, nanorod are essential for complexity in its lattice structure. Due to interdependency of large number of parameters in complicated fabrication processes, the system properties are inevitably random in nature resulting in uncertain response of the nano structures. Many researchers worked on nano structures in past but mainly focused on deterministic analysis [1–4]. The stochastic natural frequency responses are solved by random eigenvalue problem with an improved nonlocal elasticity based approach. Eringen's nonlocal elasticity theory is idealized as nanocantilever beam to study the flapwise bending vibrational frequencies of rotating nanocantilever. The longitudinal frequency responses of nanorod is determined by using methods of separation of variables on ordinary differential governing equation. The various effects of nonlocal parameters, small scale effects and nondimensional angular velocities on nanostructure are investigated in the present study.

88.2 Nonlocal Parameter and Nonlocal Elastic Model

The basic parameter in non local model e_o remains constant for each nanostructure. The parameter e_o should be taken in such way so such non local model yields certain approximation when dispersion curve due to lattice effect on plane waves

are considered. The term $e_o a$, $e_o l_c / l_a$ are the nonlocal parameters or the scale coefficient. Non-local elasticity theory has the capability to envisage the statistical behaviours on longitudinal vibration response and to predict buckling characteristics of CNTs. However, main distinguishable features of nonlocal theory is to optimize nonlocal parameter's value to get better effective outcomes from any systems. In this theory nanoscale effects are taken into consideration when lattice dimensions of atoms considered as material parameter. The most generalized form of the constitutive relations for nonlocal elasticity involves an integral over the whole body. The nonlocal linear homogeneous equations are explained by Pradhan and Murmu [19]. Based on Euler–Bernoulli hypothesis nonlocal constitutive relations for general one dimensional beam can be expressed as

$$\sigma(x) - (e_o a)^2 \frac{\delta^2 \sigma}{\delta^2 x} = E\varepsilon(x) \tag{88.1}$$

where σ and ε denote the normal stress and strain due to nonlocality respectively. E denotes Young's modulus of the element. The term $(e_o a)$ is nothing but nonlocal parameter or the scale-coefficient as depicted in a well-established literature [19].

88.3 Materials

In general, during synthesis of nanostructures, high stiff and light-weight materials are employed so that the structure can sustain large external load without failure. In the present study, single walled carbon nanotube is used because of its extraordinary material properties. It is made of graphene sheet which is idealized to construct the present model. Carbon nanotubes (CNTs) being the allotrope of carbon have a cylindrical closed packed structures. These CNTs have unusual properties, which are valuable for nanotechnology, optics, electronics industry and other fields of materials science and technology. Owing to the material's exceptional strength and stiffness, nanotubes are constructed with length-to-diameter ratio of up to 132,000,000:1 significantly larger than for any other material.

88.4 Methodology

88.4.1 Governing Equation of Motion of Rotating Nanocantilever

Let us assume a nano cantilever whose length is L and one end fixed at point O to a rigid hub of radius r rotates at a constant angular speed Ω rad/s in counter clockwise direction shown in Fig. 88.1. From different literature survey it is observed that the classical Euler-elastic beam theory offers a most reliable model for mechanical

deformation due to axial loading in one-dimensional nanostructures, such as carbon nanotubes, nanowire, nonobeam, nanoplate. Further it can also be assumed that plane of cross section normal to longitudinal axis remains plane before and after deformation and deformation due to bending is considered as very small. Thus rotating nanocantilever model is considered as an Euler–Bernoulli beam model. In the present theory the plane cross sections of the beam remain plane during flexure and that the radius of curvature of a bent beam is larger compared to the beam's depth. Here W denote deflection along flapwise deformation direction.

The moment stress resultant relation can be expressed as $M(x,t) = \int_A z\sigma_x \mathrm{d}A$

Using Eq. (88.1) and from moment relation, nonlocal moment relation can be expressed as

$$M - (e_o a)^2 \frac{\delta^2 M}{\delta x^2} = -EI \frac{\delta^2 W}{\delta x^2} \tag{88.2}$$

Now based on Euler-Bernoulli's hypothesis is applied in basic differential form of Euler Lagrange's equation as follows

$$\frac{\delta^2 M}{\partial x^2} + \frac{\delta}{\delta x}\left(\Gamma \frac{\delta}{\delta x}\right) - \rho A \frac{\delta^2 W}{\delta t^2} = 0 \tag{88.3}$$

where Γ denotes the tensile force due to centrifugal action of rotation of cantilever beam. The force $\Gamma(x) = \int_x^L \rho A \Omega^2 (r + x)\mathrm{d}x$. Here r is taken as the hub radius which denote length of beam from centre point of rotation to edge of cantilever beam, Ω denotes angular velocity with which cantilever rotates. Now The second order derivative of M is determined from Eq. (88.2) by substituting into Eq. (88.1) it is found that

$$M = -EI \frac{\delta^2 W}{\delta x^2} + (e_o a)^2 \left[-\frac{\delta}{\delta x}\left(\Gamma \frac{\delta W}{\delta x}\right) + \rho A \frac{\delta^2 W}{\delta t^2} \right] \tag{88.4}$$

Now substituting the expression for M from Eq. (88.3) into Eq. (88.4), the governing partial differential expression for vibration of rotating cantilever are obtained as

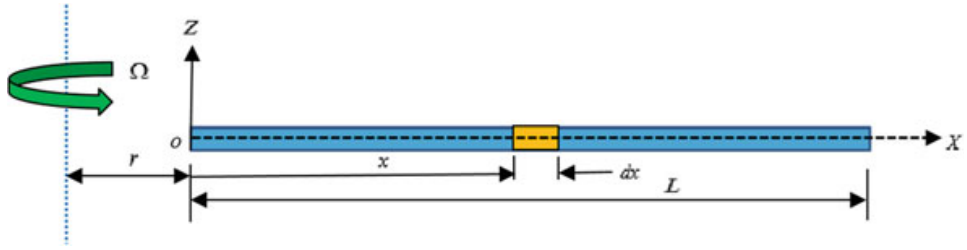


Fig. 88.1 Basic layout of a rotating nanocantilever based on Euler–Bernoulli approach

$$\frac{\delta^2}{\delta x^2}\left(EI\frac{\delta^2 W}{\delta x^2}\right)-\frac{\delta}{\delta x}\left(T\frac{\delta W}{\delta x}\right)+\rho A\frac{\delta^2 W}{\delta t^2}+(e_o a)^2\frac{\delta^2}{\delta x^2}\left[\frac{\delta}{\delta x}\left(T\frac{\delta W}{\delta x}\right)-\rho A\frac{\delta^4 W}{\delta t^4}\right]=0 \tag{88.5}$$

The four boundary conditions for a nanocantilever beam at two ends are expressed as

when $x = 0$, $W = \frac{\delta W}{\delta x} = 0$ and $\frac{\delta^2 W}{\delta x^2} = \frac{\delta^3 W}{\delta x^3} = 0$ at $x = L$,

After non-dimensionalization, the principal parameter employing differential quadrature (DQ) method [19], the Eigen value equation can be obtained as $K_{\text{total}}\{W\} = \lambda^2\{W\}$, where K_{total} denotes the cumulative DQ matrix. From the above Eigen matrix, the Eigenvalues are solved and the fundamental frequencies are calculated for longitudinal vibration of nanocantilever based on nonlocal elasticity approach [20].

88.4.2 Characteristics Equations of Nanorod

In present study, a nanorod having length L and arbitrarily selected area of cross section A is considered as illustrated in Fig. 88.2a. It is assumed that cross-sectional area along X-direction remains constant. The materials of nanorod are considered to be homogeneous and isotropic in nature. To find out fundamental frequencies of nanorods vibrating longitudinally, it is assumed to be vibrated along X-direction. For analysing of present model, it is considered as an infinitely small element having length dx at location x and the equilibrium condition is based on D'Alembert's principle as furnished in Fig. 88.2b. A normal resultant force N is applied due to normal stress σ_x in axial direction. In addition, the continuously distributed axial loads f (x, t) are neglected. On the right side of infinitesimal element elementary force $+\mathrm{d}N$ is considered.

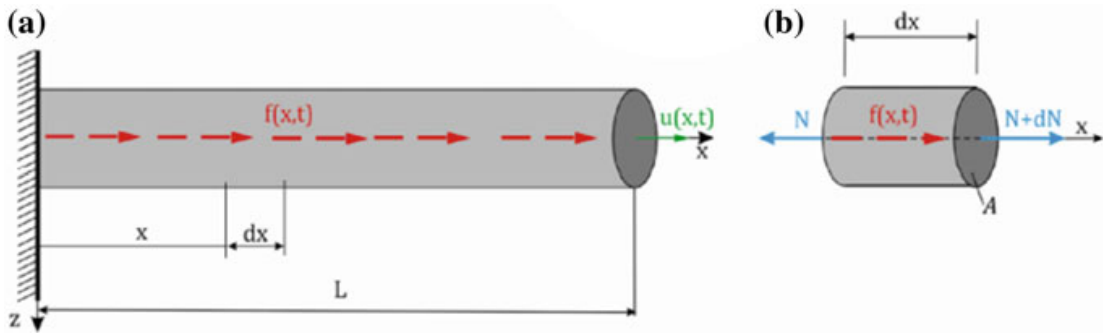


Fig. 88.2 **a** Model of a nanorod vibrating longitudinally; **b** corresponding infinitesimal element subjected to resultant stresses in axial direction [21]

By considering D'Alembert's equilibrium equation, equating sum of all forces acting along x axes is equal to zero which states

$$-N + (N + \mathrm{d}N) + f(x, t)\mathrm{d}x = \rho A \mathrm{d}x \frac{\partial^2 u}{\partial t^2} \tag{88.6}$$

where $dN = \frac{\partial y}{\partial x} dx$ is the resultant stress N of the small infinitesimal element, $u = u(x, t)$ is the axial deformation along x-direction. After some transformations the following equilibrium equation is obtained in the differential form,

$$\frac{\partial N}{\partial x} + f(x, t) = \rho A \frac{\partial^2 u}{\partial t^2} \tag{88.7}$$

where, N denotes the resultant axial stress over whole cross section area defined by $N(x, t) = \int_A \sigma_x \mathrm{d}A$, where σ_x represent normal stress in infinitesimal element.

According to Eringen's nonlocal elasticity model the constitutive equations in differential form are specified as follows

$$\sigma_x - (e_o a)^2 \frac{\partial^2 \sigma_x}{\partial x^2} = E\varepsilon_x \tag{88.8}$$

$$N - (e_o a)^2 \frac{\partial^2 \sigma_x}{\partial x^2} = E\varepsilon_x. \tag{88.9}$$

By substituting (88.7) into (88.9) the following governing differential equation of motion based on D'Alembert's principle and Eringen's model of nonlocal elasticity in terms of axial displacement is obtained

$$\rho A \frac{\partial^2 u}{\partial t^2} - f(x, t) - EA \frac{\partial^2 u}{\partial x^2} = (e_o a)^2 \frac{\partial^2}{\partial x^2} \left(\rho A \frac{\partial^2 u}{\partial t^2} - f(x, t) \right) \tag{88.10}$$

In the present study two types of boundary conditions are considered.

(a) -Clamped-Clamped- $U(0, t) = U(L, t) = 0$
(b) -Clamed-Free- $U(0, t) = N(L, t) = 0$

where $N(L, t)$ is nonlocal axial resultant force acting at the end of a nanorod. From (88.7) and (88.9) the following equation can be derived

$$N(x, t) = (e_o a)^2 \frac{\partial^3 u}{\partial x \partial t^2} + EA \frac{\partial u}{\partial x}. \tag{88.11}$$

Considering oscillatory motion of nanorod with time and using separation of variables methodology, the expression of governing equation is assumed as $\theta(x, t) = \sum_{n=1}^{\infty} X_n(x) T_n(t)$, where $T_n(t) = e^{i\omega_n t}$ is the time harmonic function

$X_n(x)$ is the mode shape function corresponding to longitudinal vibration which is governed by some end boundary condition of the system. Neglecting continuously distributed axial load $f(x,t)$ the following ordinary differential expression with corresponding mode shape function $X_n(x)$ can be obtained as

$$\frac{\mathrm{d}^2 X_n(x)}{\mathrm{d}x^2} + \alpha_n^2 X_n(x) = 0 \tag{88.12}$$

where α_n is the characteristics value of the mode shape function corresponding to longitudinal vibration, which depends on corresponding boundary condition

$$\alpha_n = \frac{\omega_n^2}{\frac{E}{\rho} - (e_o a)^2 \omega_n^2}. \tag{88.13}$$

The nonlocal frequencies of a nanorod subjected to the longitudinal vibration can be determined by using boundary condition given below (Figure 88.3).

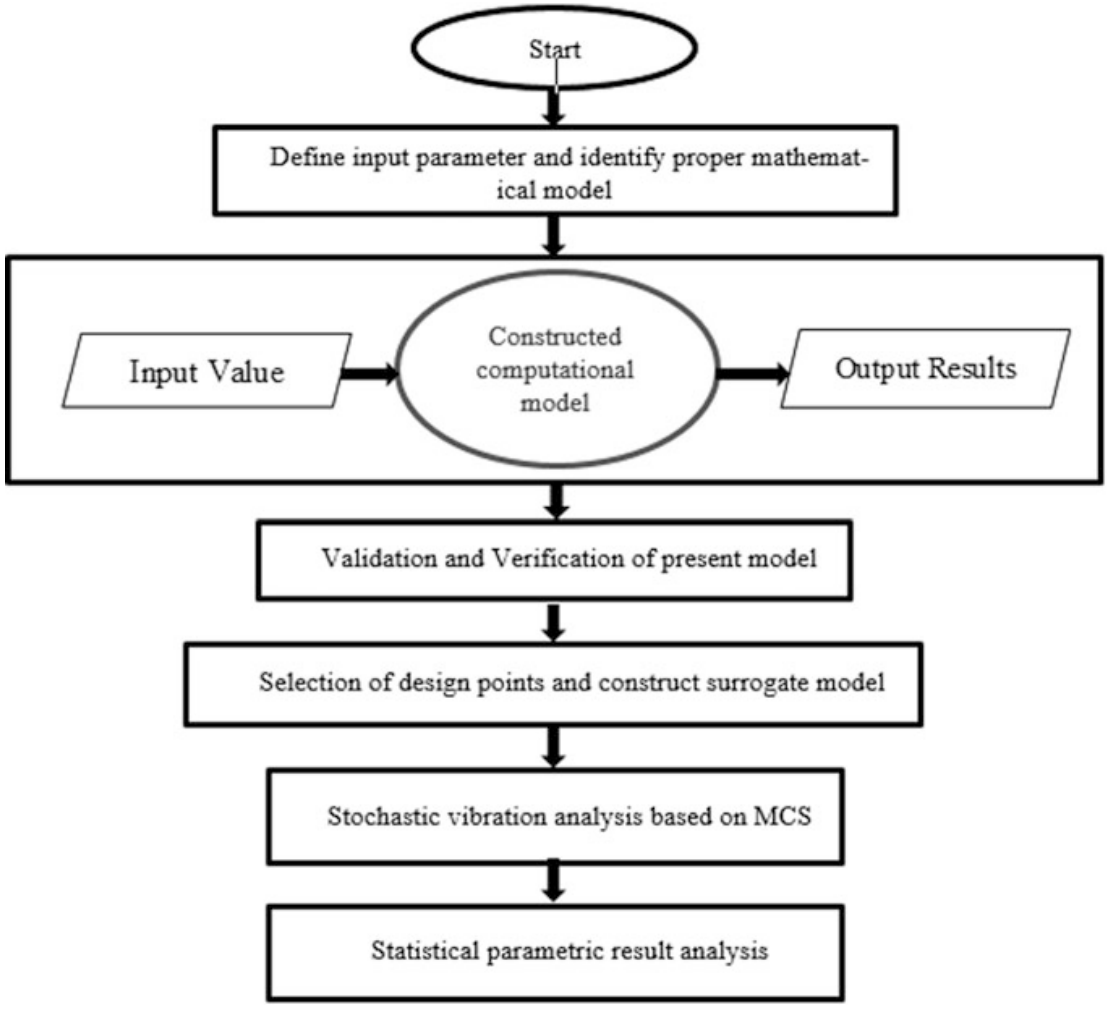


Fig. 88.3 Flow diagram of present methodology

$$\omega_n = \sqrt{\frac{E\alpha_n^2}{\rho\left(1+(e_o a)^2\alpha_n^2\right)}}. \tag{88.14}$$

88.5 Result and Discussion

In the present study, the following assumptions are considered such as, body force is neglected and employed the D'Alembert's Principle for equilibrium equation. Any contribution of shear force in cross section to transverse deformation is ignored i.e. transverse shear force is neglected, only deformation due to force applied in longitudinal is taken into consideration in this study. The method of separation of variable is considered using Fourier infinite series to solve ordinary differential equation (ODE). The fundamental mode is considered for validation and verification while the scale effect is considered to be caused by small internal molecular distance. Moreover, single walled carbon nanotubes formed by rolling graphene sheet is considered having material properties such as, Young's Modulus (E) 6.85 TPa and density (ρ) 9517 kg/m^3 [21].

Nonlocal elasticity theory is used to illustrate small scale effect on rotating nano cantilever beam. In present study, nonlocal elasticity model based on Eringen's theory is solved by differential quadrature method and the fundamental natural frequencies are obtained by considering the nonlocal parameter. After solving governing differential equation, the solution of governing differential equation is analysed by MATLAB code. The results are validated and verified based on Murmu and Pradhan [11] wherein the stochasticity analysis is carried out. The accuracy of fundamental natural frequencies and mode shapes depend on mode and material properties. The small scale effects on angular velocity and mode number is also investigated. For local model nonlocal parameter is assumed to be zero. But when non local parameter is taken into consideration then fundamental frequencies value corresponding to each curve becomes very close together. It is also observed that as the nondimensional angular velocity raised (see Fig. 88.4), non-dimensional fundamental frequencies are also found to increase due to effect of centrifugal force.

Moreover it can be observed that in nonlocal model of nano cantilever beam considered has been greatly influenced by angular velocity on the fundamental frequencies of nonobeam. Moreover local and nonlocal models is considered then more significantly small scale effect on mechanical vibration and buckling response are raised when cantilever rotates at high speed angular speed. It is also observed that higher frequency at high angular velocity zone is because of the simultaneous effect of both nonlocal parameter and angular velocity due to rotation. Fatigue effect on mechanical structure are also dependent on number of reversal of cyclic frequency applied on that particular structure. This investigation provides good information for the design of carbon nanotubes blade of nano turbine. It has been observed that when the fact of non local parameters are disregard longitudinal

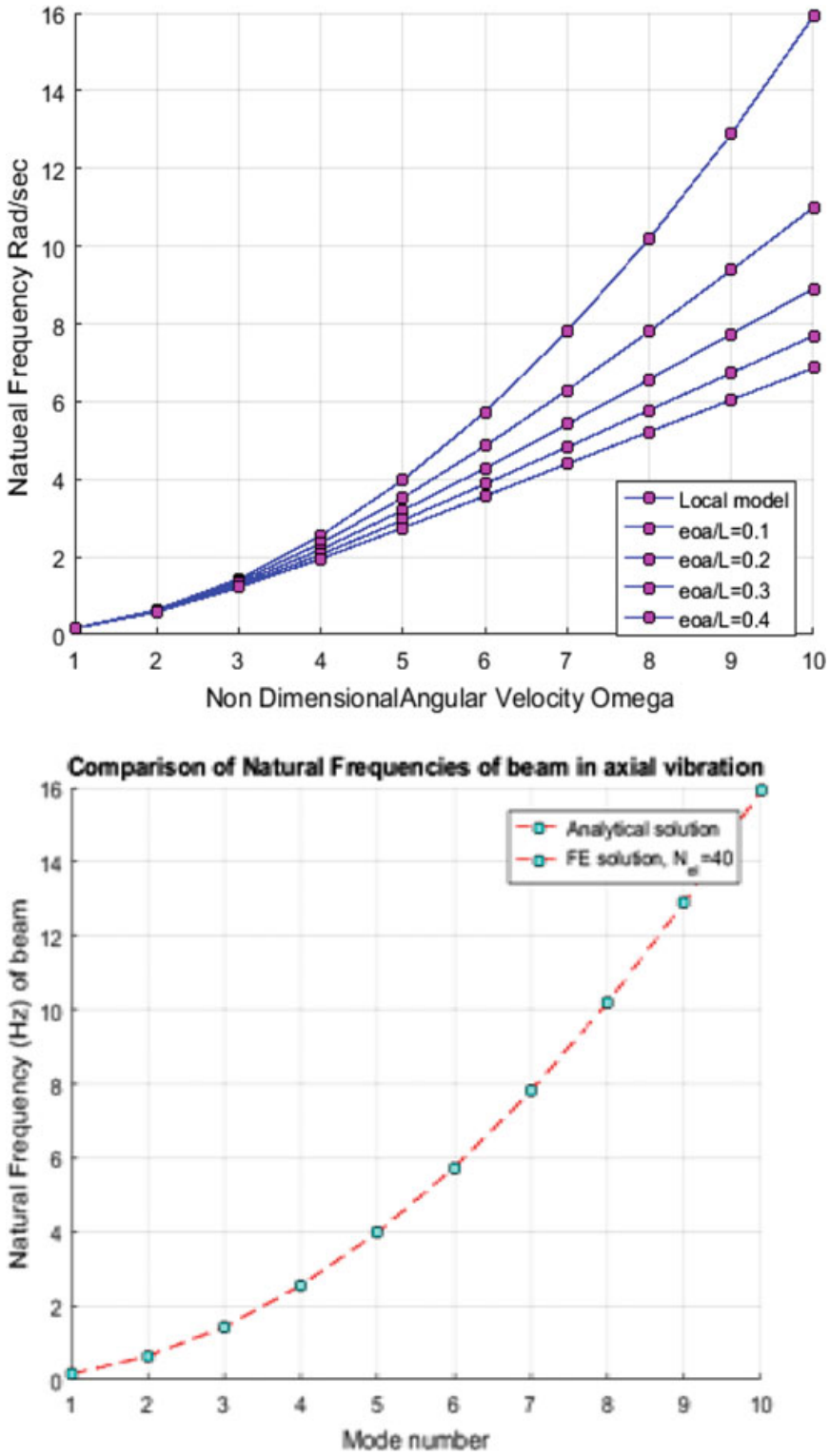


Fig. 88.4 Validation of non-dimensional natural frequency with corresponding angular velocity for different nonlocal parameters of a rotating nanocantilever beam and mode number with published literature [19]

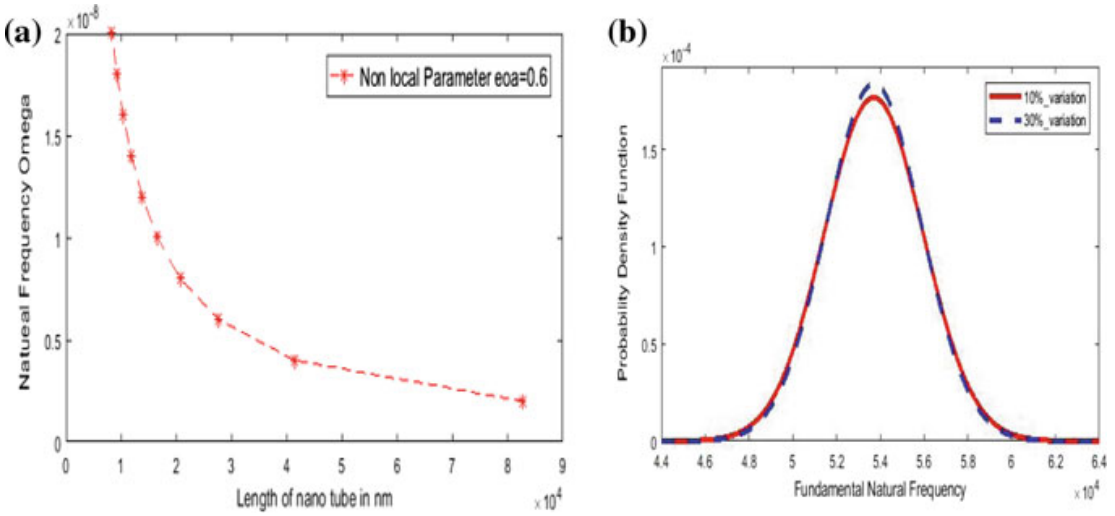


Fig. 88.5 **a** Validation of fundamental frequency changes with variation in length of nanorod with clamped free condition [21]. **b** Stochasticity analysis of fundamental natural frequencies of carbon nanorod based on randomness on input parameter (Non local parameter and Mode number)

vibrational frequencies are overestimated. Unlike transverse vibration of nanorod, natural frequency has got decreased with an increase in nonlocal parameter (see Fig. 88.5). Non local effect on longitudinal frequency of a nanorod has more prominent and distinguishable effect when any higher vibrational modes are considered. In the following table comparison and validation of the fundamental natural frequency of nanorod obtained from non local theory and those obtained from lattice dynamics based on MD simulation.

The following parameters and dimensions have been employed to validate the results as in given Table 88.1: ρ = 9517 kg/m^3, E = 6.85 TPa, L = 12.2 nm. Here fundamental frequencies are compared with the MD simulation for single walled carbon nanorod after obtaining results for clamped-free vibration using nonlocal elasticity theory. More consistent results are obtained when non local parameter goes to unity (see Table 88.1).

Since the value of internal length parameter a is fixed for SWCNT so in order to fit the result natural frequency optimization of parameter e_o is required for each mode of vibration corresponds to variation in geometrical properties of nanorod. However molecular dynamics (MD) simulation has been restricted to smaller scale because of higher time and cost consumption therefore nonlocal theory is used for more complex nanostructures analysis.

Above Fig. 88.5 indicates the effect of length non local parameter and of nanorod on fundamental frequency and it also shows characteristics of natural frequency of carbon nanorod based on random variation of material property. In present study each outcomes i.e. natural frequency on 10, 20, 30% are varied by using normal distribution and it can be observed that for small variation in material property within a certain limit natural frequency follows an normal distribution with a mean value constant and certain standard deviation. This parametric study helps us to observe uncertainty determination due to variation in material property due to

Table 88.1 Comparison and validation of nonlocal parameter influences on natural frequency of longitudinal vibration of nanorod

Non Local Parameter ($e_o a$)	Mode Number (n)	Present Analysis of Fundamental Frequency using MATLAB	Fundamental Frequency on Murmu and Adhikari published book [21]	Fundamental frequency using molecular dynamics on G. CAO et al. literature [22]	Percentage of error with Murmu and Adhikari book [21]	Percentage of error with molecular dynamics simulation on G. CAO et al. Literature [22]
0	1	0.488	0.549	0.544	11.11	10.29
0.5	1	0.536	0.548	0.544	2.19	1.47

structural inhomogeneity and anisotropic behaviour of structure in nanoscale. A notable decrease in fundamental frequencies with increase in length is observed (see Fig. 88.5). However, the fundamental nondimensional frequency has a decreasing in nature exponentially with a raises in the non-local parameter but this effect becomes insignificant with a greater length of the nanorod.

88.6 Conclusion

Present work shows the non local longitudinal frequency analysis of a single walled carbon nanorod and rotating nanocantilever, where first a non local mathematical model based on Eringen's non local stress theory has been used, then uncertainty effect based on parametric study of randomness on material property has been observed. It can be observed that nonlocal parameter has an important role when scale effect is considered on nanostructure. It has also been well established that when first vibrational mode consider, the fundamental non dimensional frequencies evaluated with nonlocal based models have higher magnitude than those obtained by local model. With increase in magnitude of angular velocity natural frequency on cantilever beam increased and also when length increases natural frequency of nanorod decreases exponentially.

The novelty of present study includes the computational investigation on stochastic frequency response analysis using computational code of a nanorod with uncertainty in the material properties i.e. young's modulus, density and non local parameter, mode number. Computational code is very accurate as compared with published work. Also when randomness is incorporated in material property natural frequency responses like a normal distribution curve then central tendency of frequency for continues variation of various material property can easily be evaluated. Stochasticity has also been investigated based on randomness in input parameter which shows uniform variation in fundamental frequency curves with very effective manner.

Present work is helpful in manufacturing point of view because Present study deals a significant contribution on stochastic behaviour of fundamental frequencies for a nanoscale structure so that structure become safe when that particular structure used in some specific domain of a machine element basically Nano Electro Mechanical System (NEMS) systems manufacturing like turbine blade, centrifugal pump blade etc. As manufacturing of any systems requires first fully structural analysis like bending, buckling, vibrational behaviours of any static or dynamic structures then according to structural analysis report manufacturings are being carry forwarded. From the physical point of view, it can be summarized that the non-local parameter dampened away fundamental natural frequencies. Fatigue response of a nanorod as well overall stiffness also depends on frequency response of the nanorod. To the best of authors' knowledge, this is the first attempt for applying Nonlocal elasticity in uncertainty quantification of nanostructure. Finally it is concluded that the present methodology can also be extended to analyse statistical behaviour of nanoplates as well as buckling behaviour analysis of nanoplate.

References

1. Lijimia, S.: Helical microtubules of graphitic carbon. Nature **354** (1991)
2. Eringen, A.C.: Differential equation of nonlocal elasticity and solution of screw dislocation and surface waves. J. Appl. Phys. **54** (1983)
3. Kozi, P., Pavlovic, R., Karlic, D.: The flexural vibration and buckling of the elastically connected parallel beams with a Kerr-type layer. Mech. Res. Commun. **56** (2014)
4. Ansari, R., Rajabiehfard, R., Arash, B.: Nonlocal finite element model for vibrations of embedded multi-layered graphene sheets. Comput. Mater. Sci. **494**, 831–838 (2010)
5. Oniszczuk, Z.: Free transverse vibrations of an elastically connected complex beam-string system. J. Sound Vib. **24**(4), 703–715 (2002)
6. Peddieson, J., Buchanan, G.G., McNitt, R.P.: Application of nonlocal continuum models to nanotechnology. Int. J. Eng. Sci. **41**, 305–312 (2003)
7. Zhang, Y.Q., Liu, G.R., Xie, X.Y.: Free transverse vibration of doubled walled carbon nanotubes using a theory of nonlocal elasticity. Phys. Rev. B **71** (2005)
8. Lu, P., Lee, H.P., Lu, C., Zhang, P.Q.: Dynamic properties of flexural beams using a non local elasticity model. J. Appl. Phys. **99** (2006)
9. Reddy, J.N., Roque, C.M.C.: Analysis of Timoshenko nanobeams with a non local formulation and meshless method. Int. J. Eng. Sci. **49**, 976–984 (2011)
10. Lee, H.L., Chang, W.J.: Surface effect on frequency analysis of nanotubes using nonlocal Timoshenko beam theory. J. Appl. Phys. **108** (2010)
11. Pradhan, S.C., Murmu, T.: Application of nonlocal elasticity and DQM in the flapwise bending vibration of a rotating nanocantilever. Physica E **42** (2010)
12. Pradhan, S.C.: Non local finite element analysis and small scale effects of CNTs with Timoshenko beam theory. Finite Elem. Anal. Des. **50**, 8–20 (2012)
13. Wang, C.M., Tan, V.B.C., Zhang, Y.Y.: Timoshenko beam model for vibration analysis of multi-walled carbon nanotubes. J. Sound Vib. **294**(4-5), 1060–1072 (2006)
14. Danesh, M., Farajpour, A., Mohammadi, M.: Axial vibration analysis of a tapered nanorod based on nonlocal elasticity theory and differential quadrature method. Mech. Res. Commun. **39**, 23–27 (2012)

15. Aydogdu, M.: Axial vibration analysis of nanorods (carbon nanotubes) embedded in an elastic medium using nonlocal elasticity. Mech. Res. Commun. **43**, 34–40 (2012)
16. Abdollahi, R., Boroomand, B.: Nonlocal elasticity defined by Eringen's integral model - Introduction of a boundary layer method. Int. J. Solids Struct. **51** (2014)
17. Karličić, D., Kozić, P., Adhikari, S., Cajić, M., Murmu, T.: Nonlocal mass-nanosensor model based on the damped vibration of single-layer graphene sheet influenced by in-plane magnetic field. Int. J. Mech. Sci. **96–97**, 132–142 (2015)
18. Bahaadini, R., Hosseini, M., Jamalpoo, A.: Nonlocal and surface effects on the flutter instability of cantilevered nanotubes conveying fluid subjected to follower forces. Physica B **509**, 55–61 (2017)
19. Murmu, T., Pradhan, S.C.: Small scale effect on buckling of single layered graphene sheet under biaxial compression via nonlocal continuum mechanics. Comput. Mater. Sci. **47**(1), 268–274 (2009)
20. Pradhan, S.C., Murmu, T.: Vibration analysis of nano single layered graphene sheets embedded in a elastic medium based on non local elasticity theory. J. Appl. Phys. **105** (2009)
21. Murmu, T., Adhikari, S., McCarthy, M.: Non-Local Structural Mechanics, 1st edn. Danilo Karličić, © ISTE Ltd. (2016)
22. Cao, G., Chen, X., Kysar, J.W.: Thermal vibration and apparent thermal contraction of single-walled carbon nanotubes. J. Mech. Phys. Solids **54** (2000)

Chapter 89
Finite Element Analysis and Three-Point Bend Testing of Indian Rice Grains with Fissure Effects

P. V. Gharat, D. K. Shinde and D. N. Raut

Abstract In today's world, one of the major problems faced by the rice-milling industries is breakage of rice during processes like de-husking/de-hulling, whitening, and polishing. For designing of milling machinery and for reducing the breakage of different rice varieties while processing, it is important to understand the rupture force of rice kernels of different varieties under the three-point breaking test. In this study, the different varieties of rice grains with different fissure conditions were investigated. The rupture force investigated by using three-point breaking test which is conducted on Instron UTM with specially designed adjustable support fixture, in which span can be adjusted for different varieties of rice grains. Five brown rice varieties—Rashi Poonam, Kolam, Komal, Basmati, and Daptari tested at a rate of 1 mm/min. The finite element analysis has helped to validate the result of the flexural stresses in the rice grain induced during test and then the crack propagation is simulated in the analysis. It was seen that out of the five rice varieties under tests, Rashi Poonam has the highest average rupture force of value 32.29 N. The results obtained from this test and computational analysis can help for developing methods and types of equipment used for whitening and polishing processing of rice with different varieties.

Keywords Brown rice · Three-point breaking test · Fissure detection

P. V. Gharat
Department of Production Engineering, Veermata Jijabai Technological Institute, Mumbai 400031, India

D. K. Shinde (✉) · D. N. Raut
Faculty Department of Production Engineering, Veermata Jijabai Technological Institute, Mumbai 400031, India
e-mail: dkshinde@vjti.org.in

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_89

89.1 Introduction

Rice is the seed of the grass species Oryza sativa, and it is one of the most important cereals in the world. The worldwide production of rice in 2014 was 741 million tonnes. India is one of the world's largest producers of rice. India's rice production is 20% of all worldwide rice production. In India, rice is pre-eminent crop and the staple food of the people of the various parts of the country. For the year 2013–2014, India's rice production was 106.54 million tonnes. In India, rice is a major crop and staple food for more than 70% of the population, thus plays a vital role in the national economy [1]. Rice passes through various processing like harvesting, de-husking, whitening, and polishing which may lead to breakage of rice grains. Rice is preferably consumed as whole grains. The percentage of unbroken rice grains gives good economical value to farmers and millers. Therefore for accurate design of the de-husking, whitening and polishing machines to minimize the grains breakage, information on physical properties such as length, width, and thickness and mechanical properties of rice grains such as breaking force and graph of force versus deformation must be provided [2].

The various physical properties like length, thickness, and mass of rice variety PR-106 were measured by Gadge. The moisture content of rice was 13.34% (wet basis). The average length, width, and thickness were 6.61, 1.75, and 1.40 mm, and unit mass volume of rice kernels were 0.017 g, and 0.051 cm^3, respectively. The mean diameter and surface area and aspect ratio were 2.52 mm, 20.1 mm^2, and 26.58%, respectively [3]. The study by Zareiforoush et al., the rupture force of paddy grains was measured under quasi-static compressive loading. The Alikazemi and Hashemi varieties of paddy rice were tested in horizontal and vertical orientations at two different loading rates of 5 and 10 mm/min. For Alikazemi variety of paddy rupture force values decreases as the loading rate increased from 5 to 10 mm/min. The rupture force values for horizontal orientation were 125.69 and 33.51 N at a loading rate of 5 mm/min and 10 mm/min, respectively. Similarly, the rupture force values for vertical orientation were 117.38 and 33.51 and the loading rate of 5 mm/min and 10 mm/min respectively [2]. In another study, Lu et al., the compressive force and bending force of rough rice, brown rice, and white rice were determined by conducting compression and three-point bending tests. It was observed that thinner kernels required low breaking force and therefore lead to more breakage during milling [4].

I. Bagheri et al., investigated the mechanical strength of 12 varieties of brown rice at four levels of moisture content 8, 10, 12, and 14 wet basis. Brown rice grains were quasi-statically loaded at two rates 10 and 15 mm/min. The results show that the breaking force decreases with increasing moisture content and also with an increase in a loading rate. The breaking of brown rice with a lower level of moisture content was a sudden failure with less deformation and breaking of brown rice with a higher level of moisture content was gradually crushing with more deformation [5]. For three different levels of amylase ratio, the mechanical properties of different varieties of rice grains were invested by Y. J. Shu et al. The mechanical properties

were measured using compression and three-point breaking test. According to their study, the rice variety with the high amylose ratio has the highest compression strength. Their study can help to develop methods and devices used for milling processing of rice with different levels of amylose to reduce the milling losses [6]. The another study carried out by Ponlakrit Kumklam, in which the strength of rice grains measured by breaking force and study done on the correlation of breaking force and head rice yield. Bending test was done by keeping fixed base support distance of 3.5 mm with flat loading head of thickness 1.5 mm and deformation rate was 1.5 mm/min. From the test, the bending force was closely related to the head rice yield [7].

The main objective of this research was to measure physical properties such as length, width, and thickness and to determine the mechanical property of five different varieties of brown rice grains in terms of rupture force under three-point breaking test and examining the effect of rice variety and number of fissures on rupture force.

89.2 Materials and Methods

The brown rice varieties Rashi Poonam, Kolam, Komal, Basmati, and Daptari used for this test were obtained from one of the rice mills from Palghar, Maharashtra, India. The hot-air oven-drying method was used to determine initial moisture content of the rice samples. The initial moisture contents of grains were 5.35, 5.69, 4.98, 6.13, and 5.42% (wet basis) for Rashi Poonam, Kolam, Komal, Basmati, and Daptari varieties, respectively.

89.2.1 Detection of Fissures in Rice Grains

This study investigates the rupture force for five different rice varieties with the number of fissures in rice grains. To detect fissures in rice grains, fissure detection technique is newly developed. Imaging of the fissures in rice kernels was done with the help of DSLR camera. Camera used was Nikon D5200 with 18–55 primary lens and close up lenses/secondary lenses (numex 52 mm close up +1, +2, +4 and +10) were additionally attached for better zooming. For better zooming without the use of secondary lenses can be achieved by attaching a primary lens in reverse direction to the camera. For better results, the camera was on a tripod stand.

The light source was white LED bulb (Wipro LED, 9 W). All samples were placed in plain brown-colored glass, and images were captured using DSLR camera. After capturing the images of rice grains, all are edited in Lightroom CC image-editing software (version 3.4) for better clarity of fissures. While editing, clarity was increased to full value, highlights, shadow, and whites reduced to zero value and exposure, the contrast was adjusted according to images and finally, the

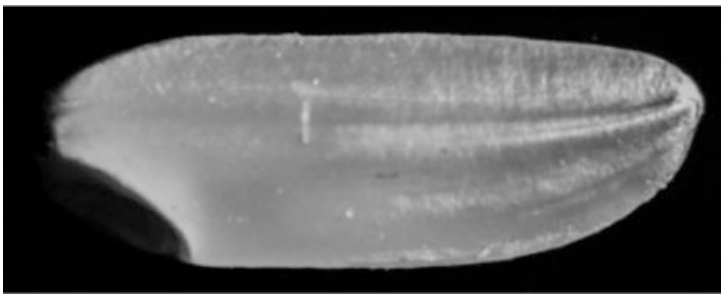


Fig. 89.1 Image of brown rice with no fissure

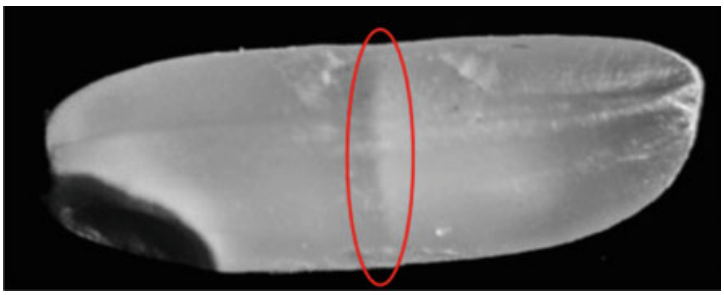


Fig. 89.2 Image of brown rice with one fissure

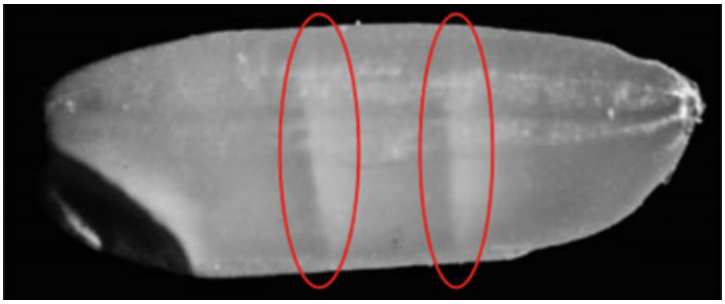


Fig. 89.3 Image of brown rice with two fissures

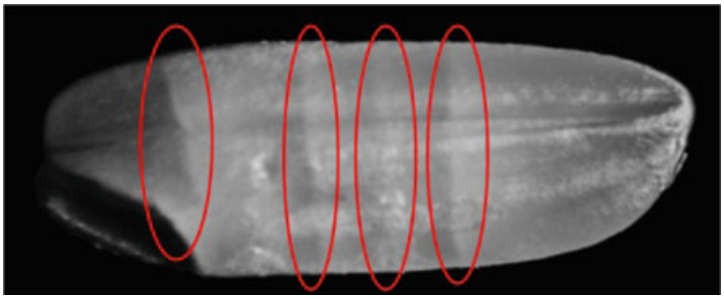


Fig. 89.4 Image of brown rice with more than two fissures

black and white effect was given to images. Figure 89.1 shows the images of brown rice with a different number of fissures. Figures 89.1, 89.2, 89.3, and 89.4 show the brown rice with 0, 1, 2, and more than 2 fissures, respectively.

89.2.2 Types of Fissures Observed in Rice Grains

Various types of fissures are mentioned in Figs. 89.5 and 89.6 show the image of brown rice grains (Basmati rice) with different types fissures.

From Fig. 89.5,

Type A fissures—Fissures perpendicular to longitudinal axis of grain—Non-terminating type: The fissures present in a rice grain which are perpendicular to the longitudinal axis of grain and ended somewhere in rice grain.

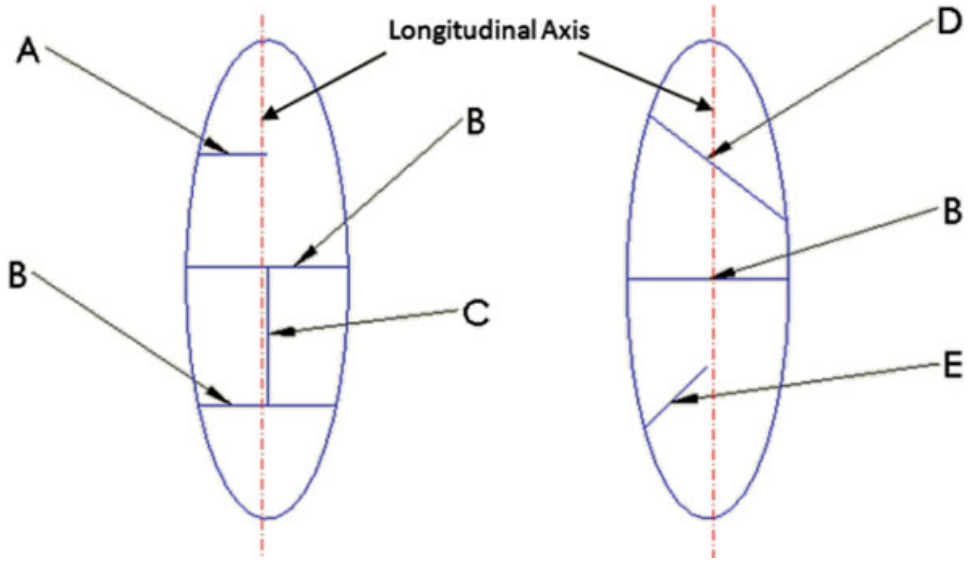


Fig. 89.5 Types of fissures in rice grains

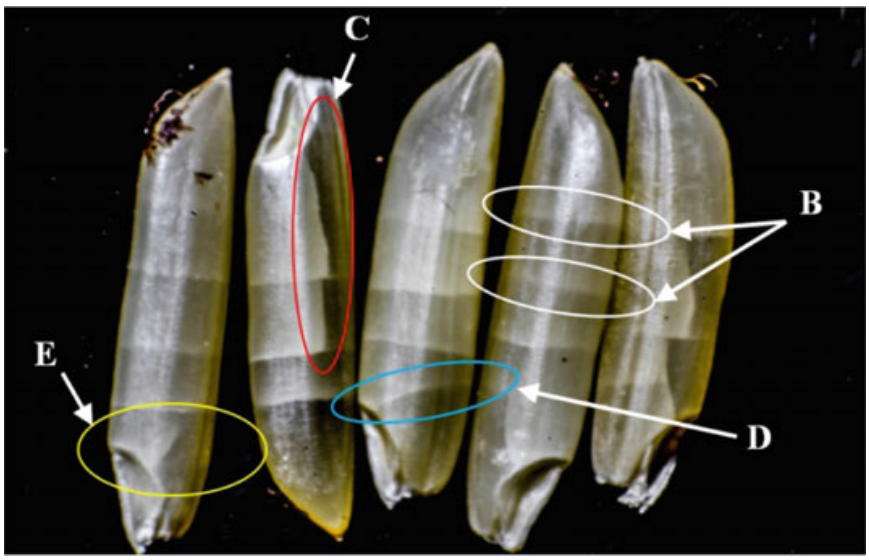


Fig. 89.6 Image of detection of different types of fissures in rice grains

Type B fissures—Fissures perpendicular to longitudinal axis of grain—Terminating type: The fissures present in a rice grain which are perpendicular to the longitudinal axis of grain and passed throughout the rice grain along that axis.
Type C fissures—Fissures parallel to longitudinal axis of grain—Terminating type: The fissures present in a rice grain which are parallel to the longitudinal axis of grain and passed throughout the rice grain along that axis.
Type D fissures—Cross fissure—Terminating type: The fissures present in a rice grain which are neither parallel nor perpendicular to the longitudinal axis of grain and ended somewhere in rice grain.
Type E fissures—Cross fissure—Non-terminating type: The fissures present in a rice grain which are neither parallel nor perpendicular to the longitudinal axis of grain and passed throughout the rice grain.

It was observed that in the case of fissured kernels the percentage of fissures perpendicular to longitudinal axis of grains (A and B-types) was more than any other types of fissures. Each and every fissured kernel had A and B type of fissures present in it. Percentage of fissures parallel to longitudinal axis of grains(C type) were more in Basmati rice variety than any other rice variety.

89.2.3 Physical Characteristics of Rice Grains

The shape of the rice kernel was found to be ellipsoid shape with three dimensions —length, width, and thickness. The physical dimensions of 20 kernels of each variety were measured using micrometer screw gauge (Mitutoyo Corporation) having least count of 0.01 mm.

The measured lengths of Rashi Poonam, Kolam, Komal, Basmati, and Daptari rice grains were found 5.79 ± 0.2 mm, 5.41 ± 0.19 mm, 6.64 ± 0.21 mm, 7.66 ± 0.33 mm, 5.45 ± 0.18 mm, respectively. From the dimensions of rice grains, it can be seen that the length of Basmati rice variety was more than other types of rice varieties. So, Basmati was long rice grain while others were short/medium rice grains. Physical dimensions of five types of rice varieties are given in Table 89.1.

Table 89.1 Physical dimensions of five types of rice for three-point breaking test

S. No.	Type of rice	Length (mm)	Width (mm)	Thickness (mm)
1	Rashi Poonam	5.79 ± 0.2	2.0 ± 0.08	1.57 ± 0.04
2	Kolam	5.41 ± 0.19	1.65 ± 0.07	1.52 ± 0.05
3	Komal	6.64 ± 0.21	2.08 ± 0.07	1.57 ± 0.04
4	Basmati	7.66 ± 0.33	1.97 ± 0.06	1.75 ± 0.06
5	Daptari	5.45 ± 0.18	1.62 ± 0.15	1.54 ± 0.05

89.2.4 Three-Point Breaking Test

The mechanical property of brown rice grains was determined in terms of average rupture force in three-point breaking test for five different varieties Rashi Poonam, Kolam, Komal, Basmati, and Daptari. The experiment was performed at a loading rate of 1 mm/min. Three-point breaking test was performed using a Universal testing machine (Instron Universal Testing Machine-3345, Fiest Laboratory, MEMS Department, IIT Bombay).

For testing, all rice grains were de-husk by hand. In each batch, 20 rice samples were taken for testing. Rice samples were sorted out according to the number of fissures present in it. Fissured were observed by fissured detection technique which was newly developed with the use of DSLR camera. For three-point bending tests, the whole rice grain was kept on an adjustable base support in which span was adjustable. The diameter of the end of support and compression tip was 1.45 mm. As rice shape is narrow toward the ends, the 15% of overhang was considered for three-point breaking test so that rice grains cannot break at their end due to the small area as compared to the middle portion. The adjustable base supports for three-point test apparatus for the breaking of a rice grain is shown in Fig. 89.7.

Different parts of the adjustable fixture for three-point breaking test are:

1. Base of fixture.
2. Two guiding shafts: for guiding movable supports.
3. Two movable supports: for support the rice grain while testing.
4. Circular tip on each support: for reducing stress concentration.
5. Main threaded shaft: half anticlockwise and another half clockwise threading.
6. Handle for rotation of the main threaded shaft.

The position of three points for breaking test:

The position of three contact points while loading is shown in Fig. 89.8. Compression Pouch is always in the middle of rice grain.

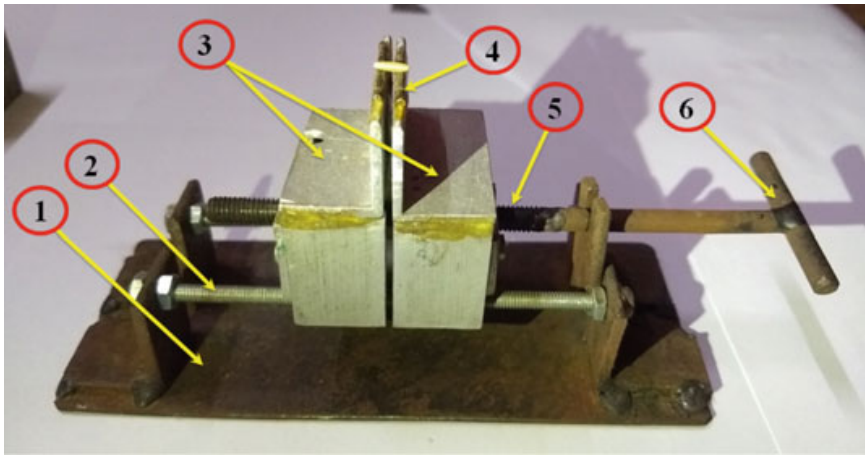


Fig. 89.7 Adjustable base supports for three-point breaking test

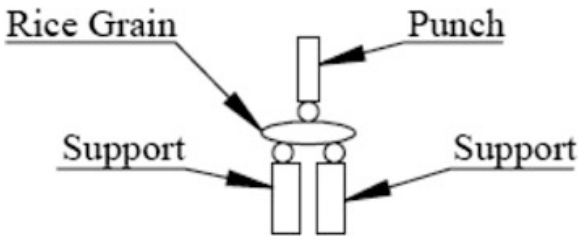


Fig. 89.8 Position of three contact points

89.3 Results

Table 89.2 shows the rupture force value for different varieties of brown rice grains. Data shows that variety and number of fissures affect the rupture force. In the case of Basmati variety, the average force to rupture the grain was obtained as 18.69 ± 6.98 N, 9.56 ± 5.63 N, 6.32 ± 3.36 N, and 3.49 ± 1.38 N for grains without fissures, with 1 fissure, 2 fissures, and more than 2 fissures, respectively. Similarly, rupture force values for Kolam, Komal, Daptari, and Rashi Poonam variety are given in Table 89.2. From data, it can be seen that rupture force value decreases as fissures increase in rice grains. Rashi Poonam variety has the highest average rupture force of value 32.29 N for grains without fissures. While Komal variety, the rupture force value 28.52 N which is much closer to the value of rupture force for Rashi Poonam variety.

89.3.1 Finite Element Analysis of Rice Grain (Three-point Breaking Test)

Finite element analysis of rice grain under three-point breaking test was done in ANSYS Workbench Software. As rice grain is having complex geometry, for analysis, rice grain is considered as a cylindrical object with diameter equal to equivalent diameter of rice grain calculated using formula:

Table 89.2 Rupture force by three-point breaking test

S. No.	Type of rice	Breaking force (N)			
		0 fissure	1 fissure	2 fissures	More than 2 fissures
1.	Basmati	18.69 ± 6.98	9.56 ± 5.63	6.32±3.36	3.49 ± 1.38
2.	Kolam	15.44 ± 6.57	10.30 ± 4.72	5.21 ± 4.69	3.81 ± 2.08
3.	Komal	28.52 ± 4.41	16.83 ± 4.39	11.29 ± 3.56	6.34 ± 2.21
4.	Daptari	16.23 ± 5.98	12.36 ± 3.87	7.12 ± 2.28	4.51 ± 1.13
5.	Rashi Poonam	32.29 ± 6.55	14.68 ± 5.89	9.90 ± 2.70	6.96 ± 3.06

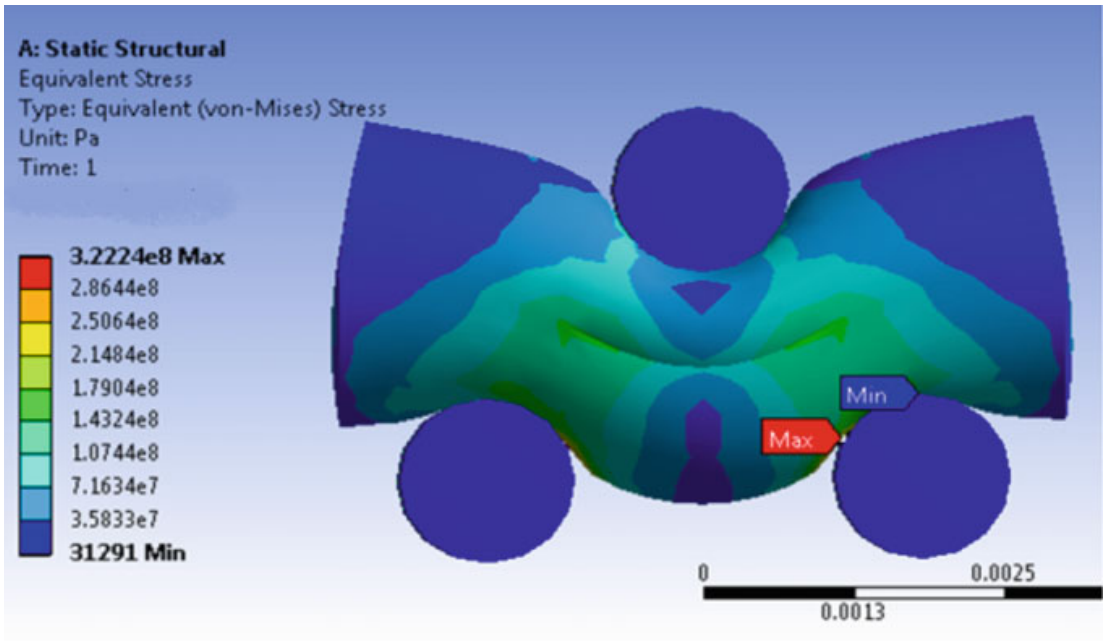
$$\mathrm{De} = \left(L * \frac{(W+T)^2}{4} \right)^{1/3}$$

where L = Length, W = Width of rice grain, T = Thickness of rice grain [8]. Mechanical properties of rice grains considered for analysis were: Young Modulus = 417MPA, Poisson's ratio = 0.4476, density 2700 kg/m^3 [9]. Physical dimensions of Rashi Poonam rice variety were taken for analysis. Figure 89.13 shows the finite element analysis of rice grain under three-point breaking test, and Fig. 89.9a, b shows the analysis results for the equivalent stress and equivalent strain, respectively.

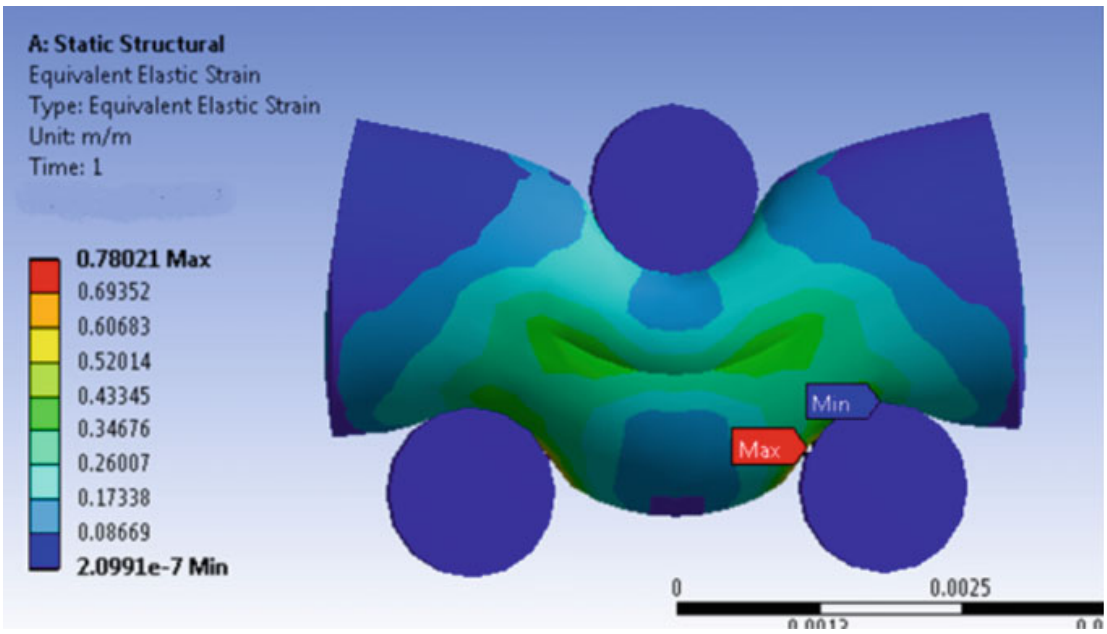
From analysis result, it can be seen that for 7% of strain, the equivalent stress in rice was 322 MPa which at the contact of rice bottom surface and support. The crack may occur at contact of support and bottom surface of rice grain where the stress intensity is maximum. In experimental results for Rashi Poonam rice variety without any fissure, the maximum stress was found was 430 MPa and maximum strain was 9%. In three-point bending test, the crack started at contact of bottom surface of rice grain and support and then propagate toward the contact of punch to rice grain as shown in Fig. 89.10. High-speed camera used for recording the crack propagation was Phantom MIRO M110 with supporting software—Phantom Camera Control Application for controlling operation of camera. Illumination is done by fiber optic dual light illuminator (White light source).

89.3.2 A Typical Graph of Load (N) Versus Deformation (mm) Obtained After Three-point Breaking Test of Rice Grains

From load (N) versus deformation graph as shown in Fig. 89.11, the testing starts at point A, i.e., punch moves downward with a set speed (1 mm/min). The region from point A to point B is almost parallel to *X*-axis. This region shows that up to point B, there is no contact of punch with rice grain. (This is due to we touch punch to rice grain manually, so for every sample, it may vary). From point B, actual loading gets started. The region from point B to point C shows the linear relationship between load and extension. At point C, the breaking of rice grains occurs.



(a) Analysis for Equivalent Stress



(b) Analysis for Equivalent Strain

Fig. 89.9 FEM analysis of the rice grain with failure area of fissure in the three-point bending loading. **a** Analysis for equivalent stress, **b** analysis for equivalent strain

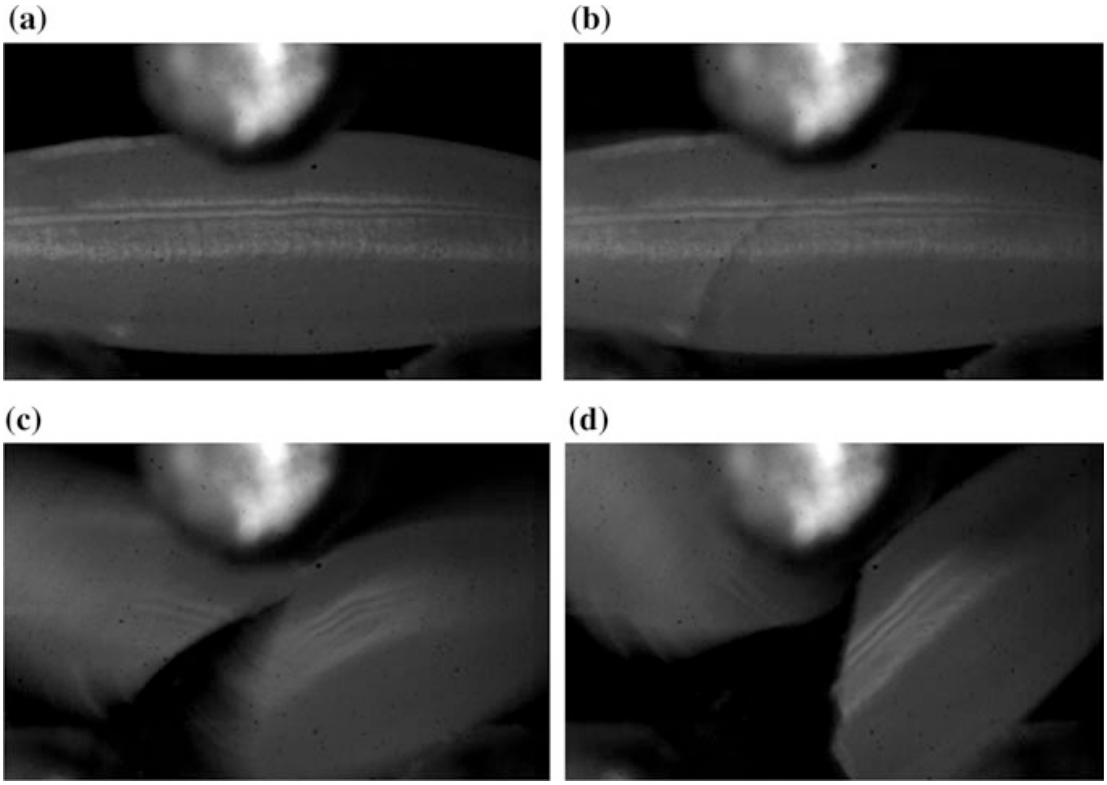


Fig. 89.10 Images of the crack propagation in rice grain while applying the load

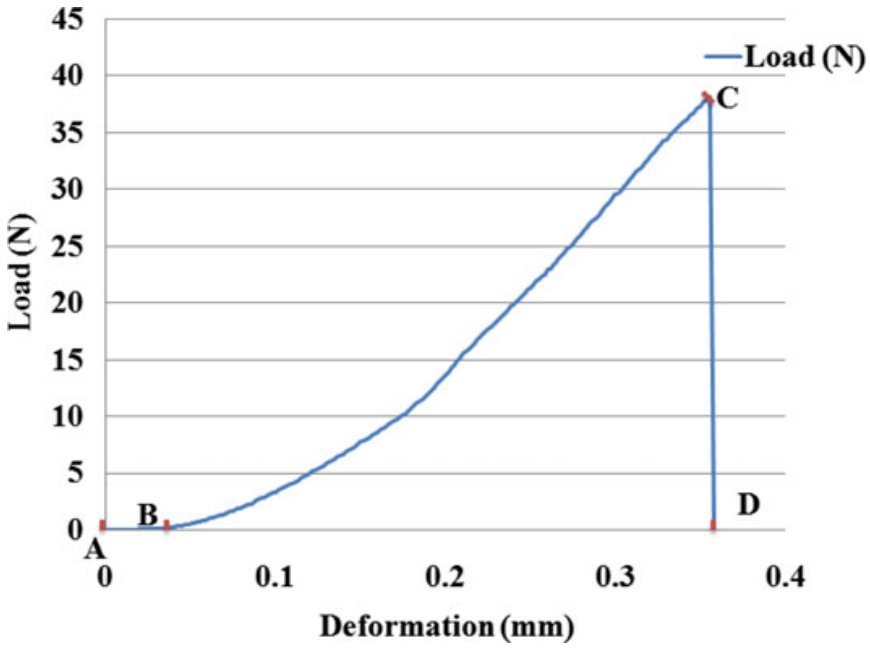


Fig. 89.11 Typical graph of load (N) versus deformation (mm)

89.3.3 Modified Graph of Load (N) Versus Deformation (mm) Obtained After Three-Point Breaking Test of Rice Grains

This graph is obtained by removing Region A-B and Region C-D is from the typical graph. This graph shows the linear relationship between load and deformation (Fig. 89.12).

89.3.4 Graph of Breaking Force Versus no of Fissures for Five Varieties of Rice Grains

Figure 89.13 shows the graph of breaking force value for five different types of rice varieties with a different number of fissures present in the rice kernel. From the

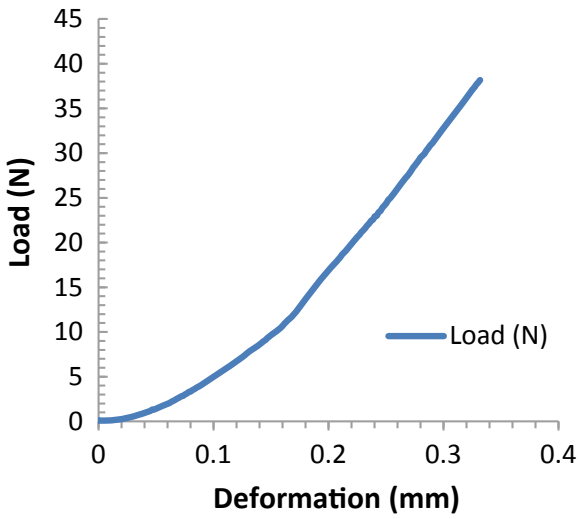


Fig. 89.12 Graph of load (N) versus deformation (mm)

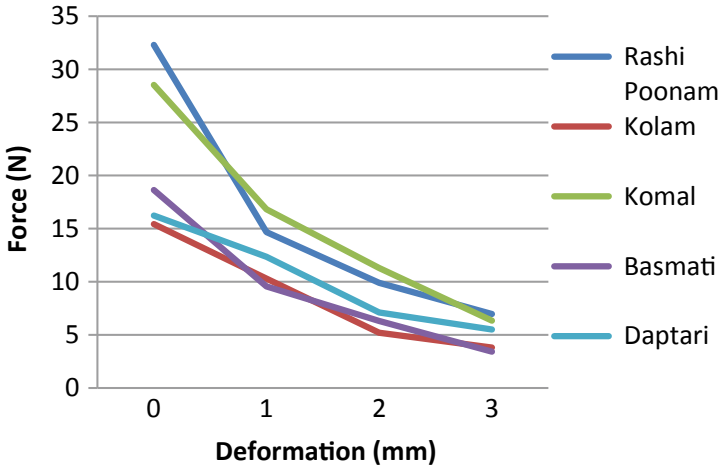


Fig. 89.13 Graph of breaking force versus number of fissures (*On X-axis: Point 3 indicates more than two fissures)

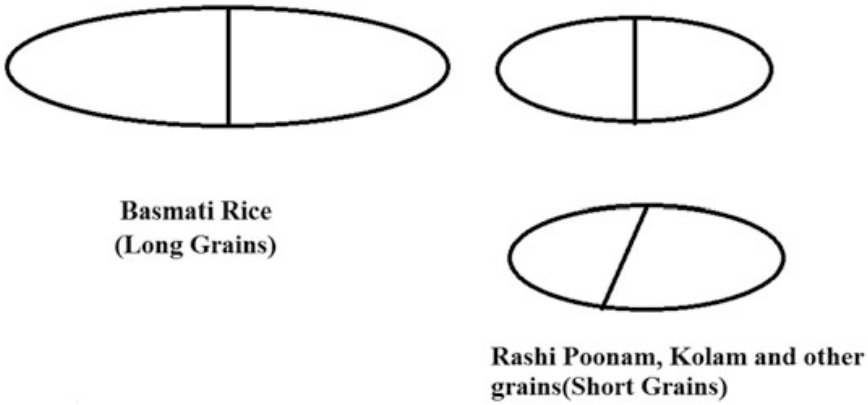


Fig. 89.14 Breaking patterns of rice grains while three-point breaking test

graph, it can be clearly seen that the breaking force required decreases as the number of fissures increases in rice grain. Rice grains become so weak if there are more than two fissures in a grain and for those rice grains breaking force values are much closed irrespective of rice varieties.

89.3.5 Different Breaking Patterns of Rice Grains While Three-Point Breaking Test

After testing and imaging of rice grain it is oberved that rice kernels with fissures subjected to 3-point breaking test, crack initiated at big fissure. When there was no fissure in rice grain, then crack is initiated from the contact point of one of the base support and propagates upward toward the point of contact between rice grain and punch. Each case of the rice grain, crack propagated from bottom surface to top surface of rice grain while loading because of bottom surface always in tension while 3-point breaking tests (Figure 89.14).

89.4 Conclusions

The three-point breaking test has been performed on five varieties of brown rice grains with different types of fissures present in brown rice grains. The values of the rupture force of brown rice grains Rashi Poonam and Komal rice varieties were taken higher than rupture force value of brown rice grains. So, it can be concluded that the Rashi Poonam and Komal varieties have more resistant to breakage than the Basmati, Kolam and Daptari varieties during the various operations such as de-husking, milling, and polishing. The results obtained from three-point breaking test of brown rice grains can help for developing of methods and types of equipment used for whitening and polishing processing of rice with different varieties.

Finite element analyis is result have validated the cracks are likely to occur due to the applied the rupture forces on the rice grain in the three point bending test. If rice millers able to find breaking force value for particular variety of rice grains before processing, then it may easy for rice millers to adjust correct pressure for processing that particular batch of rice grains so further breaking can be avoided. The testing of rice grain can be done on the low load capaicty carrying universal testing machine shall be recommeded for better rupture analysis of rice grain in the experimental testing.

Acknowledgements We thank Prof. Uday Sheorey and Prof. Rakesh Mote (Professor, Mechanical Department, Indian Institute of Technology, Bombay) for their excellent guidance throughout the project. We also thank Metallurgical Engineering and Material Science Department, IIT Bombay, for support to carry out various testing.

References

1. Mahajan, G., Kumar, V., Chauhan, B.S.: Rice production worldwide. Springer International Publishing AG, Switzerland (2017)
2. Zareiforoush, H., Komarizadeh, M.H., Alizadeh, M.R.: Mechanical properties of paddy grains under quasi-static compressive loading. N. Y. Sci. J. **3**(7), 40–46 (2010)
3. Ghadge, P., Prasad, K.: Some physical properties of rice kernels: variety PR-106. J. Food Process Technol. **3**(8), 1–5 (2012)
4. Lu, R., Siebenmorgen, T.J.: Correlation of head rice yield to selected physical and mechanical properties of rice kernels. Am. Soc. Agric Eng. **38**(3), 889–894 (1998)
5. Bagheri, I., Dehpour, M.B., Payman, S.H., Zareiforoush, H.: Rupture strength of brown rice varieties as affected by moisture content and loading rate. Aust. J. Crop Sci. **5**(10), 1239–1246 (2010)
6. Shu, Y.J., Liou, N.-S., Moonpa, N., Topaiboul, S.: Investigating damage properties of rice grain under compression load. In: International Conference on Experimental Mechanics and Twelfth Asian Conference on Experimental Mechanics, Proceedings of SPIE 923402, pp. 1–5, Bangkok, Thailand (2014)
7. Kumklam, P.: Predict head rice yield with mechanical strength of kernels. J. Appl. Sci. Res. **9** (12), 6107–6110 (2013)
8. Ravi, P., Venkatachalam, T.: Important engineering properties of paddy. Sci. J. Agric. Eng. **4**, 73–83 (2014)
9. Chattopadhyay, P.K.: The rheological properties of rice grain. J. Food Process. Eng. **17**, 1–17 (1994)

Chapter 90
Finite Element Analysis of Potential Liner Failures During Operation in Spherical Pressure Vessel

R. Pramod, N. Siva Shanmugam, C. K. Krishnadasan and K. Sankaranarayanasamy

Abstract Composite overwrapped pressure vessel (COPV) is widely used in aerospace industry to contain pressurized gases. The behaviour of pressure vessel components (liner and composite overwrap) at applied pressure and potential failures during its operation requires adequate study. The finite element (FE) analysis is performed for potential failures occurring in a 420 mm diameter aluminium alloy (AA) 6061-T6 lined spherical pressure vessel. In the present study, the design of spherical liner with apt manufacturing techniques is presented. The liner contribution towards sustaining the minimum burst load is determined and liner burst test is conducted to check the location of failure. Further, typical liner failures like buckling and bond failure occurring due to inefficient contact between the liner and composite overwrap, which eventually forms a debond region are also studied. The liner burst pressurization results predicted the location of failure in liner at transition region as the liner is undergoing thickness changes in that particular region. The mode of liner buckling, buckling load is predicted when liner is subjected to compression during depressurization cycle. Further, at debond region, the liner is unable to transfer pressure load effectively to composite overwrap experiencing severe deformation.

Keywords Finite element analysis · Liner · Buckle · Debonding

R. Pramod · N. S. Shanmugam (✉) · K. Sankaranarayanasamy
Department of Mechanical Engineering, National Institute of Technology, Tiruchirappalli 620 015, Tamil Nadu, India
e-mail: nsiva@nitt.edu

C. K. Krishnadasan
Structural Design and Analysis Division, Vikram Sarabhai Space Centre, Thiruvananthapuram 695 022, Kerala, India

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3_90

90.1 Introduction

Launch vehicle and satellite require pressure vessels during operation for managing fluids and propulsion systems. Composite overwrap pressure vessel (COPV) with inclusion of a metal liner offers less vessel weight in comparison to conventional metal pressure tanks [1]. Aluminium and titanium alloys are generally used materials for liner because of less density and corrosion-resistant properties exhibited by these materials, respectively [2]. Non-load sharing and load sharing are the two types of liner. In the first case, liner is having minimum thickness and weight designed primarily to provide permeation barrier for preventing escape of gases. In the latter case, the liner contributes towards sharing the applied pressure load with higher thickness and weight [3]. Lee et al. [4] designed a spherical pressure tank using the Ti-15 V-3Al-3Cr-3SN material. Hemispheres were made by hot spin forming and blow forming process followed by TIG welding to join the two hemispheres. Pressurization test was conducted; the strain and acoustic emission signals were obtained to examine the outcome of selected manufacturing technique on the modes of failure and working performance of the tank. Chandran et al. [5] performed FE analysis by designing and analysing the COPV using titanium liner and T-300 carbon fibre composite. The effect of debonding location and length of debond in the liner-composite interface for weld and transition region was investigated. Almeida Jr et al. [6] investigated the effect of load sharing ability of aluminium and stainless steel liner material with various thicknesses in cylindrical liner geometry. Further, analyzed the stress and strain distribution at various regions of liner and composite structure. Many research works have been carried out in developing pressure vessels using AA 6061-T6 liner. The numerical analysis performed mainly focused on the modelling techniques and analysis of composite overwrap, failure analysis of composite overwrap and methods to analyse filament wound COPV [7–9]. From the literature review, it is noted that the research work has not focused on the analysis of failure and its modes in liner under applied internal pressure. Studying the behaviour of liner by physically testing the pressure vessel by subjecting it to repeated testing shall be economically high and time-consuming as it comprises of conducting several trials to obtain in-depth information generally involved in a design process. Instead, numerical analysis is conducted using the FE method, as an alternative to experimental testing [10]. In the present case, it is emphasized on cost reduction by replacing titanium with aluminium liner material and will be helpful in meeting quality conditions effectively. Hence, in the present analysis using ABAQUS and FE analysis software, liner analysis is conducted to study its behaviour, when subjected to internal burst pressure. Also, contact failure issues such as buckling and debonding, which lead to loss of COPV are also analyzed.

90.2 Materials and Design

The selected liner material in COPV is aluminium alloy (AA) 6061-T6. Assigned physical [11] and mechanical material properties are shown in Tables 90.1 and 90.2, respectively. The spherical liner design is shown in Fig. 90.1, as being a non-contributory type liner, the thickness in dome region is kept minimum to achieve weight savings. The formed part is fabricated by forming a 3.5 mm thick

Table 90.1 Physical properties of AA 6061-T6

Young's modulus, MPa	Poisson's ratio	Density, kg/mm^3
73,100	0.34	2.76×10^{-6}

Table 90.2 Mechanical properties for base and weld metal AA 6061-T6

Part	Yield stress, MPa	Yield strain, %	Tensile stress, MPa	Tensile strain, %
Base	240	7	290	19
Weld	140	4	180	8

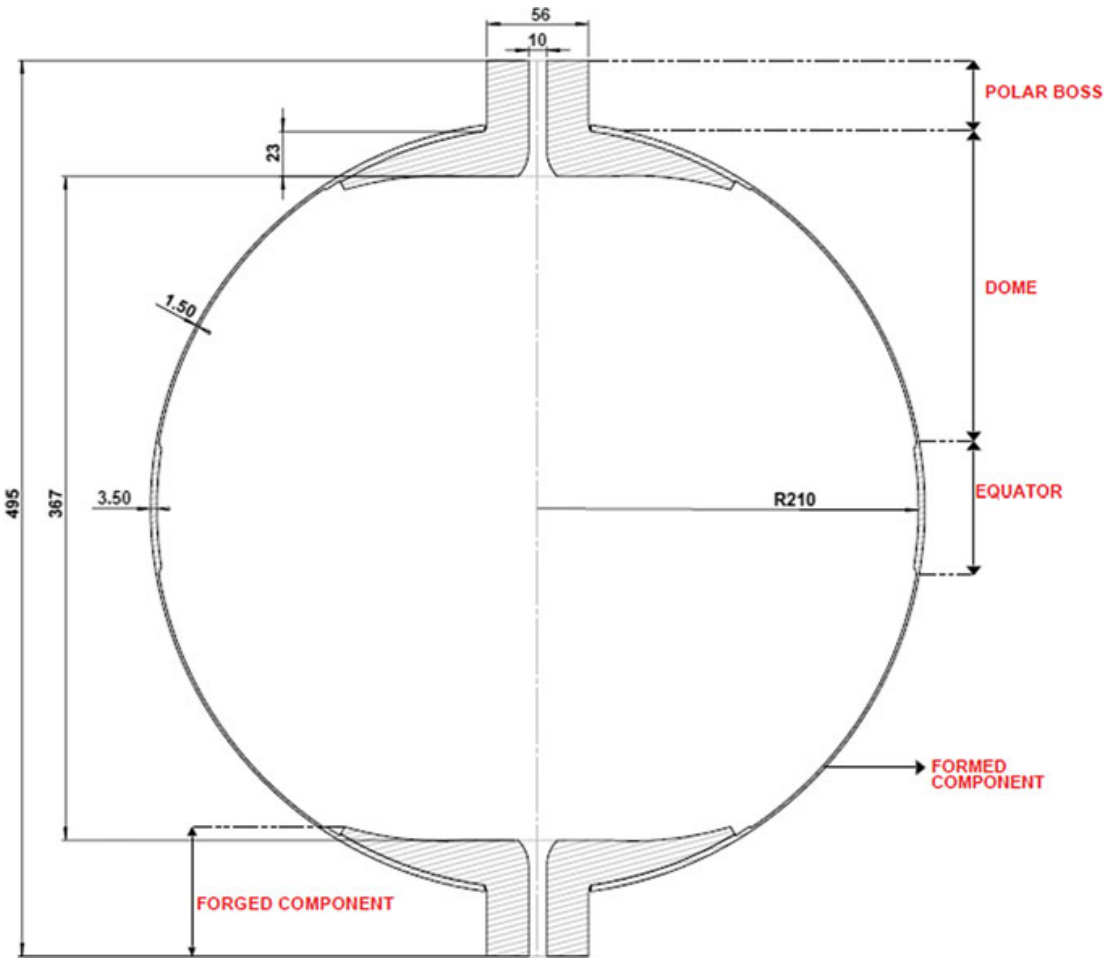


Fig. 90.1 Pressure vessel liner configuration (dimensions in mm)

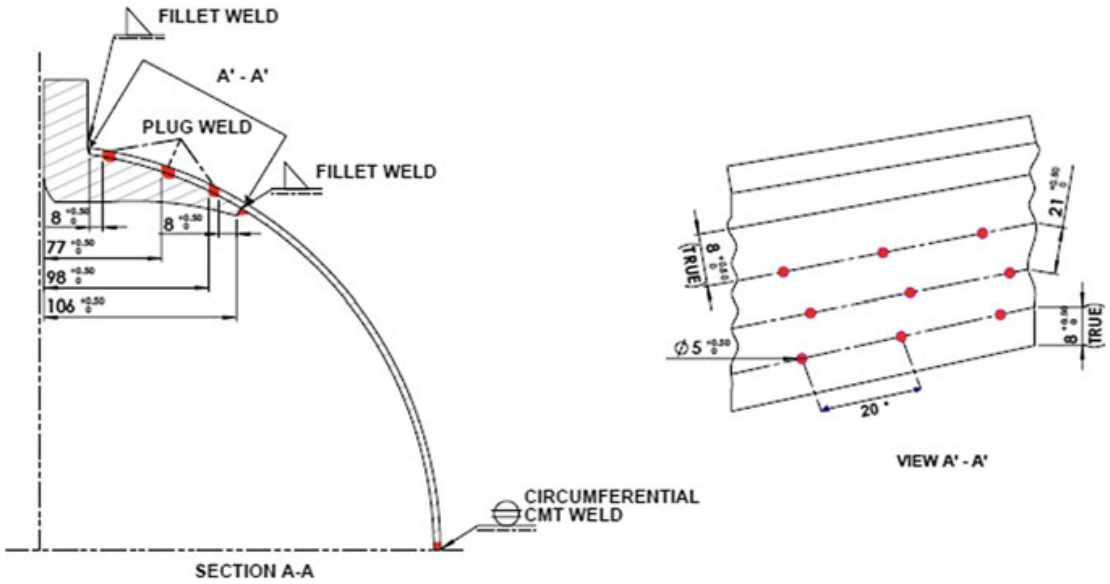


Fig. 90.2 Representation of plug, fillet and CMT welding locations in the liner

sheet using die and punch system, and forged part for end boss is obtained by machining a solid block of AA 6061-T6. The parts are further joined by plug welding technique, and welding the hemispheres circumferentially using cold metal transfer welding (CMT) technique as shown in Fig. 90.2. The true distance between the points of laying plug weld to edge of fillet weld is maintained at 8 mm with each plug having radius of 5 mm. Between each plug weld, a distance of 21 mm is maintained and positioned at three stages with pitch angle of 20° as shown in Fig. 90.2 (section A′–A′).

90.3 Finite Element Modelling

The liner is modelled as solid part. The effect of welding process changes the inherent base material properties in aluminium. Hence, by considering the effects due to welding, weld metal properties are assigned at pole and equator region, and length to be assigned is determined from an experimental trial. The length for assigning weld metal properties in liner is determined from Fig. 90.3. A miniature level AA 6061-T6 tube test specimen (replicating the large-scale liner structure) having a thickness of 3.5 mm were subjected to CMT welding to determine the weld bead width (BW) and heat-affected zone (HAZ). It was observed that BW and HAZ comprised of 15 mm (refer Fig. 90.3a) and this was measured using Struers welding expert with magnification of 20X. Due to less heat input of CMT process, various researchers have observed narrower heat-affected zone in AA 6061

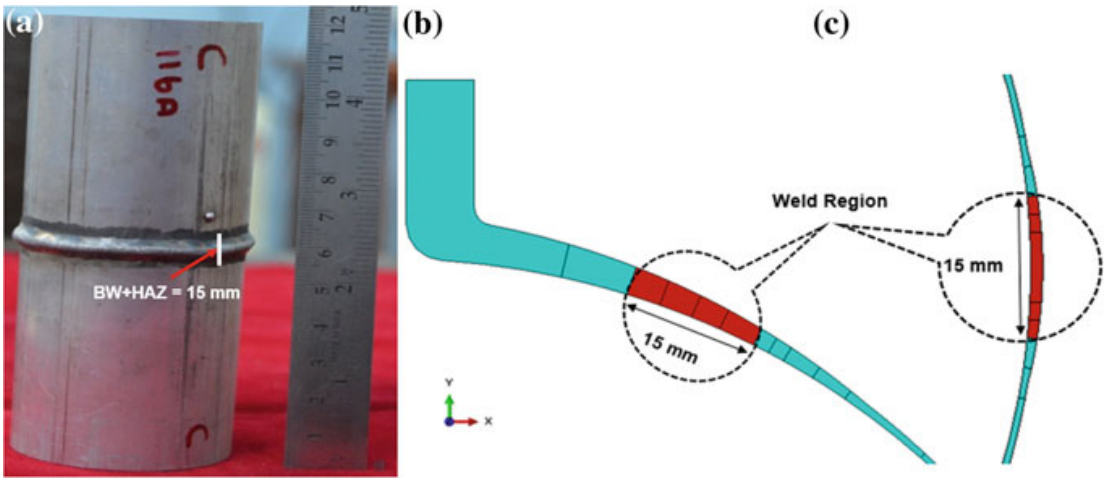


Fig. 90.3 Weld bead and HAZ measurement in CMT weld AA 6061-T6 pipe specimen (Left) and Weld region assignment in the FE modelled liner (Right)

weldment using CMT [12, 13]. In the present experimental trial, (refer Fig. 90.3a) the HAZ is found negligible as measured along with weld bead. Hence, weld bead + HAZ is considered as a single region and measured to be 15 mm in length. In the present analysis, weld material properties are assigned for a length of 15 mm as shown in Fig. 90.3b in the pole region and Fig. 90.3c in the equator region. The symmetric boundary condition is applied on side faces of liner by restricting *Z*-axis movement (U3 = 0). Also, bottom face of liner is restricted by *Y*-axis movement (U2 = 0) and inlet face movement is restricted by *X*-axis movement (U1 = 0), respectively as shown in Fig. 90.4. The internal pressure load is applied on the interior face of liner (refer Fig. 90.4). Liner burst test is conducted to determine its pressure load contribution and check the failure location due to applied load. The applied burst pressure load on the liner is determined using the Eq. 90.1.

$$P = \frac{2 * \sigma * t}{R}$$
$$P = \frac{2 * 290 * 1.5}{210} = 4.15\,\text{MPa} \tag{90.1}$$

where σ = Ultimate tensile strength of AA 6061-T6, t = Minimum thickness of liner and R = Radius of liner. Dynamic explicit analysis is conducted with non-linear geometric effects. For prediction of failure, Johnson–Cook (JC) damage material model is utilized. The failure is expected when parameter $D = 1$. The mathematical expression is shown in Eqs. 90.2 and 90.3. The constants for the JC model used in the evaluations for AA 6061-T6 is shown in Table 90.3 [14]. Load angle (ξ) effect is considered as $\xi = 0$.

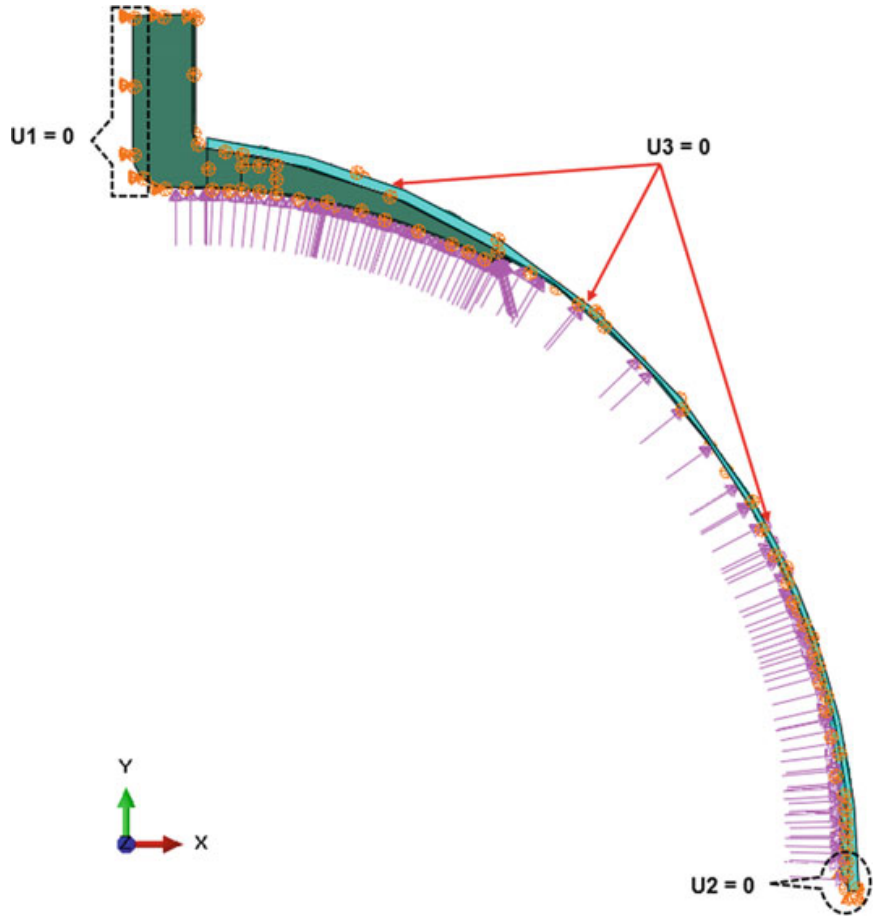


Fig. 90.4 Load and boundary conditions assigned on liner

Table 90.3 JC damage parameters assigned for AA 6061-T6

D_1	D_2	D_3	D_4	D_5
−0.77	1.45	−0.47	0.00	1.60

$$D = \sum \frac{\Delta\varepsilon}{\varepsilon_f} \tag{90.2}$$

$$\bar{\varepsilon}_f = [D_1 + D_2 \exp D_3\, \sigma^*][1 + D_4 \ln \dot{\varepsilon}^*][1 + D_5 T^*] \tag{90.3}$$

where

$\bar{\varepsilon}_f =$ Strain at fracture
$\dot{\varepsilon}^* =$ Strain rate
$T^* =$ Temperature
D_1, D_2, D_3, D_4 and D_5 are material constants.

Table 90.4 Description of meshing parameters

Part	Shape of element	Meshing technique	Elements
Plugs	Hex-dominated	Sweep	5891
Formed part	Hex-dominated	Free	7158
Forged part	Hex-dominated	Sweep	1038
Fillet	Hexagonal	Sweep	168

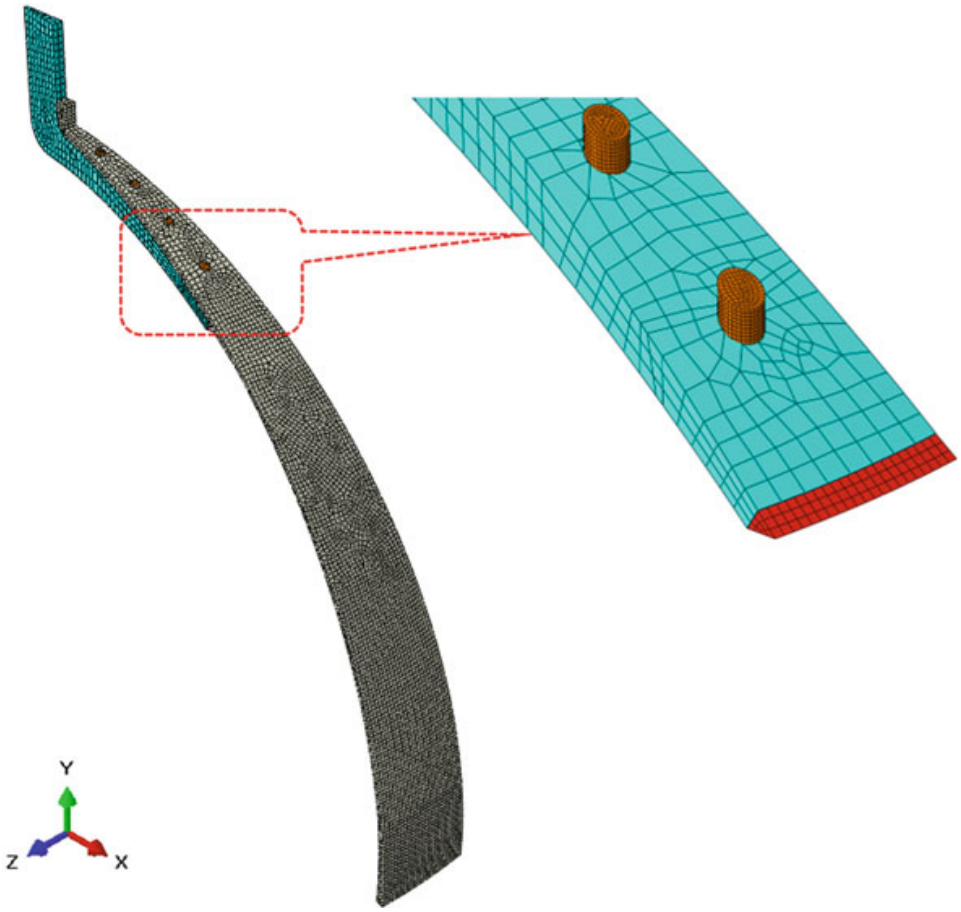


Fig. 90.5 Meshed model of liner and mesh profile in plug, fillet weld (inset)

The meshing of liner is performed by maintaining fine mesh near the plug weld region and on plug welds. The primary element type used for meshing parts is C3D8R (An eight node linear brick, reduced integration with hourglass control).

Table 90.4 shows a brief description of meshing parameters utilized in meshing of parts and Fig. 90.5 shows the meshed model of liner and mesh profile in plug and fillet weld.

The stable bond between liner and composite overwrap shall fail in case of design irregularities, complex curvatures in liner, thus creating a debond region. In the present analysis, a debond region is created by removing the contact interaction between liner and composite overwrap to study the ill-effects on liner. The eigenvalue linear analysis is performed for determining the load at which liner buckles during depressurization with subsequent buckling modes.

90.4 Results and Discussion

The burst load of 4.32 MPa applied on inner surface of liner at plastic stage showed severe yielding with Von Mises stress of 309 MPa (refer Fig. 90.6a) and plastic strain of 0.76% (refer Fig. 90.6b) at the transition regions near to pole and equator region. Further, the first failure location of liner is seen at transition region near the

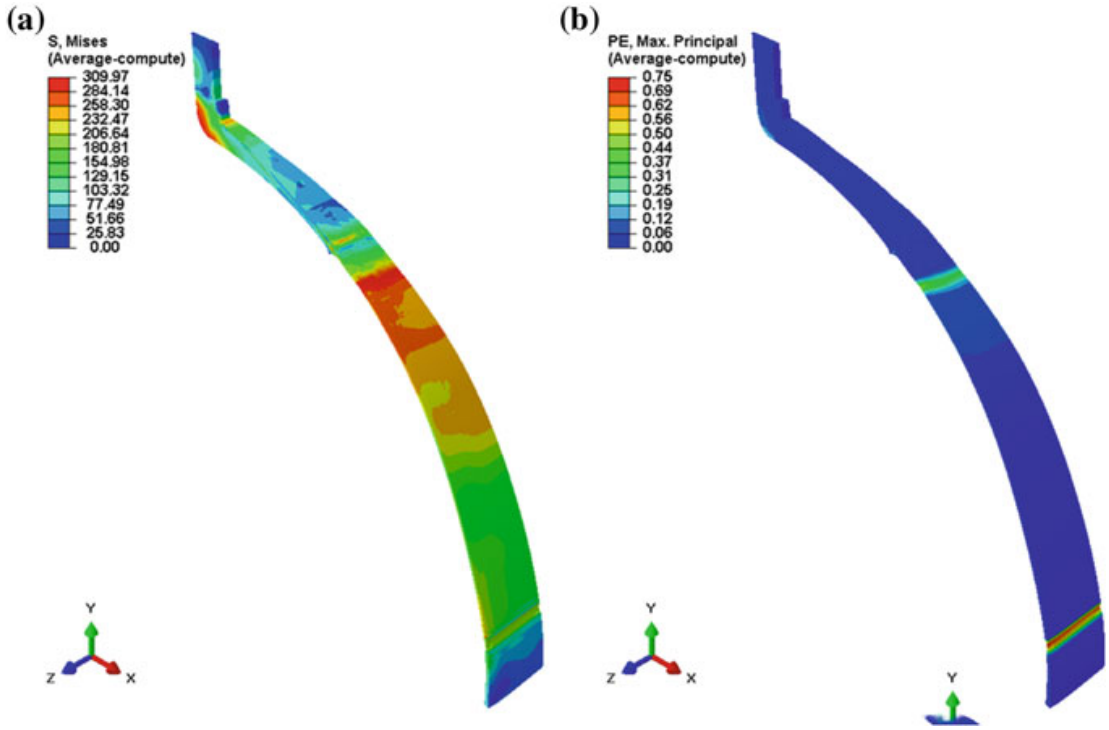


Fig. 90.6 Distribution of Von Mises stress **a** and plastic strain **b** at plasticity stage in liner

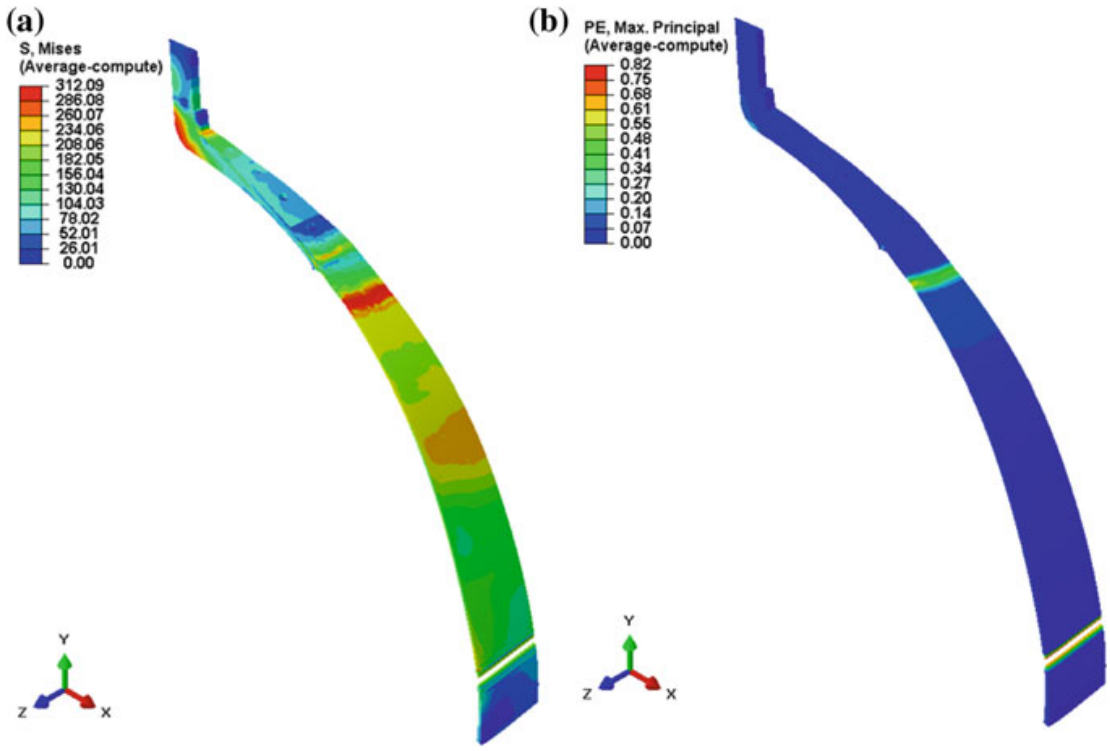


Fig. 90.7 Distribution of Von Mises stress **a** and plastic strain **b** at first failure stage in liner

equator section. Failure is observed in terms of crack appearing in liner and complete rupture is observed as equator region gets completely detached from the liner section at Von Mises stress of 312 MPa (refer Fig. 90.7a) and plastic strain of 0.82% (refer Fig. 90.7b). From the analysis, it was also required to know, if the failure is occurring in weld zones of liner as weld specimen has lower strength (refer Table 90.2). Observation showed that second failure of liner in terms of crack propagated at the transition region in dome section at Von Mises stress of 358 MPa (refer Fig. 90.8a) and plastic strain of 0.99% (refer Fig. 90.9a). Also Figs. 90.8c, d and 90.9c, d shows the enlarged view of the failure regions with complete separation of ruptured parts. It is also noted from Figs. 90.8b and 90.9b the Von Mises

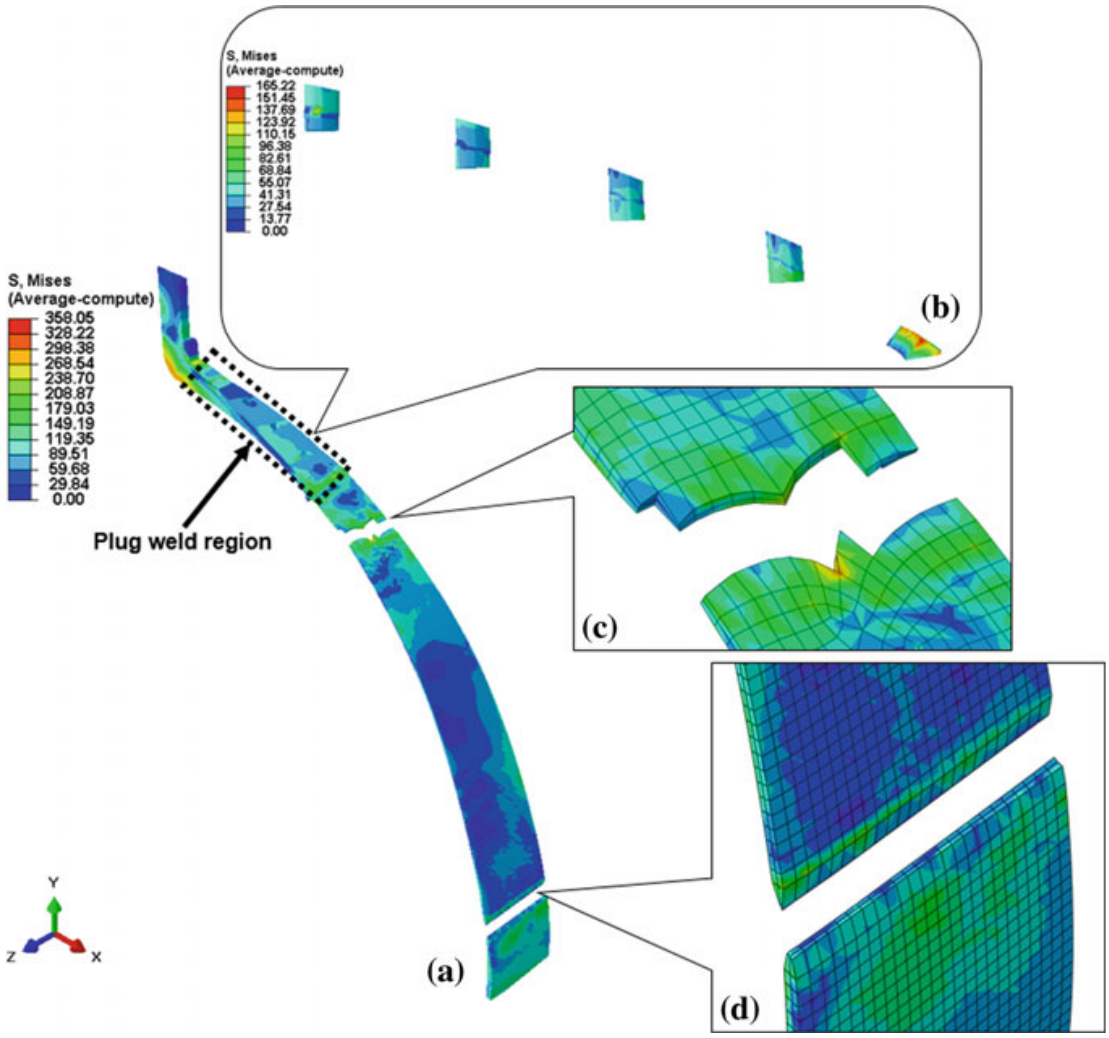


Fig. 90.8 Distribution of Von Mises stress in liner **a**, plug and fillet weld **b**, ruptured locations near dome **c** and equator **d** regions in liner

stress (165 MPa), plastic strain (0.056%) at the plug and fillet weld at the time of liner failure proved that it was intact and these welds can successfully take the impact of burst load and provide effective welding medium to join forged and formed part.

At debond regions, the internal pressure load applied on liner will not be effectively transferred to composite overwrap. This occurs because of improper bonding and insufficient bond pressure between liner-composite overwraps. Hence

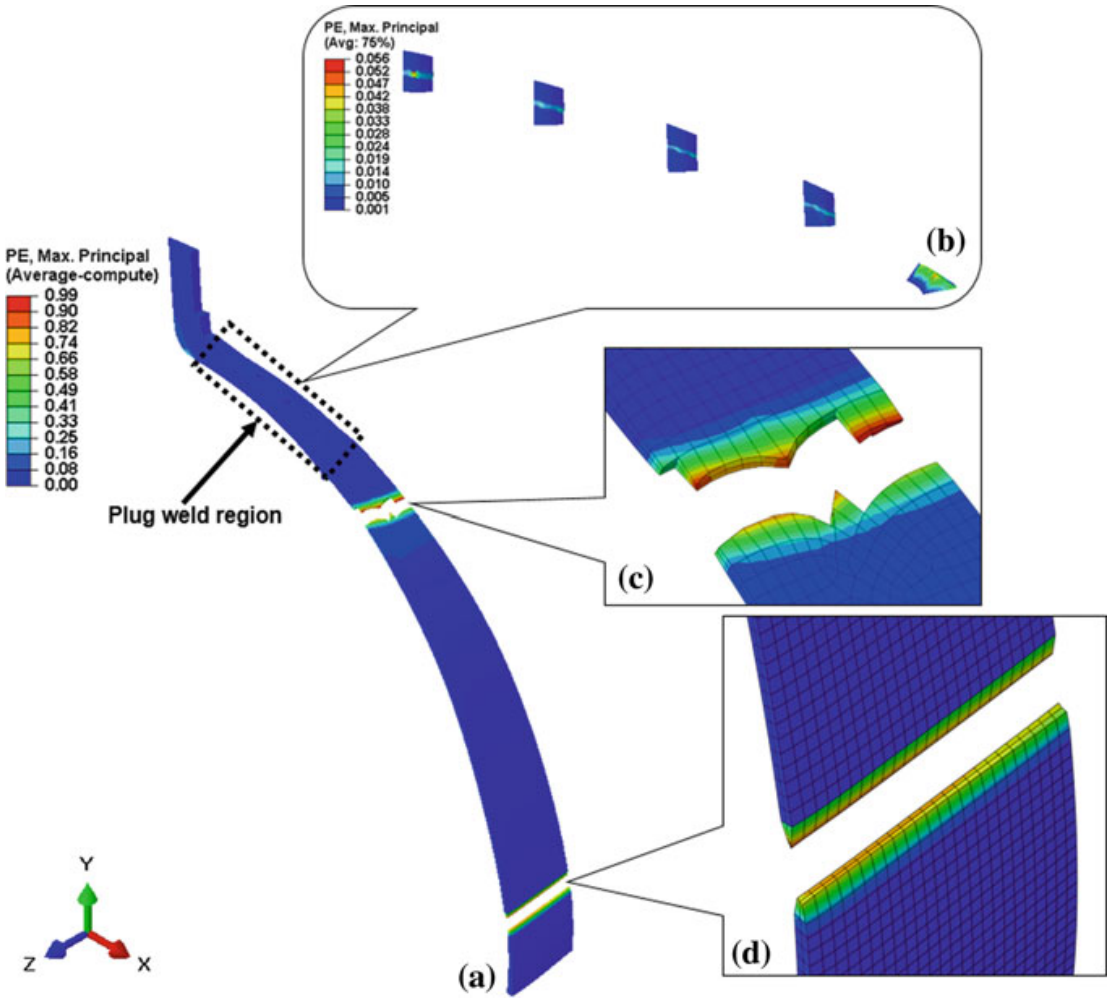


Fig. 90.9 Distribution of plastic strain in liner **a**, plug and fillet weld **b**, ruptured locations near dome **c** and equator **d** regions in liner

a void created in-between will allow liner to consume pressure load more than its capability, which will allow it to deform excessively as it cannot effectively transfer the load to composite overwrap. Peak Von Mises stress of 293 MPa (refer Fig. 90.10) is seen at the debond locations. The obtained stress of 293 MPa is 3 MPa more than the ultimate tensile strength of AA 6061-T6 proving that liner will fail in presence of debond. Graphical representation of Von Mises stress along the axial length of liner is shown in Fig. 90.11. Figure 90.12 shows the maximum deformation (7.93 mm) in liner at debond location. This plastic deformation causes

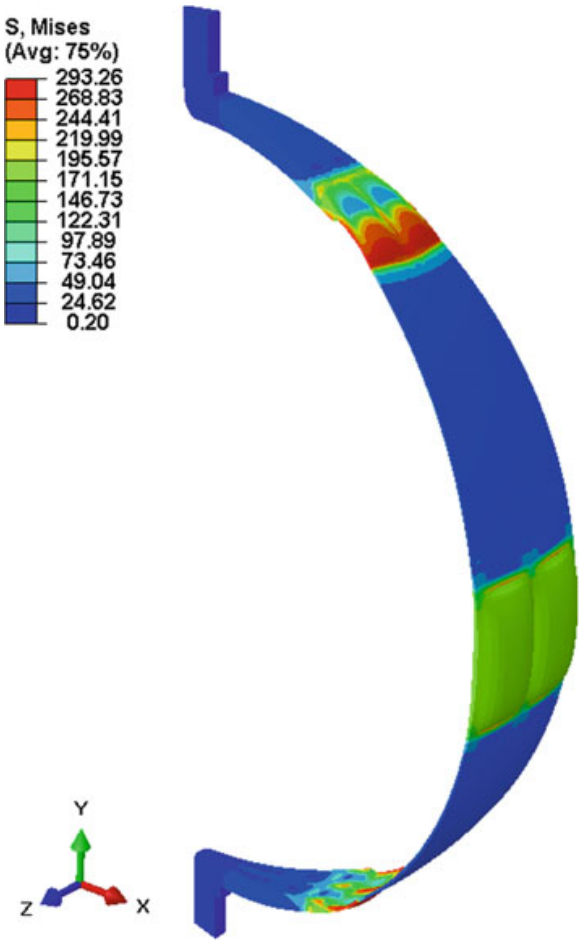


Fig. 90.10 Von Mises stress distribution in the debond model

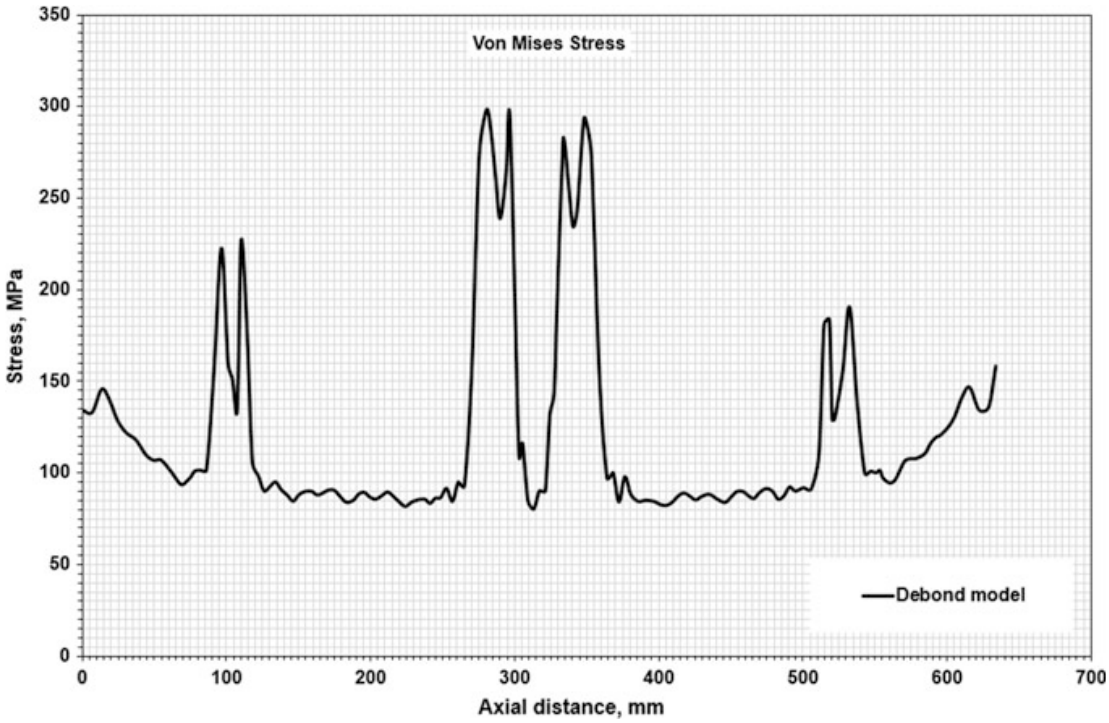


Fig. 90.11 Graphical representation of Von Mises stress distribution in the debond model

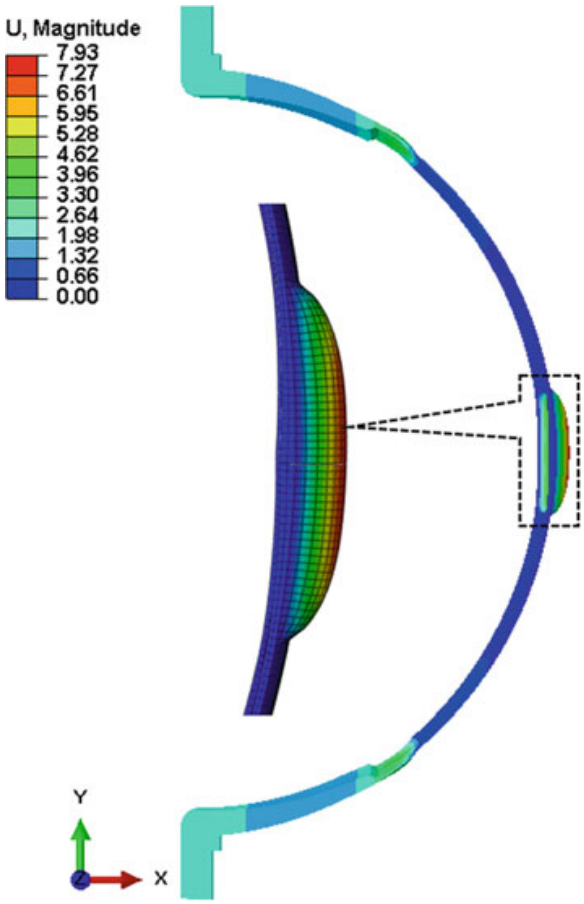


Fig. 90.12 Displacement plot in the debond model

additional problem, as during consecutive pressurization-depressurization cycles, it may cause liner to buckle; hence, bringing down the efficiency of liner and possible failure during operation. The eigenvalue linear buckling analysis is conducted to observe the buckling of liner (plastic condition) when a compressive load of 1 MPa is assigned to obtain load at which the liner buckles. During the depressurization phase, if the external load on the liner exceeds the critical buckling load of F_{cr} = 5.04 MPa the liner will buckle as seen in Fig. 90.13.

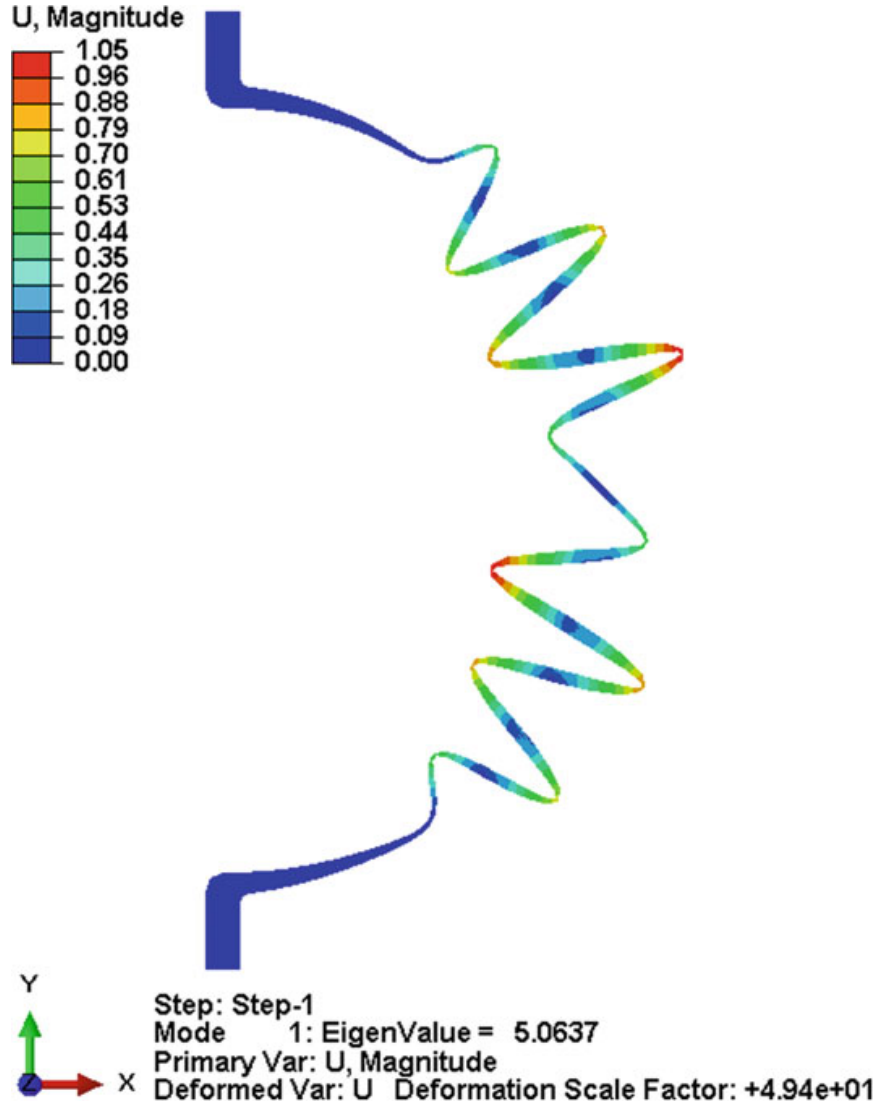


Fig. 90.13 Buckling of liner during depressurization

90.5 Conclusions

The three types of failures occurring in liner during operation are analyzed and the following conclusions are drawn:

- The presence of thickness variation in liner forms a transition region. Burst pressure load of 4.32 MPa, applied on liner showed that failure of liner occurred at these transition regions, i.e. equator region (312 MPa Von Mises stress and 0.82% plastic strain) and dome region (358 MPa Von Mises stress and 0.99% plastic strain). Further, results revealed that plug and fillet weld regions are safe from failure (165 MPa Von Mises stress and 0.056% plastic strain) due to applied burst pressure load.

- The presence of debond region had a major effect on liner as the observed Von Mises stress (293 MPa) exceed the ultimate tensile strength (290 MPa) of AA 6061-T6 proving that liner fails in presence of debond region.
- The predicted eigenvalue buckling load shows if the external pressure acting on the outer surface of liner exceeds of 5.04 MPa (depressurization phase) then the liner will buckle.

References

1. McLaughlan, P.B., Grimes-Ledesma, K.: Composite overwrapped pressure vessels. Natl. Aeronaut. Space Adm. 1–20 (2011)
2. Kawahara, G., McCleskey, S.: Titanium lined carbon composite overwrapped pressure vessel. In: 32 AIAA/ASME/SAE/ASEE Joint Propulsion Conference. Lake Buena Vista, pp. 3–12 (1996)
3. Lark, R.F.: Recent advances in lightweight filament wound composite pressure vessel technology. Natl. Aeronaut. Space Adm. 1–33 (1977)
4. Lee, H.-S., Yoon, J.-H., Park J.-S., Yi, Y.-M.: A study on failure characteristic of spherical pressure vessel. J. Mater. Process. Technol. **164**–**165**, 882–888 (2005)
5. Varun Chandran, B.V., Tiwari, S.B., Suresh, R, Krishnadasan, C.K., Sivasubramonian, B., Anoop Kumar, S.: Design and analysis of composite overwrapped pressure vessel. In: Proceedings of International Conference on Materials for the Future—Innovative Materials, Processes, Products and Applications, 109–114 (2013)
6. Almeida, J.H.S. Jr., Faria, H., Marques, A.T., Amico, S.C.: Load sharing ability of the liner in type III composite pressure vessels under internal pressure. J. Reinf. Plast. Compos. 2274–2286 (2014)
7. Son, D.-S., Chang, S.-H.: Evaluation of modeling techniques for a type III hydrogen pressure vessel (70 MPa) made of an aluminum liner and a thick carbon/epoxy composite for fuel cell vehicles. Int. J. Hydrogen Energy **37**, 2353–2369 (2012)
8. Liu, P.F., Xing, L.J., Zheng, J.Y.: Failure analysis of carbon fiber/epoxy composite cylindrical laminates using explicit finite element method. Compos. B **56**, 54–61 (2014)
9. Gray, D.L., Moser, D.J.: Finite element analysis of a composite overwrapped pressure vessel. In: 40th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit, pp. 1–15 (2004)
10. Akula, V.M., Shubert, M.: Analysis of debonding of filament wound composite pressure vessels. In: American Society for Composites—28th Annual Technical Conference, pp 341–354 (2013)
11. Kaufmann, C., Williams, K.: Optimization of a numerical simulation involving the impact of an AP-T C44 12.7 mm projectile on a semi-infinite monolithic 6061-T6 aluminium target, pp. 1–17. Defence Research and Development Canada (2004)
12. Nalajam Pavan Kumar, Arungalai Vendan, S., Siva Shanmugam, N.: Investigations on the parametric effects of cold metal transfer process on the microstructural aspects in AA6061. J. Alloys. Compd. 255–264 (2016)
13. Liang, Y., Hu, S., Shen, J., Zhang, H., Wang, J.: Geometrical and microstructural characteristics of the TIG-CMT hybrid welding in 6061 aluminum alloy cladding. J. Mater. Process. Technol. 18–30 (2017)
14. Lesuer, D.R., Kay, G.J., LeBlanc, M.M.: Modelling large-strain, high-rate deformation in metals. In: Third Biennial Tri-Laboratory Engineering Conference Modelling and Simulation, Pleasanton, pp. 1–12 (1999)

Index

A
AA6061, 95, 97, 98, 102, 103, 107
Al6061 alloy, 485, 487, 489, 494
Abaqus, 819, 821, 823, 825, 827, 851, 854, 855
ABAQUS/Explicit, 13–16
Adaptive neuro-fuzzy inference system (ANFIS), 841–843, 846, 847, 849
Additive manufacturing, 403, 405
Analytic hierarchy process (AHP), 205–208, 210–212
AISI 201-grade stainless steel, 215
Al–Cu alloy, 461, 462, 464, 466
Al–Cu cast alloy, 469, 471, 481
Alloy steel, 317–320, 322, 324–326
Al/SiC, 743–745, 747–752
Aluminium foams, 439, 440
Aluminum alloy scrap, 733, 734, 738, 740
ANOVA, 569, 579, 584, 593, 596, 600
Antenna reflector, 657–660, 664, 666, 693, 694, 700, 701
Artificial intelligence, 426, 427
A-TIG welding, 180, 189, 246, 247, 250, 252, 255
Austenitic stainless steel, 1031–1033
Autoclave, 657–660, 662, 664, 666, 693–702
Autonomous manufacturing systems, 643

B
Bagasse fly-ash, 743–747, 752
Blood warmer, 967–971
Bonobo optimizer, 153, 155, 158, 159
Boundary condition, 1012–1020
Brake disc, 973–975, 977–979, 981, 982
Brown rice, 1059–1063, 1065, 1066, 1071
Buckling analysis, 987, 991
Buckling load, 985

C
2.25Cr–1Mo steel, 245–248, 252, 254, 255
Carbon fibre-reinforced polymer (CFRP), 657–661, 664, 666
Carbon nanorods, 1045, 1047
Casting, 440
Chip morphology, 318, 319, 324, 326
Chip thickness, 531, 532, 536, 541
Circularity, 615, 617, 621, 622, 624, 626
Circular saw, 520–522, 524, 525, 527
CNC machining, 514
CNC turning, 338, 339, 344
Cohesive zone method, 832
Cold gas dynamic spray, 393–395, 397, 400
Composite lattice structures, 986
Composite materials, 947, 952
Composites, 773, 774, 778, 780, 781
Computed tomography, 603–605, 606, 613
Computer vision, 629, 630
Conical space adapter, 985
Cooling rate, 165, 167, 174–177
Cricket algorithm, 153, 155, 159
Critical velocity, 393, 400
Crystal plasticity, 349–352, 357, 360
Cuckoo search algorithm, 719, 725
Cutting-edge radius, 309–311, 314, 315
Cutting-edge wear, 629, 630
Cutting forces, 531, 532, 536, 537, 539–541
Cutting speed, 545–550, 554
Cyber-physical systems, 643

© Springer Nature Singapore Pte Ltd. 2019

R. G. Narayanan et al. (eds.), *Advances in Computational Methods in Manufacturing*, Lecture Notes on Multidisciplinary Industrial Engineering,
https://doi.org/10.1007/978-981-32-9072-3

Cylindricity, 615, 617, 621–624, 627

D
3D printing, 416, 418, 420
Damage mechanism, 1012, 1013
Decision analysis, 205, 212
Deep cold rolling, 95–97
Delamination, 829–833, 836, 837
Die casting, 447, 448, 451, 455, 457
Die casting machine, 425, 426, 428, 430, 431, 433–436
Dimensional accuracy, 693–695, 702
Dissimilar joint, 165–167, 170, 172
Dissimilar welding, 233, 236
Distortion, 657–661, 663–666
Drilling, 829, 830, 832, 833, 835–837

E
EDX analysis, 165, 167
Elastic contact, 819, 820, 824
Electrical discharge machining (EDM), 329–331, 334, 335
Electric field, 557, 559, 562, 564, 567
Electromagnetic crimping (EMC), 271
Electron beam welding, 121, 123, 125, 131, 153–156, 162, 840
Elongation, 133, 134, 136, 137, 140, 143
EN-31, 498, 500
EN31 Tool, 338–341, 343–345, 348
End forming, 65–68, 73
Energy-dispersive analysis of X-rays (EDAX), 672, 678
Erosion, 851–854, 856–859
Extrusion, 3, 4, 6–8, 10, 11

F
Factory modelling and simulation, 643
Femur bone, 371, 372, 376
Fibre-reinforced plastics, 830, 831
Fibre-reinforced polymers (FRP), 773
Fibrizer hammer, 933–943, 945
Finite element (FE), 805–808, 810, 815, 1073
Finite element analysis, 657, 819, 821, 827, 851, 853, 859, 890, 956, 957, 975, 978, 985–987, 989, 1059, 1066, 1067
Finite element method (FEM), 383, 386, 861
Finite element methodology, 393
Finite element model, 231, 233, 235
Finite element modeling, 179, 315, 350, 360, 403
Finite element simulation, 65, 283
First ply failure, 947–950, 953, 954
Fissure detection, 1061
Flexural strength, 773, 774, 777–779, 781
Flow rate and heat supply, 967, 969–972
Fluidity, 469–471, 479–481
Formability, 37–39, 43, 45, 49, 51, 47
Forming, 3, 5, 6, 8, 38, 43, 45
Forming limit diagram, 39
Fourier number, 121, 123–125, 127–129
Fractography, 417, 418, 419, 421, 424
Fracture strength, 793–795, 800–802
Free vibration, 1055
Friction stir back extrusion process, 734, 735, 738, 739
Friction stir welding, 14, 15, 133, 134, 806
Fused deposition modeling (FDM), 415, 417
Fuzzy inference system, 841, 842
Fuzzy logic, 447–450, 455

G
Gas metal arc welding (GMAW), 215–218, 225, 228
Gaussian distribution, 805, 808
General-purpose GPU (GPGPU), 891
Genetic algorithm, 215, 909
Geometrical analysis, 603
Geometrical tolerances, 621
Geometric distribution, 1023, 1025, 1026, 1029
Glass, 557–559, 562, 564, 567
Glass fibre-reinforced polymer (GFRP), 1012–1013, 1020
Graphene, 793–796, 799–802
Grey fuzzy, 569, 571–573, 575, 576, 578, 581–584
Grey relational analysis, 134, 137
Grid partitioning, 842, 849
Groove geometrical parameters, 271
Guides, 519–521, 526–530
Gyroid, 955–963

H
H-13 tool steel die and chamber, 734, 735
Harmonic analysis, 940, 942
Heat source model, 262
Heat transfer coefficient, 877, 878, 882, 886, 887
Heat treatment, 85–87, 90, 92
Hexagonal bar, 733–737, 739
Hot spot, 459–461, 466
Hot tear, 459–461, 464, 465
Human machine cooperative system, 425, 426, 430, 436
Hybrid composites, 784, 785
Hybrid multi-objective algorithm, 899, 901, 907
Hybrid welding, 259, 260, 262, 264, 267

Hydraulic autofrettage, 998
Hydroxyapatite, 669–671, 682

I
Image processing, 629–632, 637–639
Inconel 718, 231–235, 237, 239–241
Inconel alloy, 109–117
Indirect processing, 439, 440, 444
In-plane plane-stretching (IPPS), 25, 26, 32, 34

J
JIS 3302 grade steel, 85, 86, 89, 92

L
316L austenitic stainless steel, 193, 194, 198, 202
Laminates, 948–950
Laser additive manufacturing, 381–385, 387, 388, 390
Laser welding, 165–168, 172, 174
Limit dome height, 37, 41, 43, 45
Linear shrinkage, 447, 449–451, 454, 455, 457
Linear solver, 889
Liner failure, 1073, 1082
Local search, 899–903, 907
Local search heuristics, 909, 911, 912, 915, 916, 918
Low velocity impact (LVI), 1012, 1013
LS-DYNA^{TM}, 271, 272, 282

M
Machinability, 317–320, 322, 324, 326
Machine intelligent quotient (MIQ), 425, 426, 430, 432, 433, 435, 436
Machining, 309–312, 315, 363
Macro-hardness, 133, 134, 136, 137, 140, 143
Magnetic field-assisted powder-mixed (EDMMFAPM-EDM), 485–487, 489, 494
Manufacturing, 298
Marangoni number, 123–125, 127–129
Material selection, 755, 756, 762, 764
Maxwell's equation, 272, 273
Mechanical properties, 25–30, 37, 38, 40, 41, 783–785, 861, 862, 870, 874
Melt-pool dimensions, 381–386, 389
Metallic powder, 703, 705, 706, 708, 713, 715, 716
Micro-EDM, 569–571, 574, 575, 584
Microholes, 603–606, 608, 609, 613, 614
Micromachining, 349, 354, 357, 359
Microplasma arcwelding (MPAW), 587–590, 592, 593, 595, 598, 600
Microstructure, 670, 671
Microwave (MW), 706, 709
Microwave drilling, 557–561, 567
MMCs, 744, 745
Modeling, 300, 547, 554
Molecular dynamics, 349–351, 359, 360
Molecular dynamics simulation, 794, 795, 797, 798
Multi-criteria decision-making (MCDM), 334
Multiphysics, 967, 969, 972
Mutation, 899, 901–903, 907

N
Nanoindentation, 862–864, 868, 869, 871, 874
Neural network, 285
Nonlocal elasticity, 1045–1048, 1050, 1051, 1053, 1055
Non-metallic inclusions, 317, 320
NSGA-II, 899–903, 905, 906
Numerical modeling, 281
Numerical simulation, 247, 393

O
Optimization, 76, 77, 153–162, 297, 299, 302, 307, 330, 334, 335, 366, 544, 545, 554, 588, 592, 593, 598, 600
Orthogonal array, 448, 450, 455
Orthopedic surgery, 371, 372, 376
Osteonecrosis, 371, 376
Overall utility, 587, 594–598, 600

P
Parallel computing, 889–891
Perpendicularity EDM, 615, 617, 621–623, 626
Piston casting, 426
Plasma arc preheating, 806
Polymer composites, 720
Polypropylene, 784–786, 791
Porosity control, 439–444
Porous scaffolds, 956, 963
Pouring temperature, 469–471, 474–481
Preparation index (PI), 933–935, 942
Pressure drop, 877, 879, 882, 883, 887
Process capability index, 1023–1026, 1028, 1029
Processing maps, 109, 110, 114
Process parameter, 329–331, 335
Product design, 972
Pulsed thermal modelling, 167, 172

Q
Quality characteristics, 1023, 1029

R
Residual stress, 95–97, 99, 101, 103, 105–107, 338–341, 344, 345, 997–1008, 1031, 1032, 1035, 1038–1041, 1043
Rosenthal model, 123, 125, 127
Rotational autofrettage, 997–999, 1001, 1005–1008
Rotational speed, 997, 999, 1001–1008
Rough surface, 819–827, 851, 853, 854, 856, 858, 859
Run-out, 531–539, 541

S
Sampling properties, 1027
Scallop height estimation, 516, 517
Self-pierce riveting, 283, 284
Sequential drilling, 371–373, 375, 376
Serpentine microchannels, 878–883, 886, 887
Serrated cutters, 531–533, 536–541
Sheet metal forming, 26, 86
SiC, 743–747, 750–752
Silicon, 861–863, 865–867, 871, 874
Simulation, 13–16, 18–22, 459–462, 465, 466, 469–472, 474, 479, 481
Size effect, 3, 4
Sparse matrix, 890–894
Spherical pressure vessel, 1073
Springback, 75–77, 80, 82, 83, 85, 86, 89, 92
SS 316L, 231, 233, 235, 239
Stainless steel, 259, 260, 263, 266–268
Statistical analysis, 364
Steel tubes, 205
Stereolithography, 403–406, 412
Stir casting, 743, 744
Stochasticity, 1047, 1053, 1056
Stress, 557–559, 563–565, 567
Stress analysis, 974
Submerged arc welding, 1031–1033
Surface roughness, 447, 449–451, 454, 455, 457, 570, 578, 582–584
Surface topography, 511, 517
Swarm intelligence, 153, 162
Synthetic fiber, 784

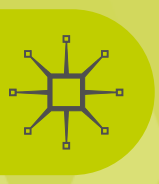
T
Taguchi, 75–80, 83
Tailor welded blanks, 13, 49
Teaching–learning-based optimization (TLBO), 485, 487, 488, 493, 494
Temperature, 283, 285–288, 290, 292–294
Tensile strength, 133, 134, 136, 137, 140, 142, 143, 774, 777, 781
Tensile test, 179–190
Tensile testing, 13, 14, 16
Thermal cycles, 381, 386, 389
Thickness gradient necking criterion (TGNC), 25, 33, 34
Three-point breaking test, 1059, 1061, 1064–1067, 1070, 1071
Ti6Al4V, 309–311, 315
Ti–6Al–4V alloy, 381–384, 386, 388, 389
TIG welding, 194–196, 198, 199, 202
Tissue engineering, 956
Tool wear, 629–631, 639
Tool wear rate, 485, 487, 490, 491
Tribological properties, 723, 730
Tube joining, 71, 72

U
U-bending, 75, 76, 80, 83
Ultra-short pulse laser, 603, 604, 613, 614
Uncertainty, 756, 757

V
Vacancy concentration, 795, 796, 802
Vacuum-assisted resin transfer moulding (VARTM), 693–695, 697, 699–702
V-bending, 85, 87
Vehicle routing problem, 909, 910
Velocity profile, 877, 885, 887

W
Weld bead, 215–219, 221, 225, 228
Welding performance, 194, 196
Welding process, 205, 207–212
Welding residual stress, 246, 247, 250, 253, 255
Weld line shift, 49, 51, 55–61
Wirecut electric discharge machine (WCEDM), 497–500, 543–545, 554
Wire wear ratio, 544–546, 549, 551
Wonji Shoa Sugar Factory, 933–935, 945
Workability, 109, 115, 117

X
X-ray diffraction (XRD), 669, 671, 673–675



Edited by
Babu George
Justin Paul

Digital Transformation in Business and Society

Theory and Cases

palgrave
macmillan

Digital Transformation in Business and Society

"This is a great book full of information for those who want to better understand the digital revolution and how it is changing the whole of society, and business in particular. This book is a good guide to see this reality in a completely different way, through the scope of technology, and to better understand the exponential future we are living in."

—Dr. Besa Shahini, *Professor of Economics and Finance, University of Tirana, Albania*

"The book offers thoughtful, comprehensive insights into the far-reaching capacity of the digital footprint to reconfigure and remap every aspect of business and societal behavior. The authors have crafted a trailblazing compendium of digital transformation thought and action related to micro- and macro aspects of business and society. The book has the potential to become the defining work in this body of literature."

—Dr. Alvin Williams, *Professor and Chair, Department of Marketing, University of South Alabama, USA*

"One of the best books I have had the opportunity to read, I strongly recommend that you take it into account for your graduate programs."

—Dr. Sergio Garcia Agreda, *Director of the Master's Program in Digital Business, Universidad Privada Boliviana, Bolivia*

"This book is a timely text that will inform readers on a wide range of business and societal applications in the area. This text will help organizational psychologists understand the human factor in business decision-making and execution to then fully grasp the potential for digital transformation both in Business 4.0 and Society 5.0."

—Dr. Mervyn Jackson, *Head, Discipline of Psychology, RMIT University, Australia*

Babu George • Justin Paul
Editors

Digital Transformation in Business and Society

Theory and Cases

palgrave
macmillan

Editors
Babu George
Fort Hays State University
Hays, KS, USA

Justin Paul
University of Puerto Rico
San Juan, Puerto Rico

ISBN 978-3-030-08276-5 ISBN 978-3-030-08277-2 (eBook)
https://doi.org/10.1007/978-3-030-08277-2

© The Editor(s) (if applicable) and The Author(s), under exclusive licence to Springer Nature Switzerland AG 2020
This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

1 **Introduction** 1
Babu George

2 **Human Flourishing in Smart Cities** 13
Maria Lai-Ling Lam and Kei-Wing Wong

3 **On How Big Data Revolutionizes Knowledge Management** 39
Asha Thomas and Meenu Chopra

4 **Social Media in Marketing Communications: A Synthesis of Successful Strategies for the Digital Generation** 61
Ogechi Adeola, Robert Ebo Hinson, and Olaniyi Evans

5 **Social Media Technologies and Export Marketing** 83
Frederick Avornyo, Mahmoud Abdulai Mahmoud, Ogechi Adeola, Robert Ebo Hinson, and Richard Boateng

6 **Data Analytics and Predictive Analytics: How Technology Fits into the Equation** 103
Brian J. Galli and Gabrielle Muniz

7 **Effective Engagement of Digital Natives in the Ever-Transforming Digital World** 113
Anju Varghese Philip and Zakkariya K. A.

8 **Robotics, Artificial Intelligence, and the Evolving Nature of Work** 127
Craig Webster and Stanislav Ivanov

9 **The Dark Side of Big Data: Personal Privacy, Data Security, and Price Discrimination** 145
Yang Liu and Connor Greene

10 **Digital Finance for Financial Inclusion and Inclusive Growth** 155
Md. Nur Alam Siddik and Sajal Kabiraj

11 **Understanding Consumer Behavior in Technology-Mediated Spaces** 169
İlayda İpek

12 **An Evaluation of the National Open Government Data (OGD) Portal of the United Arab Emirates** 191
Stuti Saxena

13 **Artificial Intelligence and Robotics Technology in the Hospitality Industry: Current Applications and Future Trends** 211
Li Yang, Tony L. Henthorne, and Babu George

14 **Risk Management in the Digital Era: The Case of Nigerian Banks** 229
Tankiso Moloi and Oluwamayowa Olalekan Iredele

15 **Digital Technology to Enhance Project Leadership Practice: The Case of Civil Construction** 247
John Ekechukwu and Thorsten Lammers

16 The Making of Data-Driven Sustainable Smart City Communities in Holiday Destinations 273
Sudipta Kiran Sarkar, Michalis Toanoglou, and Babu George

Index 297

Notes on Contributors

Ogechi Adeola is a fellow/senior lecturer in the Marketing/Operations Department of Lagos Business School Pan-Atlantic University, Nigeria.

Frederick Avornyo is a doctoral student in the Department of Marketing and Entrepreneurship, University of Ghana Business School, Ghana.

Richard Boateng is an associate professor in the Department of Operations and Management Information Systems at the University of Ghana Business School, Ghana.

Meenu Chopra is an academician and researcher by profession and spirit. She is pursuing a doctorate program as a full-time research scholar from the prestigious Delhi Technological University (DTU) (formerly Delhi College of Engineering) under the DTU Fellowship program. She has worked as an assistant professor in the Department of Management Studies, Deen Dayal Upadhyaya College, University of Delhi, teaching BBS/BMS students, which is the only undergraduate management program run by the University. Her areas of interest include knowledge management, human resource management, and advertising and branding.

John Ekechukwu is a civil engineer leading transport infrastructure design at Turnbull Engineering, Sydney, NSW, Australia. He has recently organized and facilitated the Turnbull Engineering Transport Infrastructure Design School of Excellence in partnership with the University of New South Wales. Engineering leadership and development through digital transformation is one of his primary areas of interest.

Olaniyi Evans is a lecturer in the School of Management and Social Sciences of Pan-Atlantic University, Lagos.

Brian J. Galli works as Assistant Professor of Management Engineering at Long Island University—Post (LIU Post). He also owns Apex Strategies, Ltd, a company that specializes in continuous improvement consulting and training initiatives. The author's major field of study is continuous improvement and project management integration. He has over 11 years of experience in applying industrial engineering and continuous improvement tools and concepts in a wide variety of arenas, including health care, manufacturing, transactional, and service environments.

Babu George is an associate professor in the Robbins College of Business and Entrepreneurship, Fort Hays State University, KS, USA. His research examines various dimensions of technology as it interacts with business and society. Of these, human behavior in technology-mediated spaces constitutes one of his key focus areas.

Connor Greene is an undergraduate student in the Finance and Economics Department of Southern New Hampshire University, NH, USA.

Tony L. Henthorne is a professor and Associate Dean of Graduate Programs at the University of Nevada, Las Vegas. Technological innovations in the service industry and its marketing implications constitute one of his research streams.

Robert Ebo Hinson is a professor in the Department of Marketing and Entrepreneurship, University of Ghana Business School, Ghana.

İlayda İpek is a research assistant at the Faculty of Business, Dokuz Eylül University, Turkey. Her research interests primarily concentrate on international marketing, brand management, and consumer behavior. Her articles have appeared in the *International Business Review*, *Journal of Consumer Behavior*, and *Journal of Retailing and Consumer Services.*

Oluwamayowa Olalekan Iredele holds a PhD in Accounting. He is a postdoctoral research fellow in the School of Accounting, University of Johannesburg. His interest is in environmental management accounting and reporting.

Stanislav Ivanov is a professor and Vice Rector (Research) at Varna University of Management, Bulgaria. Ivanov is the Founder and Editor-in-Chief of the *European Journal of Tourism Research* and serves in the editorial boards of more than 30 other journals. His research interests

include robonomics, robots in tourism/hospitality, revenue management, destination marketing, tourism and economic growth, political issues in tourism, and so on.

Sajal Kabiraj is a full professor at the International Business College, Dongbei University of Finance and Economics, Dalian, China. He has taught at postgraduate level in universities in Sweden, Austria, Malaysia, Canada, UAE, China and India. Information technology and digital inclusion is one of his primary areas of interest.

Maria Lai-Ling Lam is Professor of Business at Calvin College, Grand Rapids, MI, USA. She has been a senior consultant of AJ-Great Limited for several years. She has taught at two Christian universities in the United States for 14 years and several universities in Hong Kong and China for seven years She holds a PhD in marketing and organization behavior from the George Washington University, an MA degree in Religion Studies, and an MBA and a BBA from the Chinese University of Hong Kong. She has authored one book and 99 peer-reviewed articles and book chapters. Her areas of research interest are managing diversity and inclusion, yoga meditation, management education, leadership development, and corporate social responsibility.

Thorsten Lammers is a senior lecturer at the School of Information, Systems and Modelling, Faculty of Engineering and IT at the University of Technology Sydney, NSW, Australia. Informed by a previous career in strategy and technology management consulting, his main research interest lies in the impact of digital technologies on the future of work.

Yang Liu is an assistant professor in the Business Administration Department, Fitchburg State University, MA, USA. She holds a PhD in International Business, major in Marketing, from Southern New Hampshire University, NH, USA, and she also has a master's degree in Economics and Business from Sweden and a bachelor's degree in Accounting, major in Enterprise Financial Management, from China.

Mahmoud Abdulai Mahmoud is a senior lecturer in the Department of Marketing and Entrepreneurship, University of Ghana Business School, Ghana.

Tankiso Moloi is Professor of Accountancy at the School of Accounting, University of Johannesburg. He holds a PhD in Accounting and Finance as well as three master's degrees in Accounting, Financial Management, and International Relations. His interest is corporate governance, especially risk management.

Gabrielle Muniz is a freelance writer and editor. She recently graduated from Long Island University—Post (LIU Post), as she has lived in Long Island all her life. Gabrielle has extensive editing experience in the field of management engineering, with topics such as continuous improvement, project management, and change management. Aside from her freelance work, she has been writing many articles on topics that vary from dessert recipes to movie reviews for two online magazines: *Witty Serendipity* and *Her Track*. Overall, Gabrielle hopes to gain a fulfilling career in the writing industry.

Justin Paul is Professor of Marketing and International Business at Rollins College, FL, USA, and the Graduate School of Business, University of Puerto Rico. He is known as an author/co-author of text books such as *Business Environment* (4th ed) and *International Marketing, Export-Import Management* (2nd edition). Two books turned out to be best-sellers in Asia. He has served as a full-time faculty member with premier institutions such as the University of Washington, Nagoya University of Commerce & Business, Japan and Indian Institute of Management. Paul serves as a senior/guest/associate editor with the *International Business Review*, *Journal of Retailing & Consumer Services*, *Small Bus Economics*, *European Journal of International Management*, *Journal of Promotion Management*, and *International Journal of Emerging Markets*. In addition, he has taught full courses at Aarhus University, Denmark; Grenoble Eco le de Management, France; Universite De Versailles, France; ISM University, Lithuania; Warsaw School of Economics, Poland, and has conducted research paper development workshops in countries such as Austria, the United States, Croatia, and China. Paul has been a program director for training diplomats from different countries and served as a consultant to the Government of India. He has been an invited speaker at several institutions such as the University of Chicago, Fudan & UIBE-China, Barcelona, and Madrid, and has authored over 50 research papers and best-selling case studies with Ivey and Harvard. His articles have been downloaded over 300,000 times. He introduced the Conservative, Predictable and Pacemaker (CPP) model for internationalization and massage model/measure for brand management.

Anju Varghese Philip is a full-time doctoral student at the School of Management Studies, Cochin University of Science and Technology, India. Her research interests include human-computer interaction, specifically in the area of need satisfaction through technology use.

Sudipta Kiran Sarkar is an assistant professor at Sol International School, Woosong University, South Korea. The role of social media technologies in business is one of his major areas of research.

Stuti Saxena is a PhD student at the Central University of Haryana. Her research explores the use of information technology by governmental organizations.

Md. Nur Alam Siddik is an associate professor in the Department of Finance and Banking, Begum Rokeya University, Rangpur, Bangladesh. He holds a PhD degree in Finance from Dongbei University of Finance and Economics, Dalian, China. Digital finance and financial technologies is one of his primary areas of interest.

Asha Thomas is an assistant professor at Jagan Institute of Management Studies (JIMS), New Delhi, India. Her areas of research interest include knowledge management, organizational behavior, and marketing. She has about nine years of experience in teaching, as well as over three years of experience in the IT and telecom industry. She is pursuing a doctorate program as a part-time research scholar from the prestigious Delhi Technological University.

Michalis Toanoglou is the department chair and an assistant professor in International Department of Hospitality Management (SIHOM) of Sol International School (SIS) in Woosong University in South Korea. He has several years of experience as an executive in the hospitality industry in Greece.

Craig Webster is an associate professor in the Department of Management at Ball State University, Muncie, IN, USA. His research interests include the political economy of tourism, public opinion analysis, and event management. His most recent book is *Robots, Artificial Intelligence and Service Automation in Travel, Tourism, and Hospitality*, co-edited with Stanislav Ivanov. He teaches courses in hospitality management at Ball State University's Miller College of Business.

Kei-Wing Wong is a graduate student of Architecture, University of Michigan, Ann Arbor, MI, USA. She also has a Bachelor of Science degree from Calvin College, Grand Rapids, MI, USA. Her major was Environmental Science (Chemistry). Her minor was Architecture and Geography. She has worked for different sustainable cities projects. Her passion is to develop a universal and affordable design for people to flourish in hospitals and to cope with the threat of climate change.

Li Yang recently graduated from the master's program at W.H. Harrah College of Hotel Administration, University of Nevada, Las Vegas. His graduate thesis investigated how artificial intelligence might revolutionize the future of hospitality.

Zakkariya K. A. is Professor of Marketing and Human Resource Management at the School of Management Studies, Cochin University of Science and Technology, India, and is also the Director of DDU KAUSHAL Kendra, the Centre for Vocational Education in the university. With more than 20 years of teaching and research experience, he has been a member of various academic and administrative bodies of different universities in the country. He has more than 60 research publications in journals and books of national and international repute and is reviewer to many journals, including *Gender in Management* and *Journal of Management Development*.

List of Figures

Fig. 1.1	Key themes included in this volume	3
Fig. 3.1	Intersection of big data and KM in business	47
Fig. 3.2	Knowledge management—harnessing the power of big data	52
Fig. 4.1	The linkage between WOM and eWOM. Source: Authors' representation	66
Fig. 5.1	Thematic distribution of included literature	95
Fig. 5.2	Social media export marketing performance capabilities framework	96
Fig. 10.1	Impact of ATM on LOL	159
Fig. 10.2	Impact of INT on LOL	160
Fig. 10.3	Impact of MMT on LOL	160
Fig. 10.4	Step direct benefit transfer delivery model. Source: National Payments Corporation of India. Infographics by FINO PayTech	166
Fig. 15.1	Research design overview	249
Fig. 16.1	Interlinkages between sustainability and technology in the urban tourism setting	292

List of Tables

Table 8.1	Top ten job skills needed in 2020	135
Table 10.1	Variables, their proxies, legend and source	159
Table 12.1	Citizen engagement models proposed by Sieber and Johnson (2015)	196
Table 12.2	Model proposed for evaluating OGD portals (Machova et al., 2018)	197
Table 12.3	Four types of websites and their attributes	198
Table 12.4	Quality aspects and their dimensions	199
Table 12.5	Assessment of website types in terms of the attributes	199
Table 12.6	Major facilitators and hindrances in re-using datasets via the national OGD portal of UAE	202
Table 12.7	Key insights about the portal using the framework of Sorum et al. (2013)	205
Table 14.1	Current risk management approach in Nigerian banks	243
Table 15.1	Overview of literature review	250
Table 15.2	Overview of key leadership demands in construction	256
Table 15.3	Overview of challenges to meet leadership demands	264
Table 15.4	Digital technologies to assist with leadership challenges	266

CHAPTER 1

Introduction

Babu George

Introduction

Businesses are being transformed in fundamental ways as a result of the data revolution and societal drivers of change (Aral, Dellarocas, & Godes, 2013). The new era inaugurated by digital 2.0 technologies created a discontinuity and a fundamental shift in our ideas of business and society. Fundamental concepts that constituted the definitional parameters of business are undergoing a re-definition process. Technological revolutions brought to us increase in processing power, bandwidth, user-friendly interfaces with technology, the 'internet of things', to list a few benefits. Also, we live in a society that loves to consume radical change. These drivers together have transformed our situated existence in the human society (Uhl & Gollenia, 2016). For businesses, these changes mean the blurring of industry boundaries, modular business architectures, distributed innovation micro-structures, 'just-in-time' definitions of business performance, among others.

Most major businesses in our times are in the business of transforming themselves by the power of data and the technological assets that make it possible for them to use the data in the best possible ways (Matt, Hess, & Benlian, 2015). Systematic improvement in business performance is

B. George (✉)
Fort Hays State University, Hays, KS, USA
e-mail: bpgeorge@fhsu.edu

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_1

achieved by means of re-engineering every aspect of the firm, including its external relationships, with technologies that promise scope and scale enhancements and also flexibility to create new business models (Davidson, 1993). There is a widespread realization that the path to competitive advantage is through data (McAfee, Brynjolfsson, Davenport, Patil, & Barton, 2012). In this backdrop, this volume brings together different perspectives from experts on how the data revolution has shaped different regions, industries, and domains. The authors in this volume have identified and distinguished the driving forces that make radical shifts in our society and businesses possible.

The Digital Transformation of Businesses and Societies

Knowledge has gained predominantly as a factor in contemporary consumption (Antons, Joshi, & Salge, 2018). Consumers are eager to be part of the knowledge sharing process—they co-create what they consume. In many situations, the boundaries between the buyers and the sellers are now permeant, observe Ranjan and Read (2016). Also, since most of the purchases are made online, it is vital to provide these consumers sensory experiences as close to what they would get while visiting brick and mortar shops. There is huge scope for innovation in this area and digital technologies fueled by the power of big data could provide the main lever for the same (Caputo, Evangelista, Perko, & Russo, 2017).

On the one hand, we have some evidence to support the claim that the digital revolution has helped us improve not just our business enterprises but our overall standard of living as a community of people, around the world. On the other hand, constant disruptions in the technology sphere need individuals and societies to pace up relentlessly. The digital divide is stark in certain regions of the world and also in certain demographic groups. Governments and the civil society have to come up with ways and means to bring the digital have-nots up to speed. This could be achieved by means of a combination of education, training, and other incentives. The private sector's role in this regard should be examined more, too.

Digital revolution has impacted and transformed almost every realm of human affairs (Makridakis, 2017). So, almost every current topic on business and society will have a certain aspect of it. When we look at issues with a thin focus, it is natural to see certain things as outliers. However, technology myopia encountered by their leaders has caused many

businesses their very survival and we need to be mindful about unduly limiting the scope of the digital revolution (Vega-Jurado, Juliao-Esparragoza, Paternina-Arboleda, & Velez, 2015). It is impossible to make sense of our contemporary life without referencing it with technologies that laced the fabric of it (Brooke, 2016).

Will people work in the future? No, especially if we continue to define work the way we do it now (Cook, 2017). Are we ready for the new kind of work? What characteristics will define human workers and how will these be able to give them a competitive advantage over the robots? Will artificial intelligence evolve to the extent of making humans entirely replaceable in the workplace? These are but a sample from the cross section of the questions that futurists pose (Aronowitz, 1994; Collin & Young, 2000; Stiegler, 2018). These questions cannot be avoided when we discuss business implications of the fourth industrial revolution.

That said, for practical reasons, this volume needed to be delimited to certain key dimensions of it. The diagram given below is an attempt to distill the characteristic themes that the chapters contained in this volume represent. This is by no means an exhaustive or mutually exclusive list of themes. Several chapters cut across the theme boundaries—which, again, is a characteristic feature of the seamlessly networked lives of the digital natives that we are.

The key themes included in this volume are represented in Fig. 1.1.

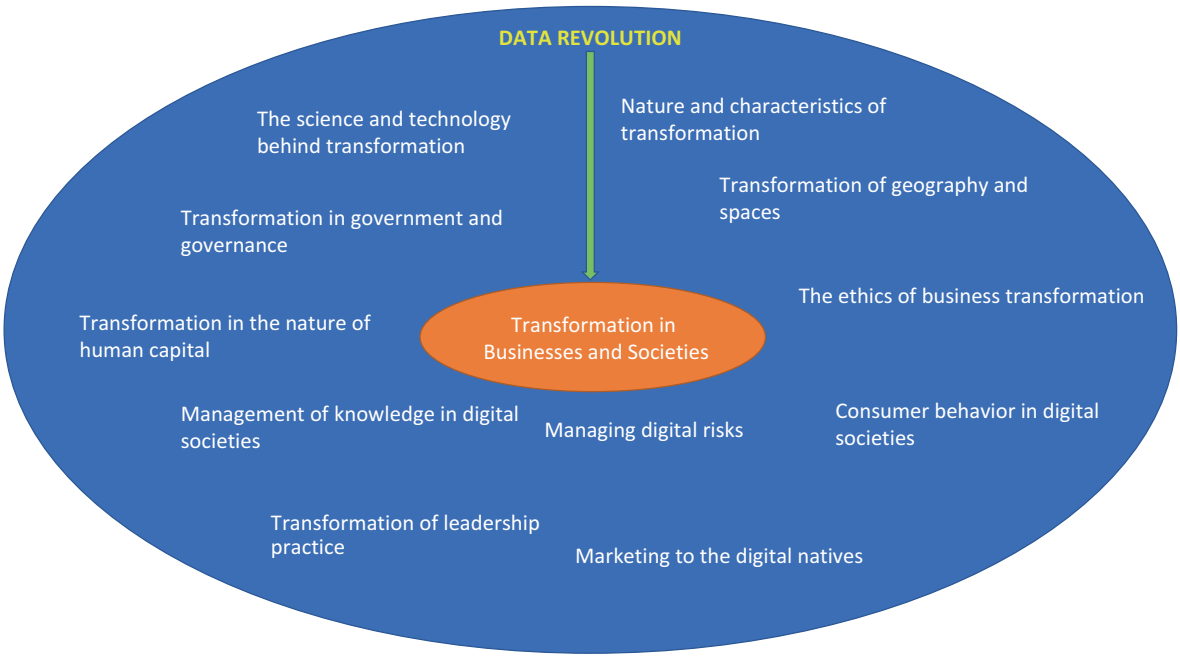


Fig. 1.1 Key themes included in this volume

A Brief Overview of the Chapters

Human ecological systems will not be the way we are currently familiar with them, as a result of our ability to manipulate big data (Graham & Shelton, 2013; Hampton et al., 2013). Ecology will not remain an esoteric pursuit, too. In the chapter titled 'Human Flourishing in Smart Cities', Maria Lai-Ling Lam calls for the importance of the development of enlightened citizens, rather than consumers, as the guardians of the smart cities managed by data-driven systems. As observed by Blagojević and Tufegdžić (2016), while technologies are disrupters, they also help preserve and refresh heritage of our geographical areas. 'The Making of Data Driven Sustainable Smart City Communities in Holiday Destinations' by Sudipta Kiran Sarkar, Michalis Toanoglu, and Babu George explains how data-centered technologies create smart and sustainable destinations for travelers. These chapters recognize the need for understanding digital technologies and data-driven systems as foundations of sustainable urban spaces.

Management of knowledge has gained a new meaning, as a result of our increased abilities to see insights in big data (Ding, Wang, & Wu, 2013). The industrial revolution 4.0 has brought in its associated phase of knowledge management 4.0 (Neumann, 2018). Asha Thomas and Meenu Chopra in their chapter 'On How Big Data Revolutionizes Knowledge Management' elaborate on this. To be able to see patterns as an emergent phenomenon needed newer heuristics and ways of thinking (Lv, Song, Basanta-Val, Steed, & Jo, 2017).

Social media is the frontline poster boy of digital transformation as the layman knows it. Social media strategy, more than social media technologies, drives the digital transformation directions of a vast majority of our organizations, observe Kane, Palmer, Phillips, Kiron, and Buckley (2015). While planning to revamp businesses through digital transformation, social media strategy cannot be let to the periphery (Westerman & Bonnet, 2015). Social media has become a central means by which to engage with the digital natives accustomed to this, and the decentralized nature of the social media makes engagement a complex enterprise (Stiver, Barroca, Minocha, Richards, & Roberts, 2015). The chapter 'Social Media Technologies and Export Marketing' by Frederick Avornyo et al. highlights aspects of it, in the particular context of changes in the age-old practice of exporting. Ogechi Adeola, Robert Ebo Hinson, and Olaniyi Evans, in their chapter 'Social Media in Marketing Communications: A

Synthesis of Successful Strategies for the Digital Generation', take a broader view and offer a synthesis of strategies for the digital generation.

Analytics is the backbone of what could be called a learning society (Larson & Chang, 2016). There is no exaggeration if someone calls it the religion of the digital society. In the chapter 'Data Analytics and Predictive Analytics: How Technology Fits in the Equation', Brian J. Galli and Gabrielle Muniz explain data analytics and predictive analytics through their various applications and examples of their benefits. Several insightful examples are provided to bring home the views. In a sector-focused chapter titled 'Artificial Intelligence and Robotics Technology in the Hospitality industry: Current Applications and Future Trends', Li Yang, Tony L. Henthorne, and Babu George explain the transformation of the hospitality industry, thanks in particular to advancements in robotics. All these chapters have a heavy emphasis upon the science and technology backbones of the data revolution that transformed businesses and societies.

Project management as an area of practice has changed dramatically, thanks to the power of information and communication technologies (Hoegl & Muethel, 2016). Strategy is traditionally distinguished from tactics and operations, in terms of time span and decisional level in the organizational hierarchy (Cascio & Montealegre, 2016). In the flat and fluid organizations in the networked economy, these distinctions have blurred away (Kahler, 2015). 'Digital Technology to Enhance Project Leadership Practice: The Case of Civil Construction' by John Ekechukwu and Thorsten Lammers portray how technologies increase the efficiency and effectiveness of project management, in the specific case of construction practices.

Organizations can now run on autopilot, except when major game changes are required (Goldspink & Kay, 2003). This also changes the role of people in organizations (Wirtz et al., 2018). To be a digital native is a radically new conception of life (Akçayır, Dündar, & Akçayır, 2016). In 'Effective Engagement of Digital Natives in the Ever-Transforming Digital World', Anju Varghese Philip and Zakkariya K. A. attempt to throw some light on how to manage digital natives both as internal as well as external stakeholders of the businesses. In the chapter titled 'Robotics, Artificial Intelligence, and the Evolving Nature of Work', Craig Webster and Stanislav Ivanov argue that humans will have to see their relationship to the job market differently and there will have to be an appropriate political response to the new economic landscape with changes in taxation and new ways of ensuring economic and political stability.

Big data does offer the promise of data-driven decision-making and change management. It has helped us identify subtle trends and relationships among key variables (Kshetri, 2009). Yet, the panacea has attracted so much unwanted attention from unscrupulous elements, too (Zhang, Xiao, Ghaboosi, Zhang, & Deng, 2012). Yang Liu, in the chapter titled 'The Dark Side of Big Data: Personal Privacy, Data Security, and Price Discrimination', shares some of these concerns. Dealing with cybercriminals needs new tactics and tools—many law enforcement departments are at the trailing edge of the curve in this regard. Given this, community policing has found a new reason for existence. Nonprofits and private sector corporations have taken the lead over government agencies in cyber defense—which, many consider to be problematic. On a related topic, 'Risk Management in the Digital Era: The Case of Nigerian Banks' by Tankiso Moloi and Oluwamayowa Olalekan Iredele examines an associated issue—the management of risks in the banking industry, using digital technologies.

Harford (2014) wonders if big data is a big mistake. Say, are we safe from computer hackers, and is the government doing enough to prepare? While the answer is not simple, this is a question worth asking, as also observed by Cuzzocrea (2014). Despite the risks, the inclusive process of knowledge creation and free flow of knowledge together could make our organizations significantly more responsive to stakeholder expectations and environmental cues (Boyd & Crawford, 2012). In 'Digital Finance for Financial inclusion and Inclusive Growth', Md. Nur Alam Siddik and Sajal Kabiraj elaborate on how technologies are levelers that could expand the inclusiveness of the financial sector.

Once upon a time, digital governance meant merely building websites for government agencies and using emails for communications (Pardo, 2000). West (2005) indicates that digital technologies have improved the efficiency and effectiveness of governments worldwide. Open data promises to open new vistas, expand our horizons, and stimulate innovation as never before. E-governance has replaced many traditional structures in governance, brought in transparency and interoperability (Gottschalk, 2009). It has made decision-making quicker and qualitatively better. 'An Evaluation of the National Open Government Data (OGD) Portal of the United Arab Emirates' by Stuti Saxena is an attempt to portray the success of one such initiative.

Current level of development of digital technologies offer new possibilities of thinking and doing. A particular issue that we need to overcome is the fact that, even as our societies have become data rich, they are still

connection poor. 'Understanding Consumer Behavior in Technology Mediated Spaces' by İlayda İpek details the relations between consumers and sellers have transformed by means of various technologies that permit co-creation experiences.

Conclusion

Human beings were always social animals, but our wealth and its transfer could not have been embedded into the social relations until very recently. Digital currencies like bitcoins decentralized wealth from the clutches of the central banks and the nation states (Tapscott & Tapscott, 2016). Barter system is making a new homecoming, albeit with digital coins. The digital technologies have made innovations happen in unexpected corners of the world (Iansiti & Lakhani, 2017). Also, innovations have become more democratic and decentralized thanks to the 'networked brains' behind many of these innovations. The dissemination of innovations now happens at pace hither to unimaginable—actually, many new products launches now happen in the peripheral corners of the world, even before people in the developed countries get to buy them.

The increased choices online imply reduced loyalty (Connor, 2018). Cut throat competition also means reduced product lifecycles (De Bettignies & Duchêne, 2015). The existence of abundant online spaces for consumer activism means increased need for investing in customer relationship activities. Corporations increasingly employ big data analysis to gain key insights about consumers and markets. It is interesting to observe the challenges and solutions in digital marketing, as they evolve in increasingly digitized societies. In fact, most contemporary discussions on digital transformations directly or indirectly refer to digital marketing in general and social media marketing in particular. There is no better combined reflection of the impact of technological revolutions upon business and societies than social media and no wonder several chapters in this volume mentions social media in some way or other.

Digital technology integration into more traditional products means there is potential for disruption there, too (Yaqoob et al., 2016). Overall, digital technologies are enablers of disruption in every aspect of life. For instance, the idea of leadership is not anymore tied to hierarchy and structure. Knowledge leaders are a new breed. Isolated people sitting in different corners of the world but working together virtually as a team was unimaginable until recently. Geography is no longer a barrier. So also, is

the traditional hierarchy built around mechanical principles: seamless cooperation and collaboration across organizations mean greater chances to synergize organizational resources.

With people spending significant portions of their lives online, criminals too migrated to the cyberspace (Saini, Rao, & Panda, 2012). Digital crimes do not end in financial fraud and malpractices. People's physical safety and security could seriously be compromised, too. Intrusions into the privacy of individuals have similar dimensions. Our increasing reliance upon big data results in the sidetracking of morals and meaning from the equations. Also, big data cannot effectively identify the micro-level nuances impacting grassroot levels—be it in organizations or in the civil society. The massive surge of data analytics as a profession is fueled by our need to tie everything to measurable data. If data science does not grow up enough to reveal the 'truth beyond data', many futurists believe we will be faced with stale robotic existence.

It is our collective amnesia that makes us think technological revolutions did not have a history. It is to be noted that mankind has used technologies ever since the dawn of civilizations. Most often, technologies have helped us advance—or, advances in technologies reflect our quest for leading better lives. However, there are also waypoints in our technological development wherein we made mistakes. Certain technologies not just failed, but also failed us. It is important to learn from these mistakes, as we design our future societies around a new set of technologies.

References

Akçayır, M., Dündar, H., & Akçayır, G. (2016). What makes you a digital native? Is it enough to be born after 1980? *Computers in Human Behavior, 60*, 435–440.

Antons, D., Joshi, A. M., & Salge, T. O. (2018). Content, contribution, and knowledge consumption: Uncovering hidden topic structure and rhetorical signals in scientific texts. *Journal of Management*. https://doi.org/10.1177/0149206318774619

Aral, S., Dellarocas, C., & Godes, D. (2013). Introduction to the special issue—Social media and business transformation: A framework for research. *Information Systems Research, 24*(1), 3–13.

Aronowitz, S. (1994). *The jobless future: Sci-tech and the dogma of work*. University of Minnesota Press.

Blagojević, M. R., & Tufegdžić, A. (2016). The new technology era requirements and sustainable approach to industrial heritage renewal. *Energy and Buildings, 115*, 148–153.

Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society, 15*(5), 662–679.

Brooke, H. (2016). Inside the digital revolution. *Journal of International Affairs, 70*(1), 29–53.

Caputo, F., Evangelista, F., Perko, I., & Russo, G. (2017, September). The role of big data in value co-creation for the knowledge economy. In *10th Annual Conference of the EuroMed Academy of Business.*

Cascio, W. F., & Montealegre, R. (2016). How technology is changing work and organizations. *Annual Review of Organizational Psychology and Organizational Behavior, 3*, 349–375.

Collin, A., & Young, R. A. (2000). *The future of career*. Cambridge University Press.

Connor, J. M. (2018). Loyalty: The emotion of future expectation, felt now, based on the past. In *Emotions, everyday life and sociology* (pp. 37–51). Routledge.

Cook, N. (2017). *Enterprise 2.0: How social software will change the future of work.* Routledge.

Cuzzocrea, A. (2014, November). Privacy and security of big data: Current challenges and future research perspectives. In *Proceedings of the First International Workshop on Privacy and Security of Big Data* (pp. 45–47). ACM.

Davidson, W. H. (1993). Beyond re-engineering: The three phases of business transformation. *IBM Systems Journal, 32*(1), 485–499.

De Bettignies, J. E., & Duchêne, A. (2015). Product market competition and the financing of new ventures. *Management Science, 61*(8), 1849–1867.

Ding, G., Wang, L., & Wu, Q. (2013). Big data analytics in future internet of things. *arXiv preprint arXiv:1311.4112.*

Goldspink, C., & Kay, R. (2003). Organizations as self-organizing and sustaining systems: A complex and autopoietic systems perspective. *International Journal of General Systems, 32*(5), 459–474.

Gottschalk, P. (2009). Maturity levels for interoperability in digital government. *Government Information Quarterly, 26*(1), 75–81.

Graham, M., & Shelton, T. (2013). Geography and the future of big data, big data and the future of geography. *Dialogues in Human Geography, 3*(3), 255–261.

Hampton, S. E., Strasser, C. A., Tewksbury, J. J., Gram, W. K., Budden, A. E., Batcheller, A. L., … Porter, J. H. (2013). Big data and the future of ecology. *Frontiers in Ecology and the Environment, 11*(3), 156–162.

Harford, T. (2014). Big data: A big mistake? *Significance, 11*(5), 14–19.

Hoegl, M., & Muethel, M. (2016). Enabling shared leadership in virtual project teams: A practitioners' guide. *Project Management Journal, 47*(1), 7–12.

Iansiti, M., & Lakhani, K. R. (2017). The truth about blockchain. *Harvard Business Review, 95*(1), 118–127.

Kahler, M. (Ed.). (2015). *Networked politics: Agency, power, and governance.* Cornell University Press.

Kane, G. C., Palmer, D., Phillips, A. N., Kiron, D., & Buckley, N. (2015). Strategy, not technology, drives digital transformation. *MIT Sloan Management Review and Deloitte University Press, 14*, 1–25.

Kshetri, N. (2009). Positive externality, increasing returns, and the rise in cybercrimes. *Communications of the ACM, 52*(12), 141–144.

Larson, D., & Chang, V. (2016). A review and future direction of agile, business intelligence, analytics and data science. *International Journal of Information Management, 36*(5), 700–710.

Lv, Z., Song, H., Basanta-Val, P., Steed, A., & Jo, M. (2017). Next-generation big data analytics: State of the art, challenges, and future research topics. *IEEE Transactions on Industrial Informatics, 13*(4), 1891–1899.

Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures, 90*, 46–60.

Matt, C., Hess, T., & Benlian, A. (2015). Digital transformation strategies. *Business & Information Systems Engineering, 57*(5), 339–343.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: The management revolution. *Harvard Business Review, 90*(10), 60–68.

Neumann, G. (2018). Knowledge Management 4.0—Implications of the fourth industrial revolution on knowledge management in supply chains. *Theory and Applications in the Knowledge Economy*, Poland: Poznan University of Economics and Business 452–461.

Pardo, T. (2000). *Realizing the promise of digital government: It's more than building a web site*. Albany, NY: Center for Technology in Government.

Ranjan, K. R., & Read, S. (2016). Value co-creation: Concept and measurement. *Journal of the Academy of Marketing Science, 44*(3), 290–315.

Saini, H., Rao, Y. S., & Panda, T. C. (2012). Cyber-crimes and their impacts: A review. *International Journal of Engineering Research and Applications, 2*(2), 202–209.

Stiegler, B. (2018). *Automatic Society: The future of work*. John Wiley & Sons.

Stiver, A., Barroca, L., Minocha, S., Richards, M., & Roberts, D. (2015). Civic crowdfunding research: Challenges, opportunities, and future agenda. *New Media & Society, 17*(2), 249–271.

Tapscott, D., & Tapscott, A. (2016). *Blockchain revolution: How the technology behind bitcoin is changing money, business, and the world*. Penguin.

Uhl, A., & Gollenia, L. A. (Eds.). (2016). *Business transformation management methodology*. Routledge.

Vega-Jurado, J., Juliao-Esparragoza, D., Paternina-Arboleda, C. D., & Velez, M. C. (2015). Integrating technology, management and marketing innovation through Open Innovation models. *Journal of Technology Management & Innovation, 10*(4), 85–90.

West, D. M. (2005). *Digital government: Technology and public sector performance*. Princeton University Press.

Westerman, G., & Bonnet, D. (2015). Revamping your business through digital transformation. *MIT Sloan Management Review, 56*(3), 10.

Wirtz, J., Patterson, P. G., Kunz, W. H., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: Service robots in the frontline. *Journal of Service Management, 29*(5), 907–931.

Yaqoob, I., Hashem, I. A. T., Gani, A., Mokhtar, S., Ahmed, E., Anuar, N. B., & Vasilakos, A. V. (2016). Big data: From beginning to future. *International Journal of Information Management, 36*(6), 1231–1247.

Zhang, Y., Xiao, Y., Ghaboosi, K., Zhang, J., & Deng, H. (2012). A survey of cyber crimes. *Security and Communication Networks, 5*(4), 422–437.

CHAPTER 2

Human Flourishing in Smart Cities

Maria Lai-Ling Lam and Kei-Wing Wong

Introduction

What cities *Homo sapiens* will choose to live will reflect the expectations of what our good life will be. The European Parliament (2014, p. 9) defined a smart city as "a city seeking to address public issues via ICT-based solutions on the basis of a multi-stakeholder, municipally based partnership." Smart cities, with all information and communication technologies (ICT), are assumed qualities to make us live well. Hollands (2015, p. 62) criticized that smart cities hid many issues, as information technology was assumed to "automatically make cities more economically prosperous and equal, more efficiently governed and less environmentally wasteful." Are we truthfully living happier in smart cities? Are we becoming a part of a predictable, efficient, and controllable system? Will we be disempowered and even lose our humanity in the highly rationalized system? What norms and values about a good life have been presented in the smart-city litera-

M. L.-L. Lam (✉)
Calvin University, Grand Rapids, MI, USA
e-mail: ml36@calvin.edu

K.-W. Wong
Michigan University, Ann Arbor, MI, USA
e-mail: wkei@umich.edu

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_2

ture? How should we deconstruct these norms? Are there other alternatives for us to live happier in smart cities?

During the period between 2010 and 2017, we reviewed the literature, visited more than 40 international cities in Asia, Europe, and North America, interviewed their citizens about their respective cities, and explored the thinking of how we should lead our students in colleges to have critical engagement with the pervasive and affordable digital technology in their respective cities. We chose two leading smart cities in Asia to illustrate how economic prosperity and global competitiveness are used to frame their well-being. Singapore and Songdo in Republic of Korea were chosen, as their respective national governments were highly ranked in the area of provision of online services, telecommunication connectivity, and human capacity by the United Nations' 2016 E-government survey (United Nations, n.d.). Singapore and Republic of Korea both were named as Asian Dragons in the 1980s and aggressively adopted information technology to sustain its economic growth in the twenty-first century. They also compete with each other for global investments and are often ranked very closely in terms of smart cities development. Their governments have invested heavily in the development of smart city as their national industrial strategy (Central Policy Unit, 2015; Kim, Jung, & Choi, 2016; Team Finland Future Watch Report, 2017). They both inspire what future cities in China will be and what needs to be done for human flourishing in our higher education.

This chapter proceeds with five sections. Section "Socially Constructed Norms and Values of Smart Cities" describes the socially constructed norms and values of smart cities. Section "A Tale of Two Cities: Singapore and Songdo" describes Singapore and Songdo. Section "Critique of Smart Cities and Human Flourishing" includes the critiques of smart cities and human flourishing. Section "A Paradigm for Human Flourishing in Smart Cities" proposes a new paradigm of the development of smart cities that enhances human well-being. Section "Implications to Public Policy and Higher Education" discusses the necessary changes in our public policy and our higher education.

Socially Constructed Norms and Values of Smart Cities

The smart city has become popular when cities are aggressively marketed to create economic prosperity, global competitiveness, environmental sustainability by major ICT (information and communications technology)

firms, engineering, property development, and construction companies (Hollands, 2015). The well-being of human lives will be improved in the densely connected, highly rational, and largely real-data analytic and programmable decision-making process in the smart cities (IBM, n.d.; Cicso, n.d.). White (2016) critically reviewed the marketing materials of smart cities provided by these global large ICT corporations, including Cisco and IBM, and learned that the technological intervention in the development of a smart city is legitimated by these three recurring crises: resource pressures brought on by rapid urbanization and an aging population, global climate change, and fiscal austerity and interurban competition for foreign direct investment and highly skilled workers. The need for the development of smart cities has been reframed as essential when the fear of quick urbanization, climate change, aging, meager public resources, and keen intercities competition are legitimated as the truth. Many issues related to data protection acts, inclusivity, costs, benefits re-distribution, and civil participation are not addressed in the rhetoric of technology-driven smart cities. The following section will explain how the development of smart cities has been legitimated and become part of our criteria for having a good life.

Characteristics of Smart Cities

The concept of smart cities is still quite vague. Smart cities are also named as intelligent, rational, green, sensor, wired, creative, sentient, digitized, or ubiquitous (u) cities…etc. The smart cities are equipped with all advanced technological devices such as the internet of things and multiple sensors and have become a promising discipline, industry, and academic discipline (Vanolo, 2014). Caragliu, Del Bo, and Nijkamp (2009, pp. 47–49) reviewed the literature about smart cities and summarized the characteristics of a smart city that is common to many research findings:

1. The utilization of networked infrastructure to improve economic and political efficiency and enable social, cultural, and urban development.
2. An underlying emphasis on business-led urban development.
3. A strong focus on the aim to achieve the social inclusion of various urban residents in public service.
4. A stress on the crucial role of high-tech and creative industries in long-run urban growth.

5. Profound attention to the role of social and relational capital in urban development.
6. Social and environmental sustainability as a major strategic component of smart cities.

Thus, a smart city relies on the use of the internet of things to improve economy and political efficiency. It mainly relies on business-led urban development through high-technology and creative industries. It provides many new business opportunities that lead to urban development. These opportunities are like Airbnb and Uber. It also emphasizes social inclusion, and social and environmental sustainability in the urban development. It provides technology-enabled participatory platforms such as New York City Big Apps. The smart-city proponents advocate the effective implementation and integration of smart technologies in a smart city will "improve quality of life, boost economic growth, mitigate the negative effects of climate change, and even foster more active citizenship" (Gaffney & Robertson, 2018, p. 49). People will flourish when there are synergies between economic growth and environmental protection in the smart cities.

Frost and Sullivan (2013), a global research and consulting company, estimated the size of smart-city market would be US 1.565 trillion by 2020. It defined a smart city consisting at least five of the following eight elements: "smart energy," "smart buildings," "smart mobility," "smart healthcare," "smart governance," "smart education," "smart security," and "smart citizens" (p. 2). Citizens are assumed to be better when they are smart enough to adopt various smart concepts or products in their daily lives. The smart-city industry can drive the development of other industries such as ICT, construction, communication, transportation, banking, insurance, health care, and education industry. Thus, smart city can be used as an instrument to improve national competiveness, which enhances community and quality of life. Kim et al. (2016) verified the positive impact of the smart-city industry on the national economy of Korea in terms of higher production, value-added, and employment induction effects. They learned that the key to the success of the smart-city industry was to acquire the competitive edge of the ICT and construction industries. The national wealth would be assumed to spill over into many people in different locations in the activated knowledge networks. They concluded that their study and numerous empirical studies have showed that "the various ICT applications in smart cities ultimately led to the build-up of intellectual capital or the creation of wealth by means of

the interconnection of human and social capital via activated knowledge networks" (p. 16).

Chourabi et al. (2012) reviewed six working definitions of a smart city and concluded that a smart city is like a large organic system connecting many subsystems and components to collect, connect, and analyze real-live data in different urban domains for a better city. They proposed eight critical factors for the success of government-driven smart-city initiatives: management and organization, technology, governance, policy context, people and communities, economy, built infrastructure, and natural environment. Sadly, the current practices of smart cities in many countries are used to prioritize technology more than other factors.

Indeed, there is no clear shared definition or global consensus about the concept of smart cities. The concept of smart city can be easily manipulated as a government software-driven urban managerial tool and an opportunistic public policy (Cowley, Joss, & Dayot, 2017). There is no such research about people and communities in smart cities. Gaffney and Robertson (2018, p. 49) condemned the vague definitions and outcomes of smart cites as follows:

> This lack of clarity and consensus in the academic literature opens space for city government and ICT companies to apply the term 'smart city' flexibly according to their needs and desires. Furthermore, the stated aims of smart cities—such as sustainability, livability, intelligence—are often abstract goals that are difficult to measure and define in themselves. The abstract aims allow cities to declare smartness without great accountability.

Neirotti, De Marco, Cagliano, Mangano, and Scorrano (2014) analyzed 70 international cities' smart initiatives and learned that the main focus was in the area of transportation and mobility, natural resources, and energy. "Government is the domain in which the cities included in the sample report the lowest number of initiatives" (p. 18). Asian cities have exhibited a lower coverage in the smart initiatives related to human capital and government practices than European cities. They concluded, "cities that are more active in the domains that are aimed at improving their capacity to 'sense and act' through ICT systems are less likely to differentiate the initiatives launched for soft domains related to human capital, cultural heritage, and innovation" (p. 24). Without much investment in the human capital and government in the civil society, will the wealth creation in the industries in Asian smart cities be more concentrated than in European smart cities?

Logical Reasons for the Rise of Smart Cities

The rise of smart cities is framed as an effective strategy to solve many problems in the urban cities given the demand of new urbanism and the supply of inexpensive interconnected of things. The business-led technological innovations proposed by the smart city seemed to be a response to the failure in the city development in the 1980s in the United States: suburban sprawling, car-oriented cities, serving elites all at the expense of the public. The rise of the smart city is related to the concept of smart growth and new urbanism which originated in the United States in the 1980s and later transferred to Europe. Vanolo (2014, p. 887) stated, "New urbanism in planning was aimed at improving the urban environment and the quality of life in cities by promoting communitarian ideas and limiting urban sprawl, land consumption, and proliferation of forms of development inspired by the logic of the automobile and personal mobility." Later, a group of researchers in Vienna ranked 70 medium-sized smart cities by six criteria: smart economy, smart mobility, smart governance, smart environment, smart living, and smart people. The ranking of cities in the European smart cities' reports determine the competiveness of a country. There are many voices about smart cities such as World Bank Smart Cities Knowledge Silo Breaker (KSB) and European Union Smart Cities Assessment. The objective behind the development of smart cities in Europe is to compete effectively with rival global economics as China and India aggressively develop smart cities (European Parliament, 2014). Smart city is becoming a tool to sustain its own country's economic growth and results in higher human capital and quality of life.

The supply of inexpensive processing devices, wireless coverage, rich analytic data, easy apps, and visualized output enable policy makers to estimate the risk level of each location and efficiently allocate resources. Many data in the cities are collected and translated to relevant information in a format of simple diagrams. The large information and communication companies are pursuing smart cities as their core strategies. For example, Siemens has "sustainable cities" logo, and IBM has "smarter planets and smarter cities" logo. They are assumed to provide many technocratic solutions to solve our pressing urbanization problems and even deal with natural disasters such as earthquakes. More and more professional conferences and initiatives are disseminated to use smart cities as our panacea. The rhetoric of smart cities is changing the criteria of measuring a city. A successful city needs to fulfill the criteria of being smart. Thus, smart cities

may help many governments to achieve their social, environmental, and economic targets at lower cost. This is the testimony from the assessment of smart city in the European Union (European Parliament, 2014, p. 19):

> In view of the challenges associated with growing European urbanization, as well as the wider agenda to tackle economic recovery poverty, unemployment and environmental damage, *the Europe 2020 strategy incorporates a commitment to promote the development of Smart Cities throughout Europe and to invest in the necessary ICT infrastructure and human and social capital development*. Smart Cities may play a part in helping to meet the targets set out in Europe 2020 by adopting scalable solutions that take advantage of ICT technology to increase effectiveness, reduce costs and improve quality of life.

The Europe 2020 strategy believes that the incorporation of advanced ICT technologies in smart cities will improve global competitiveness, sustainability, diversification of functions (or authority), and quality of life because European cities will also use ICT to enhance human and social capital in the universities, industries, government, and civil societies. One Polish scholar told the author that the term "smart city" was mainly used by her government to access funding from the European Union and be aligned with the Europe 2020 strategy. Smart city can be easily politically manipulated for particular design and management of cities (Vanolo, 2014).

Why Do Smart Cities Matter?

By 2030, 60% of world's population will live in urban areas and 95% of urban expansion will take place in the developing world. Fresh water supplies, sewage, public health, and the living environment will be seriously affected by the rapid urbanization. China and India will build more and more smart cities to deal with increasing urbanization problems (Central Policy Unit, 2015). According to IHS Markit (2014) research, the number of smart cities will rise from 13 to 88 within 12 years. Some notable global large information and communication corporations aggressively pursue the concept of the "smart city." IBM has "smarter planet" and "smarter cities." CISCO has "the internet of everything for cities." These big corporations always promise to have an open, distributed, collaborative, and accessible global neural system. They are assuming leadership roles in the collaboration process among government, corporations, and

civil society. They continue to play prominent roles into urban governance structure around the world. They advocate that the densely connected, highly rational, and largely real-data analytic and programmable decision-making process in the smart cities improve the well-being of their clientele.

The development of an inclusive, safe, resilient, and sustainable city is one of the 17 United Nation's sustainable development goals (United Nation, 2017). With the usage of technological innovations and the high density of cities, smart cities may be a good alternative to achieve the United Nation's sustainable development goals. The local government can collect sundry real-time data and monitor the urban system in a more centralized, holistic, and automated manner. Engineers, scientists, designers, and city planners can use the digital technology to model the realities and visualize better structures that use less energy and are more resilient to disaster. These abstract visual models can be manipulated and socially communicated to the public for the development of an inclusive, safe, resilient, and sustainable city.

The smart cities can provide the technology-enabled participatory platforms such as New York City Big Apps and Urban Mechanics initiated in Boston. These platforms are used by local governments to keep citizens informed, solicit input from citizens about improving the city, and collaborate with the citizens. Thus, smart cities can redefine how the local government connects with citizens, fosters thoughtful collaboration, and improves the internal function and quality of public service delivery. Desouza and Bhagwatwar (2014) examined the role of government and citizens in the 25 most populated US cities' technology-enabled participatory platforms and learned that there were four archetypes: citizen-centric and citizen data, citizen-centric and government data, government-centric and citizen data, and government-centric and citizen-developed solutions. They expect that there will be more citizen-initiated technology-enabled platforms that shape, implement, and evaluate public policy and the governance mechanism in smart cities.

A Tale of Two Cities: Singapore and Songdo

Over the last decade, there are favorable conditions for the investment in smart cities in Singapore and the Republic of Korea: "a centralized governance favoring shorter decision-making processes for public investments and more rapid development time for their executive, a higher economic development rate, a low political risk, and unique weather conditions that

determine particular needs, with special regard to transportation systems" (Neirotti et al., 2014, p. 15). Singapore and Korea central governments are key drives for the development of smart cities as a global competitive industry (Central Policy Unit, 2015; Kim et al., 2016; Team Finland Future Watch Report, 2017). Will these two smart cities, Singapore and Songdo in Republic of Korea, allow people to flourish in all-efficient designs? Will these two cities treat citizens' involvement as necessary benefits or barriers? Are citizens engaged in the process of designing their own cities and making changes for better future?

Singapore

Singapore is a city state and has committed to develop its country and city around the internet. It has Singapore's iN2015 and wants to become an intelligent country (Arun & Yap, 2000). It is claimed to be the smartest city on the earth according to a 2016 UK research (Juniper Research, 2016) and has a high financial well-being but a low score in well-being according to the Gallup Study (2015). Singapore's smartness is mainly driven by its long-term development in an information technology policy (Central Policy Unit, 2015). During 1981–1985, the Singapore government funded information technology initiatives to computerize government ministries, improve public service, and produce a group of computer experts. During 1985–1990, the government developed a strong export-oriented information technology industry through its National Technology Plan. During 1991 to the present, the city/state has to be transformed as information technology permeates into every part of life. Hollands (2008) concludes that the basic objective of the development of smart Singapore is to enhance national competitiveness and the quality of life of citizens. Its success is through progressively institutionalizing information technology in the government, business sector, and civil society. When one author visited Singapore in 2011, one university professor told her, "We had to work harder and harder as our economy grew faster in the past 10 years." One smart-city engineer told the author in a smart-city conference in August, 2016 in Hong Kong, "In terms of smart city practice, European cities are doing much better than Singapore as there are more privacy protections and community involvements in Europe." One Singaporean also lamented, "We have many intelligent devices in the city. We have to pay a higher tax." Singaporeans experience much less chance to be consulted about smart-city strategy or develop citizen initiatives than Europeans do

(Council of the European Union, 2010). One expert who has developed smart cities in Europe and Singapore told the author that Singapore has just moved from its baby-stage smart-city development since it enacted the Personal Data Protection Act 2012. A group of Finland experts stated the main hurdle in Singapore's smart-city development was the lack of citizens' participation. They said, "Singapore's smart city endeavors are nearly entirely state-directed, and the lack of consultation with citizenry and other constituents could result in applications falling short of needs" (Team Finland Future Watch Report, 2017, p. 6).

Songdo, Republic of Korea

Songdo in the Republic of Korea is always claimed as an exemplar of smart cities. It is named as "the city of the future," "the world's smartest city," or "Korea's high-tech utopia" (Overview of Eco Cities, n.d.). This city is a symbol of synergy between economic growth and environmental protection. It is driven by the central government's ubiquitous city policy and system. It represents a new paradigm shift as it has been moved from the use of conventional infrastructure to an intelligent one (Shwayri, 2013). It is framed as a city that consists of the best of different cities in the world. There are combinations of commercial buildings, residential areas, parks, streams, animals, humans, and interconnected technologies. It is "smart. It is the model for next generation cities" (Songdo IBD website, n.d.). It was built from the reclaimed wetland from the Yellow Sea and was near the Incheon international airport starting in 1994 after numerous political fights between the Central government and the local Incheon City government during 1991 and 1994. During the financial crisis in 1997, the central government had set appropriate legislations to attract global companies and related foreign business using high-end technologies and to create a world-class city in Songdo. The local and municipal governments had to secure foreign investments and expand exports. The Incheon City had to choose to partner with the private sector for the development of buildings in Songdo. "In 2001, the local government had to turn to POSCO Engineering and Construction, a South Korea chaeobol, granting it the development rights with specific instructions to partner with an American developer to plan and build Sondgo International Business District (IBD), an area covering six square kilometers from the 53 square kilometers reclaimed from the Yellow Sea to build Songdo city" (Shwayri, 2013). New Songdo International City, Inc. was formed by New York

City-based Gale International (61%), POSCO E&C (30%), and Morgan Stanley Real Estate (9%). New Cities (2014) quoted the speech of the CEO of Gale International, Mr. Stanley Gale, from metropolis now: "We want to crack the code of urbanism, then replicate it. We want to build at least twenty Songdo ourselves: G20-Gale 20." Mr. Gale showed his ambitious dreams of producing 20 Songdo smart cities in the future. CISCO, a key information technology company, was chosen to connect and optimizes everything. The city was expected to be completed by 2013.

However, when the author and her research assistant visited the city and interviewed the marketing manager of New Songdo International City, Inc. in May, 2016, we were surprised there were many empty buildings and few residents in the park. The city had to be finished by 2020. The Gale International did not succeed to attract enough foreign investors in the district. The Korean government had to change the targets and to provide more infrastructure to attract residents to relocate in the city. People have to move in after the buildings are built. One developer of an information technology engineering company told us the company was blamed when it really installed smart devices because the city just wanted to look environmentally friendly to attract foreign investment. Many smart garbage collection bins were demolished because there were not enough maintenance and residents. The Songdo city is too big to fail. The government must use taxpayer money to finance this project and continue maintaining the area. During 2003 to 2013, the investment in the city was more than US$ 100 billion, while the budget amount was US$ 35 billion. The city is framed as a smart city with all computerized accessible buildings and electrical sensors such that the city can be responsive to the needs of residents. However, it was centralized developed and mainly constructed for the attraction of multinational investments (Townsend, 2013).

As more multinational investments are in China, Songdo city is now used to attract Chinese investors to reside in this city. The development of Songdo still needs daily huge investments, while South Korea would like to keep Songdo as an exemplar of a smart city. The cost of the removal of huge habitat land for birds and the resettlement of many fishermen near Yellow Sea have not been addressed in the Songdo's development report. The city was originally designed to enable foreigners or tiny proportion of local employees working for foreign corporations to enjoy a quality of life that was like the expatriates' Western cities. In summary, the key criteria to measure the success of this smart city is to attract foreign investments and increase its international competiveness. The quality of life is mainly

measured in terms of economic success created by the development of smart cities. Kim et al. (2016) reviewed the negative critiques of Songdo and specified that Songdo is "criticized as being motivated by real estate initiatives, in which the 'smart' city label is reduced to the mere adoption of ICT features for the apartments" (para. 1 under section 2.3). Many civil voices about the establishment of the city were suppressed. The top-down approach only "promotes private business interests while ignoring social needs" (Shwayri, 2013, p. 43). This is the comment done by Sennett (2012, para 7):

> Songdo represents the stupefying smart city in its architectural aspect—massive, clean, efficient housing blocks rising up in the shadow of South Korea's western mountains, like an inflated 1960s British housing estate—but now heat, security, parking and deliveries are all controlled by a central Songdo 'brain'. The massive units of housing are not conceived as structures with any individuality in themselves, nor is the ensemble of these faceless buildings meant to create a sense of place.

Lessons Learnt from Singapore and Sondgo

Singapore and Sondgo both are driven by the centralized government policies for national competitiveness and economic interests. Sondgo in the Republic of Korea is established from scratch and is an assemblage of pre-existing urban imaginaries (Vanolo, 2014). The image of Songdo cannot relate to its history and culture. It is placeless. It can be reproduced and was implemented without paying attention to the social and political inequality in different context. The seamlessly connected world in Singapore and Songdo's virtual images ignores how the existing social and political structure limits many people's human capacities to make choices and to live well. It seems that the smart-city model can be reproduced and solve cities' problems without knowing the causes of those problems. The development of smart city continues a McDonaldization effect and just wants to have more people to be in the city. Citizens can work as sensors to help their government to identify potholes or other problems by using their smartphones. Citizens are treated as consumers and entrepreneurs in these smart cities. When cities are created to be responsible for being smart according to a defined criterion in relation to technological development, environmental protection, and economic attractiveness, the public policy makers can be easily tempted not to invest their resources in helping

their citizens to increase their capacities to enjoy the benefits of smart cities. Sometimes citizens might be treated as commodities to fulfil the political agenda. Citizens have to become smart people in order to help their cities to have higher ranking and assessment of the smartness for a better global competitive score.

Critique of Smart Cities and Human Flourishing

Many smart cities are controlled by corporations, and there is no distinction between public and private space when data is collected everywhere (Boyer, 2015). The public space is coded and the privatization of public space is disturbing. Individual's right of privacy is violated when data protection act is not enacted. Cities are assumed to be efficiently managed through the centralized operating system, data analytics, big data, and urban informatics. The complexity of a city can be abstract and distanced as simple informatics. The urban development was skewed toward the positive side of technology and attracted global capital for making industrial cities. Many alternatives of promising urban development are ignored (Caragliu et al., 2009; Hollands, 2008). The management of urban development is saturated with technological devices instead of the development of social and human capital. The city of the future has been sold to the corporations without much public officers' accountability to their people. Once corporations finish their projects, they often leave these to ill-prepared local government to run and maintain with all advanced technologies.

Many citizens are vulnerable to cyberattack and ransoms without much government protection. It is difficult to attribute the cause of the problem when many sensors are interconnected and there are time gaps between the causes and results in environmental and social performance. No one has complete expertise about the function of the complex technological system adopted in the smart cities. There are multiple uncertainties, multiple normative frameworks, and multiple scenarios of future in the complex technological system (Sarlwitz, 2016). Citizens with less resources living in highly connected technologies are foreseen to be more vulnerable than before the implementation of smart cities.

Cowley et al. (2017) reviewed the role of the public in five prominent smart UK cities and found that "the public is positioned as part of the city's routine functionality, rather than assigned 'co-creational' agency" (p. 12). These five-star smart cities emphasized service-user publicness and entrepreneurial publicness and ignored political and civic publicness. The

practice of smart city in the public policy was quite opportunistic and did not draw much attention from the public. The development of smart-city initiatives in the United Kingdom comes from the Department of Business, Skills, and Innovation, not from the Department of Communities and local government. Public policy makers are eager to use the smart cities' initiative to provide economic and political efficiency. Smart cities are "politically inclusive and culturally creative in only limited ways" (Hollands, 2008, p. 304).

There are serious problems of surveillance, discrimination, and freedom infringement in the development of smart cities. The big data collected by the smart cities can be a powerful tool to exclude or include particular data. The entire process of keeping or deleting particular data can be very political. There are significant power differences between knowledge of experts and citizen operators, the public and managers of information in these highly rationalized cities. Many data in smart cities are collected, coded, and operated with algorithms that are invisible or incomprehensible to the users and actors. The self-contained step-by-step set of operation in the algorithms is programmed without much transparency and public involvement. The algorithm may come from the set of data that affirms the existing social and political discrimination. It is questionable whether the public has the capability to meaningfully interact or engage in the discussion of the existing algorithm that governs the daily operation of the city. How can the existing algorithm be undergoing continuous improvement when so many people are incapable to criticize it even when the algorithm is claimed to be transparent to the public? It is the question who owns the data and who will not be tempted to sell the data. There are different social structures using different algorithms to make sense of data as information and form opinions about a person. There is information and power asymmetric in different social structures that can result is serious reputation losses to those who have less power. Sadly, those who cannot access the facilities in smart cities will be excluded from the network and may even bear a higher cost of living because the presence of many smart buildings and technology in their cities.

Rio de Janeiro (Brazil) can be a good case to illustrate how it was created in the idea of IBM and poorly managed by the local government. The city got the "2013 Smart City of the Year" in the Smart City Expo (Enbysk, 2013), is the exemplar of IBM's "smarter planet" and "smarter cities" marketing campaign. Gaffney and Robertson (2018) examined critically the implementation and functioning of two "smart

cities" systems in Rio de Janeiro and learned that the smart technologies exacerbated pre-existing digital and social economic inequalities and served the rich more in limited wealthy areas. The notable security development Rio de Janeiro was around sporting venues, tourist areas, and key transportation infrastructures. The information collected was not used for long-term decision-making. More crimes happened outside the monitored areas. The smart technology was used to "consolidate the territorial, financial, and political gains realized during a major event" (p. 58). When there were protests in 2013–2014, the government used the smart technology to disrupt activist network. The smart technology only consolidates power, accumulates wealth, and generates honor for mayors and a few people through mega sports events in Rio de Janeiro! The smart technology does not solve the existing problems of radical inequality or systematic poor governance. Gaffney and Robertson (2018, p. 48) concluded as follows:

> the smart city paradigm is not capable of addressing the most pressing needs of cities with chronic deficits in urban infrastructure and an absence of robust civil society institutions. As such, smart city systems may actually contribute to the securitization and fragmentation of urban space, exacerbating social and political divides.

Another case is the authors' home town, Hong Kong. It was ranked fourth in the Asia Pacific after Seoul, Singapore, and Tokyo in the 2013 Cohen study of smart cities (Central Policy Unit, 2015) according to Boyd Cohen's definition of smart cities:

> Smart cities use information and communication technologies (ICT) to be a more intelligent and efficient use of resources, resulting in cost and energy savings, improved service delivery, quality of life, and reduced environmental footprint—all supporting innovation and low-carbon economy. (Footnote 3, p. 1)

Hong Kong is effective at using technology to improve operation efficiency in daily lives by adopting the Octopus electronic transaction system and its robust ICT infrastructure. The mobile and household broadband penetration rates are very high. However, the quality of life of people has dropped in the recent 10 years while the city is regarded as the world's second-most unequal city in terms of household income distribution (South China Morning Post, 2017, June 9). More and more people are

complaining about the public policy and over-compact living environment. Although Hong Kong Planning department has advocated smart, green, and resilient strategy in the Hong Kong 2030+ document (Hong Kong 2030+ Website, 2016), there is not much public engagement and excitement about the policy. The public treats it as an opportunistic public policy. The 2030+ blueprint was criticized to run "counter to the goal of a true livable city" (South China Post, 2016, November 9). It was criticized not to develop actions to reach the standards listed in the 2030+ plan. Holloway (2017), CEO of Habit for Humanity Hong Kong, comments Smart Hong Kong Consultancy study only increases the income disparity as the inclusiveness of the city is missed in the 2030+ blueprint. The claim of smart-city development in Hong Kong cannot give comfort to the public who have to deal with serious economic and political inequality in their daily lives.

A Paradigm for Human Flourishing in Smart Cities

The new paradigm must be community-led urban development, not business-led urban development. The city is governed by the public who are empowered by technology. The development of smart cities should not be opportunistic in public policy and needs the involvement of the public and the awareness of the public about the values behind the decision that was made. This new paradigm embraces the investment in the soft domain of smart city including human capital and governance structure in Asian countries. Increasing the quality of people and inclusivity of the community should be the heart of the smarter city process. The government should develop data protection act to enact the rights of privacy of citizens.

ICT infrastructure and technology are used to empower the people and community to exercise their choices and to be engaged in the economic, social, and political activities in the smart cities. These technological companies need to inform citizens how their data are used and protected. Internet of people should exceed over internet of things. Multiple stakeholders must work within their perceived capabilities and values. They have to do civic experiments and also encounter technological problems in the limitations of our human brains and bodies. "But building smart cities is going to take time. It will by necessity be a long, messy, incremental process" (Townsend, 2013, p. 11). Developing smart cities for people needs to understand the perception of people about the cities and the priorities of values to precede the development. It values connectivity and

practices empathy in a culture of care and justice. Human flourishing, human connectivity, and inclusiveness would be seriously addressed.

In this new paradigm, a sustainable flourishing city is derived from the social meaning of the people in the community (James, 2015). The vision of the future of the city should come from the people, not from those who sell the technology. The evolving expectations of households and neighbors about their cities will be captured in the community dialogues. In this new paradigm, the inhabitants' symbolic meaning and collective historical memory of the city are captured and highly advocated by Christopher Alexander in his numerous books (Alexander, 1979; Alexander et al., 1977; Alexander, Davis, Martinez, & Corner, 1985; Alexander, Neis, Anninou, & King, 1987). He advocates the new paradigm of human flourishing as the rediscovery of the lost "pattern language" from the hidden memories of ordinary people about the buildings and the cities. He explains this new paradigm in urban planning and building design as the timeless way and "quality without name":

> A building or a town will only be alive to the extent that it is governed by the timeless way….This quality in buildings and in towns cannot be made, but only generated, indirectly, by the ordinary actions of the people, just as a flower cannot be made, but only generated from the seed (Alexander, 1979, ix-xi).

Without the help of architects or planners, if you are working in the timeless way, a town will grow under your hands, as steady as the flowers in your garden (Alexander, 1979, 8).

For example, when Christopher Alexander designed Eishin School in Tokyo, he listened to the voices of all future users including principals, school teachers, and students; he also designed the building on the site based on the hand-on experience. Thus, the design of smart cities for people needs to be evolved slowly based on the needs of users and achieve "timeless way" and "quality without name" through the in-depth understanding of pattern languages commonly used in the city.

The new paradigm has been practiced by Lynch (1960) in his book titled *The Image of City*. He has visited the places and interviewed the residents of Boston, New Jersey, and Los Angeles cities. He has found that people in urban cities orient themselves through the means of mental maps, which provide them emotional security. The designer needs to understand how the environment is constructed through the interrela-

tions of different groups and individuals and provides obvious structure with open patterns that allows the incorporation of participants' more complex mental representations of cities:

> Cities are the habitat of many groups, and only with a differentiated understanding of groups and individual images and their interrelations can an environment be constructed that will be satisfying to all until such knowledge is at hand, the designer must continue to rely upon the common denomination, or public image, and otherwise provide as great a variety of types of image-building material as he can devise. (p. 157)

> A city must have both an obvious structure that can be grasped immediately and also a potential structure which allow one gradually to construct a more complex and comprehensive picture. (p. 158)

In his book titled *What Time Is This Place* (Lynch, 1972), he describes how our innate sense of time affects the way we view and change our physical environment, especially in cities. He said, "Since we cannot be certain what will be most relevant in the future, we have an obligation to save some characteristic evidence of every major period—to establish an environment archive." To practice this new paradigm is to enable people to be connected to the historical meaning of the place and also to allow for future new developments.

The new paradigm is human-centered development and accepts our community-led urban development within our social and ecological limits. Ordinary people need to be equipped with the skills to know how to use the technology and to be included. They also need to be educated to know the pattern, quality, symbolic images, and mental representations of their cities when they are empowered to govern their own cities. Thus, the community-led growth is related to the grass-root smart growth movement that promotes sustainability by design and inclusive engagement from the public. "Sustainability emphasizes finiteness, limits, and ways of life that will advance that priority, rather than a continued emphasis on pursuing traditional forms of growth, but in a 'smarter' manner" (Beatly & Collins, 2000, p. 297). Thus, the symbolic meanings of the city need to be translated in a comprehensive human-centered city vision. The vision has to be articulated to the citizens, end-users, finance and corporate ICT firms, engineering, property development and construction companies, urban planners, and government.

The vision of the city is evolving and the design of the city for people comes from doing. Technology is only one of the alternatives that affects the future prosperity and development of human communities. Each individual is allowed to build the city which promotes environmental friendly habits by creating walkable neighborhoods containing a wide range of housing and job types. The government needs to be regulated by the community and then develop expertise to know how to equip the communities to develop their own cities for being resilient (Jacobs, 1961). Newman, Beatley, and Boyer (2009) urge cities to be prepared for short-term contingencies such as alternative transport availability, alternative fuel programs, and household awareness programs. Without these preparations, cities will begin to crumble and fall apart.

In summation, in the new paradigm of the development of smart city for people, we need to be patient with the complexity and ethical uncertainties in the process of developing a sustainable flourishing community that is aligned with existing social capital and structure. Policy makers must develop governance structure to protect citizens' data privacy. They also need to mobilize more engagement from the public, develop a process for social reflections about the values behind the public choices, need to develop contingent consensus, and learn by doing in the policy making. The new paradigm also challenges the major corporate ICT firms, engineering, property development and construction companies, and government to develop the city for people, not for quick financial interests and economic competitiveness.

Implications to Public Policy and Higher Education

Through the research of smart cities and the Asian exemplars, including Singapore and Songdo, we affirm the importance of development of citizens rather than consumers in our teaching and research. We need to facilitate our college students to practice community-led urban growth and embrace their civil responsibilities. We need to help our college students to wisely navigate through a vast ocean of digitized urban technology and solutions (Eshelman, Lam, & Cook, 2012; Lam, 2017; Vaill, 2007). We have to invest more smart initiatives in people and community, and governance in our home town, Hong Kong. It is essential for Hong Kong policy makers to develop trusting relationships with their citizens through responsible data policy, open and collaborative governance structure, and heavy investment in human resources if the economic benefits of

a "smart city" are expected to spillover and improve the quality of life of local citizens. Citizens flourish when they can live in cities with dignity and wisely exercise their human capacities, including rational, emotional, bodily, sensory, relational, imaginative, and language-using abilities. Their basic capabilities of humanity need to be nurtured and respected before they can exercise their freedom and make their own choices in smart cities.

Here are some suggestions to Hong Kong policy makers.

Responsible Data Policy

Need to address the issues of data ownership, data protection, data interpretation, data regulation, and data security in the Data Policy. Citizens have rights to be deeply informed about how their personal data is collected and is selected for particular decisions. Data providers should know their risks of losing their reputation when their personal data have been chosen to form opinions about themselves. The open data, big data, the chosen algorithms used in the smart-city development should be managed with integrity and be accountable to citizens. The local government needs to learn from the best practices of other smarter cities and practice a citizen-focused data policy (Eurocities, 2016). The local government needs to strengthen its internal government and provide e-government solutions for the citizens. For example, citizens can access the real data about the gas emission of our monopolized utility companies.

Open and Collaborative Governance Structure

Need to foster community-level social capital that allows citizens to develop their own cities and treasure the symbolic meaning of the city with the support of technology. Cities should not be written in the language of a few powerful ICT companies and real-estate developers. They need to practice transparent governance and allow citizens to be partners, collaborators, and co-creators in the design and implementation of the city that has her "quality without name," and comprehensive human-centered city vision. Leaders from the private and public sectors need to examine what can be measured, monitored, and addressed in the smart city having significant moral implications to humanity. For example, the tender specifications of private projects in the development of smart cities need to be consulted. The public and private parties have to be accountable for the quality of lives of citizens who have to live with the consequences of the projects.

Heavy Investment in People and Community

They need to develop robust civil society institutions; need to reach out to, engage with, and empower people to find better solutions for urban problems and develop communities with compassion; need to educate people to be creative users of those ICT solutions for better community and quality of life; need to invest people's basic knowledge and skills of using those technologies; and need to allow people to question what "smart city" models are being used and what social norms and codes that guide the practices for the smart cities. For example, Hong Kong 2030+ plan should allow people to have numerous feedback about the proposed smart-city system and question the public values behind the plan. People should be empowered to develop social capacities for reflection about the social norms or algorithms for the interpretation of data as information, and participate in the political processes balancing the needs of the community, local government, and corporations in the development of smart cities.

In the development of smart cities, citizens require new public services, regulations, and accountability from policy makers and corporations. A city in which people reside is the smartest city when people care and work on it. "We want cities that work well enough, but are open to the shifts, uncertainties, and mess which are real life" (Sennett, 2012). Hirschman (1965, p. 35) reminded us of our high degree of initial ignorance and uncertainty when we implement the technological innovations to solve the problems of the city. He said,

> The term "implementation" understates the complexity of the task of carrying out projects that are affected by a high degree of initial ignorance and uncertainty. Here "project implementation" may often mean in fact a long voyage of discovery in the most varied domains, from technology to politics.

Educators need to enrich students' interpretations of these visual data of urban city solutions and guide them to ask many questions about human well-beings in smart cities. These are some suggested questions: (1) What does it mean for us humans to live well in smart cities? (2) What various issues can arise from an account that makes room for such a variety of human lives in smart cities? (3) What practices are used that undermine the flourishing of our families, our communities, our political groupings, and our culture in smart cities? (4) Will we continue to use these open-ended human capacities to construct valuable lives that yet are not seen in smart cities?

References

Alexander, C. (1979). *The timeless way of building*. New York: Oxford University Press.

Alexander, C., Davis, H., Martinez, J., & Corner, D. (1985). *The production of houses*. New York: Oxford University Press.

Alexander, C., Ishikawa, S., Silverstein, M., Jacobson, M., Fiksdahl-King, I., & Angel, S. (1977). *A pattern language: Towns, buildings, construction*. New York: Oxford University Press.

Alexander, C., Neis, H., Anninou, A., & King, I. (1987). *A new theory of urban design*. New York: Oxford University Press.

Arun, M., & Yap, M. (2000). Singapore: The development of an intelligent island and social dividends of information technology. *Urban Studies, 37*(10), 1749–1756. Retrieved from http://www.jstor.org/stable/43196456

Beatly, T., & Collins, R. (2000). Smart growth and beyond: Transitioning to a sustainable society. *Virginia Environmental Law Journal, 19*(3), 287–322.

Boyer, C. (2015). *Public space and the smart cities in the 21st century*. Speech delivered in Hong Kong Polytechnic University on November 27, 2015.

Caragliu, A, Del Bo, C., & Nijkamp, P. (2009). Smart cities in Europe. In *3rd Central European Conference in Regional Science—CERS* (pp. 45–59).

Central Policy Unit. (2015). *Research report on smart city*. The Government of the Hong Kong Special Administrative Region. Retrieved from June 20, 2017 from http://www.cpu.gov.hk/doc/en/research_reports/CPU%20research%20report%20-%20Smart%20City(en).pdf

Chourabi, H., Nam, T., Walker, S., Gil-Garcia, J. R., Mellouli, S., Nahon, K., … Scholl, H. J. (2012). Understanding smart cities: An integrative framework. In *2012 45th Hawaii International Conference on System Science (HICSS)* (pp. 2289–2297). IEEE.

Cisco Website. (n.d.). Smart+Connected Communities: Embrace digital innovation to create new revenue and better serve your citizens. Retrieved June 19, 2017, from http://www.cisco.com/c/en/us/solutions/industries/smart-connected-communities.html

Council of the European Union. (2010). Europe 2020—Public consultation: Overview of responses. Retrieved June 20, 2017, from http://ec.europa.eu/eu2020/pdf/overview_responses.pdf

Cowley, R., Joss, S., & Dayot, Y. (2017). The smart city and its publics: Insights from across six UK cities. *Journal of Urban Research & Practice*, 1–25. https://doi.org/10.1080/17535069.2017.1293150

Desouza, K. C., & Bhagwatwar, A. (2014). Technology-enabled participatory platforms for civic engagement: The case of U.S. cities. *Journal of Urban Technology, 21*(4), 25–50.

Enbysk, L. (2013, November 27). Rio de Janeiro wins top smart city honors. *Smart Cities Council*. Retrieved June 20, 2017, from http://smartcitiescouncil.com/article/rio-de-janeiro-wins-top-smart-city-honors

Eshelman, G., Lam, M., & Cook, M. (2012). Three contributing factors to effective utilization of technology in management education and practice: Personhood, Mindfulness, and Meditation. *Journal of the North American Management Society, 6*, 24–34.

Eurocities Website. (2016). Smarter cities: City-led, citizen-focused. Retrieved June 21, 2017, from http://nws.eurocities.eu/MediaShell/media/EUROCITIES%20stmt_smarter%20cities_June2016.pdf

European Parliament. (2014). *Mapping smart cities in the EU study*. Retrieved November 7, 2016, from http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/507480/IPOL-ITRE_ET(2014)507480_EN.pdf

Frost and Sullivan. (2013). Strategic opportunity analysis of the global smart city market: Smart city market is likely to be worth a cumulative $1.565 trillion by 2020. Retrieved June 21, from http://www.egr.msu.edu/~aesc310-web/resources/SmartCities/Smart%20City%20Market%20Report%202.pdf

Gaffney, C., & Robertson, C. (2018). Smarter than smart: Rio de Janeiro's flawed emergency as a smart city. *Journal of Urban Technology, 25*(3), 47-64. Retrieved July 24, 2019, fromhttps://doi.org/10.1080/10630732.2015.1102423

Gallup Study. (2015). Singapore ranks high in financial well-being, lags in other measure. Retrieved November 7, 2016, from http://www.channelnewsasia.com/news/singapore/singapore-ranks-high-in/1941492.html

Hirschman, A. (1965). *Journeys toward progress: Studies of economic policy-making in Latin America*. New York: Anchor Books.

Hollands, R. (2008). Will the real smart city please stand up? Intelligence, progressive or entrepreneurial? *City, 12*(3), 303–320.

Hollands, R. (2015). Critical interventions into the corporate smart city. *Cambridge Journal of Regions, Economy and Society, 8*, 61–77.

Holloway, H. (2017, November 1). Homes and hope: Smart cities. *Ming Pao*. Retrieved November 7, from https://news.mingpao.com/pns/dailynews/web_tc/article/20171101/s00017/1509473399797

Hong Kong 2030+ Website. (2016). Smart, green, and resilient city strategy. *Planning Department*. Retrieved June 19, 2017 from http://www.hk2030plus.hk/document/Hong%20Kong%202030+%20A%20SGR%20City%20Strategy_Eng.pdf

IBM Website. (n.d.). Smarter Cities: New cognitive approaches to long standing challenges. Retrieved June 19, 2017, from https://www.ibm.com/smarterplanet/us/en/smarter_cities/overview/

IHS Markit. (2014). *Smart cities to rise fourfold in number from 2013 to 2025*. Retrieved October 26, 2016, from http://press.ihs.com/press-release/design-supply-chain-media/smart-cities-rise-fourfold-number-2013-2025

Jacobs, J. (1961). *The death and life of great American cities*. New York: Random House.

James, P. (2015). *Urban sustainability in theory and practice: Circles of sustainability*. London and New York: Routledge.

Juniper Research. (2016, May 22). Juniper names Singapore as "smartest" city globally. Retrieved November 7, 2016, from https://www.metering.com/news/juniper-names-singapore-as-smartest-city-globally/

Kim, K., Jung, J., & Choi, J. (2016). Impact of the smart city industry on the Korean national economy: Input-output analysis. *Sustainability, 8*, 649. Retrieved June 19, 2017, from http://www.mdpi.com/2071-1050/8/7/649

Lam, L. (2017). Empathetic leadership as an alternative paradigm for responsible supply chain management. In N. Meunjohn & A. McMurray (Eds.), *The Palgrave handbook of leadership in transforming Asia*. London: Palgrave Macmillan.

Lynch, K. (1960). *The image of the city*. Cambridge, MA: MIT Press.

Lynch, K. (1972). *What time is this place*. Cambridge, MA: MIT Press.

Neirotti, P., De Marco, A., Cagliano, A. C., Mangano, G., & Scorrano, F. (2014). Current trends in smart city initiatives: Some stylised facts. *Cities, 38*, 25–36.

New Cities. (2014, December 28). South Korea conceptualizes the ultimate smart city. Retrieved November 7, 2017, from https://newcities.org/cityquest-songdo-south-korea-conceptualized-ultimate-smart-sustainable-city/

Newman, P., Beatley, T., & Boyer, H. (2009). *Resilient cities: Responding to peak oil and climate change*. Washington, Covelo and London: Island Press.

Overview of Eco-cities Website. (n.d.). Focusing future. Retrieved June 19, 2017, from http://focusingfuture.com/reader/Overview-of-Eco-Cities.html?page_n27=5

Sarlwitz, D. (2016). Innovation, the Techno-human condition, and ethics in radical uncertainty. Presentation at the *Sixth International Society of Business, Economics, and Ethics (ISBEE) World Congress*, July 13–16, Shanghai, China.

Sennett, R. (2012, December 4). No one likes a city that's too smart. *The Guardian*. Retrieved June 17, 2017 from https://www.theguardian.com/commentisfree/2012/dec/04/smart-city-rio-songdo-masdar

Shwayri, S. (2013). A model Korean ubiquitous eco-city? The politics of making Songdo. *Journal of Urban Technology, 20*(1), 39–55. Retrieved July 24, 2019, from https://www.tandfonline.com/doi/full/10.1080/10630732.2012.735409?scroll=top&needAccess=tre.

Songdo IBD Website. (n.d.). Smart. Retrieved June 21, 2017, from http://songdoibd.com/about/#smart

South China Morning Post. (2016, November 9). A green Hong Kong in 2030? Not according to this plan. Retrieved June 19, 2017, from http://www.scmp.com/comment/insight-opinion/article/2044045/green-hong-kong-2030-and-beyond-not-according-plan

South China Morning Post. (2017, June 9). What hope for the poorest? Hong Kong wealth gap with record high. Retrieved June 19, 2017, from http://www.scmp.com/news/hong-kong/economy/article/2097715/what-hope-poorest-hong-kong-wealth-gap-hits-record-high

Team Finland Future Watch Report. (2017). *Singapore, the 1st smart nation by 2025*. Retrieved June 20, 2017, from https://www.tekes.fi/globalassets/julkaisut/Future_Watch_Singapore.pdf

Townsend, A. (2013). *Smart cities: Big data, civic hackers and the quest for a new utopia*. New York: W.W. Norton & Company Ltd.

United Nations. (2017). Goal 11: Make cities inclusive, safe, resilient and sustainable? Retrieved February 3, 2017, from http://www.un.org/sustainabledevelopment/cities/

United Nations Website. (n.d.). UN E-Government Survey 2016. Retrieved June 19, 2017, from https://publicadministration.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2016

Vaill, P. B. (2007). Organizational epistemology: Interpersonal relations in organizations and the emergence of wisdom. In E. Kessler & J. Bailey (Eds.), *The handbook of managerial and organizational wisdom* (pp. 327–355). Thousand Oaks, CA: Sage Publications.

Vanolo, A. (2014). Smartmentality: The smart city as disciplinary strategy. *Urban Studies, 51*(5), 883–898. Retrieved November 7, 2017, from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.850.4942&rep=rep1&type=pdf

White, J. M. (2016). Anticipatory logics of the smart city's global imaginary. *Urban Geography, 38*(4), 572–589.

CHAPTER 3

On How Big Data Revolutionizes Knowledge Management

Asha Thomas and Meenu Chopra

Introduction

The present century being a knowledge economy, there is significant reliance on information and the availability of intellectual capital (IC). In this scenario, knowledge management (KM) has emerged to become an integral part of every organization. Increasingly, organizations are relying on the knowledge possessed by their employees for decision-making processes. Erstwhile KM practices and platforms were driven toward capturing and codifying the available explicit knowledge into digital artifacts, which made the collective organizational knowledge available to a much wider set of practitioners. These KM practices and tools were extremely successful across diverse industries, particularly in knowledge-driven sectors such as Consulting, Information Technology (IT), life sciences, and others. However, they were limited in that they were primarily focused on codifying the explicit knowledge artifacts and providing services for searching and retrieving these artifacts from the database.

A. Thomas
Jagan Institute of Management Studies, New Delhi, India

M. Chopra (✉)
Delhi Technological University, New Delhi, India

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_3

In the present scenario, one of the major challenges facing organizations is the reconstruction of knowledge from one form to another. To successfully navigate this challenge, organizations are required to take consistent and strident steps to attain their identified goals through the efficient use of tacit knowledge. This leads to the creation of a contextual relationship between big data analytics and the management of the explicit knowledge (Davenport & Patil, 2012; Khan & Vorley, 2017). It is in this context, Khan and Vorley (2017) have identified the seminal role that big data analytics plays in capturing, acquiring, and sharing large volumes of explicit knowledge and how they may be interpreted through tacit insight. It has thus been proposed that big data and KM have a similar objective of disseminating knowledge and producing conclusive outcomes for organizations.

Technology-Driven Business Transformation

Technology has helped transform business processes and models over the years. Enterprise resource planning (ERP) and customer relationship management (CRM) helped to digitize organizational data and streamline flow of information, the new technologies like cloud computing, innovative, and cost-effective communication systems have made it possible even for small and medium enterprises (SMEs) to join the technology highway. As a result, the entire value chain has been digitized.

All the links in the chain—suppliers, distributors, retailers—were now in position to stay connected by means of these technologies. This led to a streamlined delivery process and higher efficiency of the entire system. As more and more organizations joined this technology highway, the cost of tech products decreased, starting a virtuous cycle of high demand and low price. The result was adoption of these technologies in the mainstream across the entire ecosystem.

Another simultaneous change was in progress—customer technology was being commoditized very rapidly. Widespread use of smartphones, e-commerce platforms, and availability of mobile internet were some of the major contributors in bringing technology closer to the end consumer. This new generation of consumer was not only shopping online, they were also pretty open about voicing their choices and opinions and sharing feedback on social media platforms. This information came directly from the horse mouth hence was not only true and reliable but also instantaneous. The intersection of digital technology across all ends

of the spectrum is generating colossal data every minute, and this data, if processed and analyzed, can provide extremely useful information for all stakeholders.

Advantages:

1. Transformation of business models
2. Generate new sources of revenue
3. Resource optimization
4. Maximize shareholder wealth

Challenges:

1. A centralized approach for identifying and analyzing big data may not be present—though there are talks about it, not many organizations have the process and technology to identify, capture, and analyze the data coming in from various sources.
2. Talent is another limiting factor. Capable data scientists with knowledge about the industry/firm ecosystem may not be available.
3. Change needed in strategic intent—using knowledge assets to build practices around insights drawn from data is required.

The knowledge pyramid of conversion of data–information–knowledge–wisdom holds immense opportunity for organizations, countries, and the world at large. The emergence of new technologies, proliferation of the internet and smartphones, and dominance of the industry by tech giants like Apple, Amazon, Facebook, etc. has produced humungous amount of data. This has made it imperative for the field of knowledge management to develop new ways and means of analyzing and utilizing this information for global transformation.

This revolution is not just restricted to technology—it has made rapid advances to areas like manufacturing, warehousing, and logistics. Professionals of the area believe that the emergence of Industry 4.0 will be ruled by the following four:

- data volume
- computational power and connectivity
- business intelligence and analysis capabilities
- human machine interaction advances like touch screen, voice recognition, and augmented reality

Knowledge Management

The study of KM is close to touching a 25-year mark. With initial talks about knowledge in organizations in early 1990s, "knowledge creation company" by Nonaka and Takeuchi (1995)) provided a solid ground for KM to be established as a discipline in the industry. The concept garnered a lot of attention, and numerous research papers have been published in the area since.

The literature on KM speaks about these two views of a firm—resource-based view (RBV) (Barney, 1991; Penrose, 1959; Wernerfelt, 1984) and knowledge-based view (KBV) (Grant, 1996; Kogut & Zander, 1992; Spender, 1996; Teece, 1998). While the RBV focuses on the tangible and intangible resources an organization possess, the KBV takes stock of knowledge that exists within an organization and propounds that this knowledge is the source of unique and sustainable advantage the firm enjoys in the market. Both these views are based on the premise that an organization's performance is dependent more on its internal factors and less on its product positioning in the market place.

KM and intellectual capital (IC) together identify and manage knowledge assets to provide a competitive edge to a firm. IC clearly demarcates between human capital, structural capital, and relational capital (Bontis, 1999). Human capital refers to job-related skill set, structural capital refers to knowledge that exists within the organization, and relational capital exists by virtue of relations with customers, suppliers, etc.

The study of KM focuses on:

1. Tacit/explicit knowledge—based on easier or harder to capture/share (Nonaka & Takeuchi, 1995)
2. Complexity and stickiness (Kogut & Zander, 1992; McEvily & Chakravarthy, 2002)
3. Tools and techniques employed for KM like communities of practice (COPs)
4. Organizational factors like culture, trust, infrastructure, and so on, which influence the existence and smooth flow of KM processes in the organization.

Choice of the correct KM approach is a complex process and varies with a firm's circumstances.

KM refers to the process of identification of collective knowledge that exists in an organization and leveraging it to help the organization compete in the market place (Von Krogh, 1998).

From the above, it is clear that KM and strategy must work hand in hand to achieve the desired results. Linkage between the two has been suggested earlier (Hansen, Nohria, & Tierney, 1999; Hunter, Beaumont, & Lee, 2002; McDermott & O'dell, 2001; Pan & Scarbrough, 1999). KM in an organization should link to and flow from the organization's overall strategy. Researchers have emphasized on enablers, processes, and organizational performance as part of KM strategy (Lee & Choi, 2003).

KM processes typically refer to basic operations of knowledge—creation, sharing, storage, dissemination, and usage; KM enablers provide the right environment, support, and infrastructure for these processes to exist in the organization.

It is generally recognized that KM is cross-functional and touches various aspects of an organization. Many researchers have worked on this integration between different links of KM enablers, KM processes, and organizational performance (Andreeva & Kianto, 2012; Darroch, 2005; Gloet & Terziovski, 2004; Gold, Malhotra, & Segars, 2001; Hislop, 2003; Kianto, 2011; Lee & Choi, 2003; Palacios Marqués & José Garrigós Simón, 2006; Singh, Shankar, Narain, & Kumar, 2006; Zack, 2009).

Big Data and the 3V Model

In 2001, Doug Laney, an analyst at Garter (formerly META), put forth a preliminary definition of big data using a 3V model, namely, volume, velocity, and variety. Almost a decade later, in 2011, the International Data Corporation (IDC) proposed another definition for big data, stating that "big data technologies describe a new generation of technologies and architectures, designed to economically extract value from very large volume of a wide variety of data, by enabling the high-velocity capture, discovery, and/or analysis" (Sumbal, Tsui, & See-to, 2017). In recent years, much speculation and interest has surrounded big data, largely due to the availability of colossal amounts of data. This has been made possible due to the widespread dispensation of information through electronic devices and rising innovation in the field of IT.

Big data has been identified as having the potential to revolutionize the large amount of data available in present times (Davenport & Patil, 2012).

This large volume of available data has opened business organizations to the realization that there is a critical need to manage this data efficiently by storing and using it for improving accuracy in decision-making (George et al., 2014). While traditionally businesses have been dependent on the transactional data stored in relational databases for making key decisions, critical data can, however, serve as a potential reservoir for non-traditional, less-structured data forms in the present times such as emails, weblogs, social media, photographs, sensors, etc. that can unearth useful information. What makes this data even more accessible is the declining cost of storage as well as computer power. Consequently, companies everywhere are leaning toward incorporating unconventional yet potentially very valuable data, or big data, alongside the traditional enterprise data that was hitherto being used in their business intelligence analyses (Rajpathak & Narsingpurkar, 2013).

McAfee, Brynjolfsson, Davenport, Patil, and Barton (2012) observe that big data creates a potential for basing decisions on factual information as opposed to instincts, which can serve companies advantageously by providing them with a competitive edge. The vast amount of data available in present times cannot be processed adequately by the existing and predominantly traditional technologies, and it is here that big data steps in as a superior tool to extricate knowledge from this data (Kabir & Carayannis, 2013). This has led to a realization within businesses and organizations that big data holds tremendous scope in enhancing the competitive edge by serving as a treasure trove of novel ideas, valuable insights, and new knowledge.

However, it needs to be underscored that the true potential of big data can be achieved only when it is used in decision-making processes. But it is significant to note that the conventional approach of data analysis practices cannot be applied to big data. It is here that the 3V approach to big data—Variety, Velocity, and Volume—becomes significant (Kabir & Carayannis, 2013). Each of these aspects is explained as follows:

- *Variety*: The proliferation of IT has caused a significant increase in the sheer number of sources for data in present times, which in turn has led to an immense variety in the data available. Further, these resources are classed into three distinct categories, viz., structured, semi-structured, and unstructured. Structured data is that which is stored in a secure file and can be readily accessed. Data collated from vague sources and is difficult to analyze is unstructured data. Semi-structured

data is that which is not organized in any special format but nevertheless contains some tags or identifiers which make it easier to process as compared to raw unprocessed data.

- *Volume*: The volume of data available is a challenge for organizations as the amount may run into terabytes and petabytes, making it difficult for the prevailing techniques of data analysis to process them.
- *Velocity*: Another challenge nowadays is the speed at which the data is generated and analyzed. Faster speeds of analysis are of utmost importance to organizations, as it will enable them to implement this processed data for better decision-making for the benefit of the organization.

The Contextual Relationship Between Big Data and Knowledge Management

Knowledge has been acknowledged as a fundamental aspect of any theorization on big data (Pauleen & Wang, 2017). Big data and analytics will fail to achieve the desired results in the absence of knowledge, as it is the latter that has led to the expansion and identification of the true potential of big data. Collection and analysis of data through the use of the appropriate techniques is dependent on the experience of the humans. It is human knowledge which will determine how the information extracted from big data is used and whether it will be applied in operational, tactical, or strategic domains. Consequently, the impact of knowledge on the role that big data will play in different sectors cannot be underestimated. This is further elucidated through an analysis of the major objectives of KM and the role that big data plays in their realization.

One of the key goals of KM is effective decision-making, which is currently accomplished by using the knowledge generated from big data analytics. This also serves to evidence the relationship between KM and big data (Crane & Self, 2014; Sumbal et al., 2017). As big data has been viewed as a potential source for generating new knowledge, it is a challenge for organizations to identify and positively exploit the true potential of digital transformation. For this, it is important to undertake a cohesive strategy to set up effective communication systems and processes to facilitate the exchange of information and data analytics, which will create a holistic approach toward KM in the digital context and foster a culture of data-driven decision-making within organizations.

Another goal of the KM system is the effective codification and sharing of the key knowledge assets. Unlike in the case of the traditional KM system, big data and the information herewith can be a source of new knowledge. That both KM and big data are concerned with some intangible asset in the form of data, information, knowledge or intelligence has also led to the development of what has been perceived as a "natural connection" between the two (Erickson & Rothberg, 2014). To ensure that these resources are efficiently utilized and meet the advanced requirements of the organizations, it has been recommended that advanced KM systems replace the erstwhile systems of rudimentary text mining tools, document analysis mechanisms, and knowledge-sharing systems (Intezari & Gressel, 2017). To generate higher value for the existing KM systems of organizations, these advanced KM systems will need to assimilate big data into their knowledge databases.

Big data analytics also becomes of crucial importance in the big picture, playing a vital role in capturing and sharing explicit knowledge by furthering its dissemination via tacit cognizance using big data analytics. In decision-making within organizations, this leads to gaining new insights into knowledge for effective and fastidious actions (Khan & Vorley, 2017).

The model presented below underlines projecting light on the contextual relationship between big data and knowledge management. At the base of the KM pyramid lies data that is composed of discrete facts and when viewed from the perspective of explicit knowledge, it refers to the movement of actions and processes in any organizational system, without any context. However, when the data is viewed in terms of a frame of reference or context, it acquires the form of information, often as metadata. When this information is further transformed through intuition, reflective thinking, personal experience, and learning of the user, it is converted into knowledge. In contrast, big data refers to the structured, unstructured, or semi-structured data collated from different sources. The application of analytics to large amounts of operational and transactional data, or explicit data, by relating it to its context gives rise to meaningful and valuable information, which directs the user toward the complete picture by studying and organizing the underlying trends and patterns revealed from the information (Fig. 3.1).

Clearly, information is a core aspect of knowledge. When meaningful information is supplied to business intelligence tools, it leads to the creation of actionable knowledge, or in other words, KM.

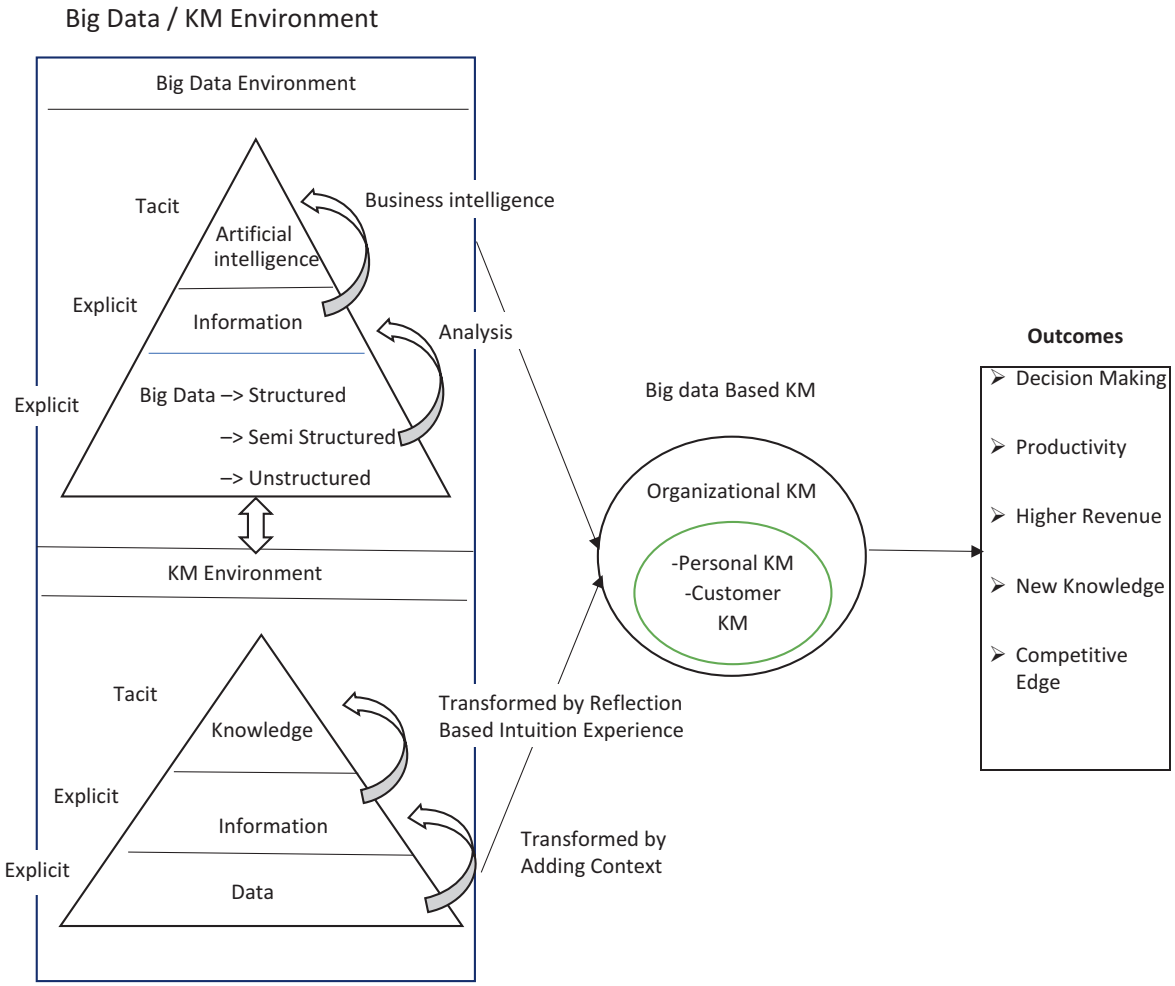


Fig. 3.1 Intersection of big data and KM in business

According to the Socialization, Externalization, Combination, and Internalization (SECI) model proposed by Nonaka, knowledge can be distinguished into two types: explicit knowledge and tacit knowledge. The former, also known as digital knowledge, is the formal and codified form of knowledge, which is captured in databases and archives and is similar to big data. On the other hand, tacit knowledge is that which exists in the minds of people and is hence much more difficult to codify. For any organization, it is the tacit knowledge of experienced employees which is significant for decision-making using big data. By bringing together the tacit knowledge of employees with the explicit knowledge that is obtained from big data, the creation of actionable knowledge and innovation is stimulated (Sumbal et al., 2017).

Clearly, big data in the form of digital data generates explicit knowledge. However, this occurs only when the available data is managed,

analyzed, and embedded in the proper context. This is done by individuals, thus making tacit knowledge a pivotal aspect of big data management (Kudyba, 2014). To this has been added the generation of predictive knowledge from big data as an ancillary goal for KM decision-making (Sumbal et al., 2017).

Big data analytics has tremendous potential for organizations, and this can be realized only by leveraging it to boost decision-making (Gandomi & Haider, 2015; Sasson, Ravid, & Pliskin, 2017) Additionally, both KM and big data were conceptualized as agents that would provide deeper insights and promote better decision-making, improve productivity, generate higher revenue, and provide a competitive edge to organizations (Elgendy & Elragal, 2014; Ohlhorst, 2012; Pauleen, 2017). The conception of novel information and knowledge or explicit knowledge from big data and its implementation in decision-making alongside the use of tacit knowledge can be extremely beneficial for enhancing the accuracy of decision-making in organizations. By focusing on such aspects as the storage and dissemination of knowledge, a more informed insight may be obtained into when and how big data can serve organizations better. At the same time, a study of variables such as the nature of knowledge (tacit or explicit), can be useful for predicting the types of knowledge apt for use in different situations and industries. This will also enable the users to gain a better understanding of when and where further inputs of big data may be advantageous (Erickson & Rothberg, 2014). Analytics can also serve useful for effective KM, leading to further improvement in the functioning of organizations, thus promoting a stronger case for the intersection of big data and KM (Sumbal et al., 2017).

Role of KM in Big Data

Role of Big Data in Personal Knowledge Management (PKM)

The manner in which individuals build knowledge by utilizing such knowledge management tools which involve the collection, absorption, and the use of knowledge and the process of innovatively using knowledge has been referred to as personal knowledge management (PKM). In PKM, the focus is on the individual and his/her quest to imbibe more information, enhance work efficiency, and socialize. One of the key attributes of PKM

is that it provides individuals with an opportunity to improve how they manage knowledge processes and interaction and collaborate to exchange knowledge and information with others (Huang, Pauleen, Scahill, & Taskin, 2018). Razmerita, Kirchner, and Sudzina (2009) have determined that the goal of PKM is to facilitate the work of individual knowledge workers instead of aiding the formation of organizations. Through this, PKM encourages the creation of personal knowledge for individuals by helping them discover and value information as a significant variable.

In recent times, with the application IT tools across sectors, several studies have sought to evaluate PKM in the context of the internet and Web 2.0, especially in terms of their use and management of personal knowledge (Rao & Chen, 2013; Razmerita et al., 2009) still PKM is an under-researched area (Pauleen, 2009). One of the core components of PKM is technology that enhances the efficiency and effectiveness of individual workers. This technology is useful for classifying ideas and information, and for archiving communication such as emails and other items so that they are easily accessible. In PKM, technology tools are considered to be imperative for enhancing the efficiency and confidence of the users as well as for the development of more effective information and knowledge management skills and dissemination and receipt of up-to-date information (Agnihotri & Troutt, 2009; Benitez & Pauleen, 2009). Liu, Wang, and Lin (2017) in their study on the applications of big data in PKM concluded that five domains:

1. *Time management*: Increased information sharing through facilitation of provision of resources which were earlier not available like software for converting the earlier not-so-easy convertible files to now free availability of such resources
2. *Computer usage efficiency management*: Retrieval of contextual information while using computers and internet gets facilitated as an individual's behavior is understood and reciprocated to through the use of Big Data.
3. *Mobile devices using behavioral management*: Behavioral parts of individuals can be analyzed using mobile phones without time and space constraints, this data is of massive use to organizations and individuals.
4. *Healthcare management*: Individuals users of mobile technology have access to unlimited apps and software, which may be of help in managing personal health by providing timely updates and reminders

about food, medicine, walk, etc. Technology has gone to the extent of being able to provide even ECG of the wearer, which may go a long way in avoiding cardiac problems by early diagnosis (Murdoch & Detsky, 2013).

5. *Browsing surfing management*: Interface between big data and an individual's surfing behavior can help in finding relevant information faster than usual. Tech devices can be made to remember the more frequently used, searched, visited websites/places hence improving the surfing experience.

The Role of Big Data in Customer Knowledge Management(CKM)

One of the key assets for any organization is customer knowledge, and hence, customer knowledge management or CKM is an important aspect for business. CKM facilitates the exchange of customer knowledge both within firms and between customers and firms in order to facilitate learning from, about, and with customers (Gibbert, Leibold, & Probst, 2002; Martín-de Castro, 2015). When customers are involved in the firm's practices through CKM, the external environment of the firm is linked to its internal environment. This, in turn, enables the transfer and sharing of information within the firm and with customers. However, there have been identified several challenges in converting customers' tacit knowledge (experience, ideas, information, difficulties, requirements, data, etc.) to explicit knowledge (useful ideas for resolving the problems of customers, for innovative services, or for upgrading current services (Taherparvar, Esmaeilpour, & Dostar, 2014).

Major retail companies such as Amazon.com, eBay, Walmart, and others have been able to successfully manage large amounts of knowledge and communicate with customers through the inclusion of real-time KM for the generation of knowledge using big data and analysis techniques (Davenport & Patil, 2012; McAfee et al., 2012; Waller & Fawcett, 2013; Sumbal et al., 2017). In these companies, the data accumulated from different sources is converted into business intelligence and knowledge via analytics. Big data analytics is also crucial in the assessment of the behavioral patterns of consumers, which is accomplished using such technology as machine-to-machine platforms and social media. The main objective here is to encapsulate the structured and unstructured data

through big data in a manner that will transform it into actionable insights by using CKM and lead to the generation, sharing, and storage of knowledge (Chan, 2014). This integration of big data and CKM for the generation of actionable insights for businesses support the creation of value for organizations.

Integration of Knowledge Management and Big Data

While both these systems are being implemented in activities like consumer survey, performance appraisal, capacity enlargement, etc., it can be observed that the independent framework of these two technologies share association with each other at multiple levels with KM playing the foundational role because of its earlier inception and time-tested performance in assisting with the installation of big data systems. However, big data is slowly adding benefits to KM in exponential ways that will make it a productive necessity for the future prospects of KM. Figure 3.2 presents the integration of the two. This integration is based on the premise that knowledge is central to all meaningful human activity. While analyzing big data, use of knowledge management is imperative as the problem or opportunity analysis must be clearly stated and form the base of the entire exercise. In the former case, big data analytics may be utilized to plan for agricultural management through geographical information or planning a city through the use of traffic patterns and trends. Opportunity identification may be purpose at hand, Google cloud platforms and Amazon web services help in the same. Organizations IT infrastructure, collaborative technologies act as the backbone for such analysis. Knowledge possessed by the employees of an organization (Tacit) and easily available coded/explicit knowledge helps us enter into the realm of big data possessed with the right weapons. Big data sources such as media, social networking sites, sensor data, archives, and log data provide humongous amount of raw data, which can be processed to find meaningful trends and opportunities by identifying and classifying the useful from the rest.

The process yields extremely useful actionable insights for managers who are in a better position to make informed decisions now. This in turn results in various benefits for the organization in terms of economic value, supply chain efficiency, and innovation.

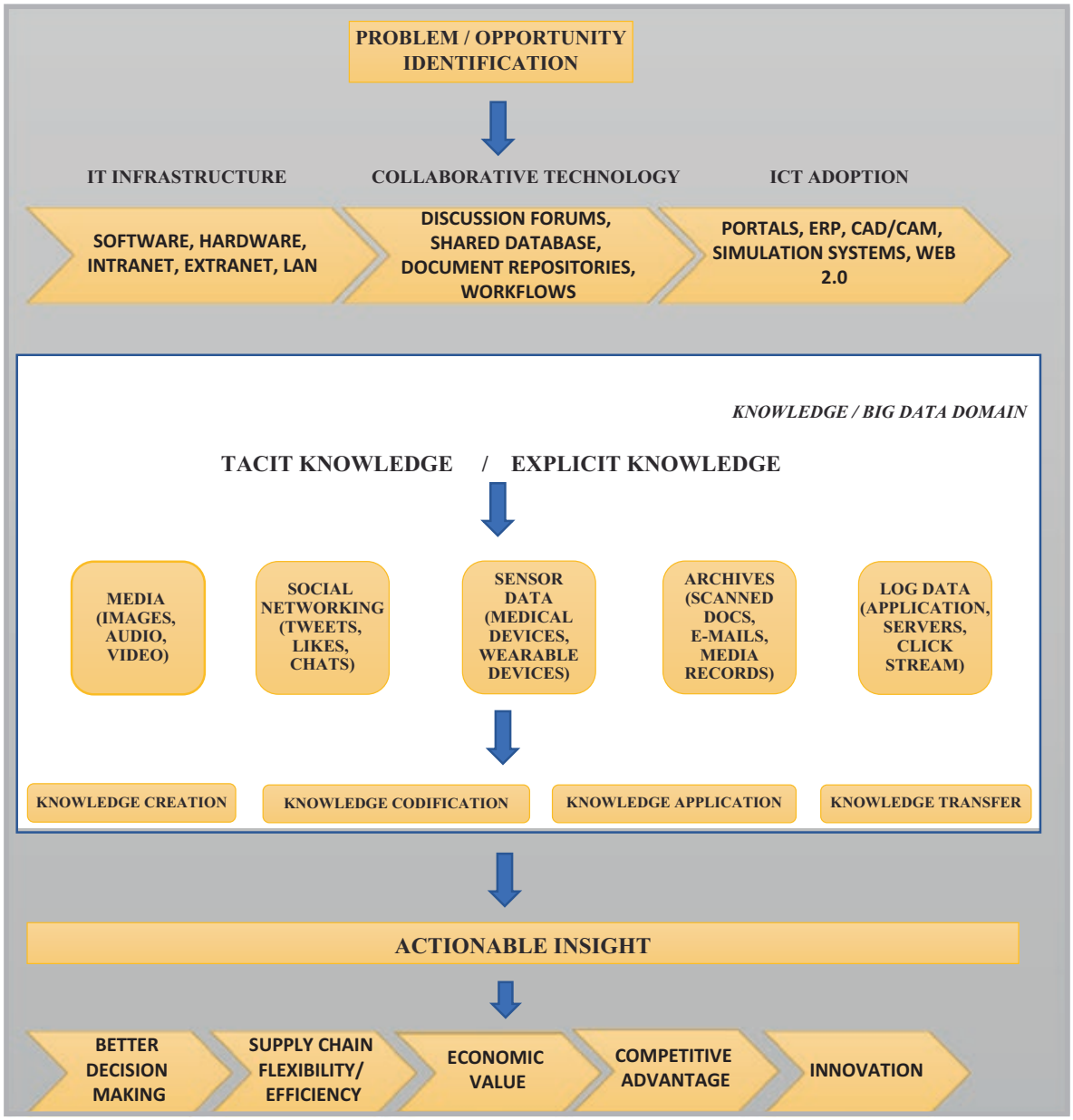


Fig. 3.2 Knowledge management—harnessing the power of big data

Therefore, the interaction of two different yet associated practices paved way to the illustrated model that can provide new business value and faster return on economic investments, enhance business operations through better decision-making, incorporate trust, promote personal and organizational-level learning, and ultimately the perks of a hybrid system that is adaptable to future needs is what comes as a bonus.

Conclusion

The twenty-first century has witnessed a major transition toward a knowledge economy of unlimited scope. During the past few years, the world has rapidly grown to connect products and services that were earlier perceived as impossible and all this can be attributed to the rising technological advancements of the second millennium. Along the same timeline, the inclusion of KM as a pivotal partner in productive practices inside companies and organizations have taken a huge leap to withstand the competition. The inevitable use of big data in a world that is run by networking has also penetrated into KM and its future, making it prone to theoretical and practical evolution on a continuous basis.

Digital transformation is inevitable in the current times. The need of the hour is that organizations understand and realize the importance of potential learning opportunities that may result from the analysis of big data. Organizations with a futuristic leadership have already woken up to the fact that "digital" is the new mantra. They are in the process of taking a stock of their current capabilities and working on a digital agenda to reskill their employees and introduce them to new learning environments and opportunities. They are in full recognition of the myriad changes' customer expectation has undergone during the last decade and appreciate the fact that some major changes will be required to meet them.

The traditional organization trained and updated their employee base by providing technical and managerial skill set training programs from time to time. While recognition of gaps in technical skills could be done using skill matrix and competence models, and appropriate training could be provided; managerial skills could be honed through communication, negotiation, and decision-making programs.

These programs were effective and yielded the desired results until the rate of change in business was incremental and the conventional theories of economics applied on the markets. With the current disruptive changes occurring in business environment across the globe, the need of the hour is development of short, situation-specific, on-the-go solutions and suggestions being provided to the employee, which may be used and referred to at their convenience of time and space.

Organizations' collective knowledge must be digitized and prioritized according to the current demands of the market. Knowledge management needs to be innovative and creative in its approach to adapt and realign itself to the changing information needs of organizations. Setting up

centers to leverage knowledge assets of the organization, data driven insights which aid decision making and alignment of strategy with business objectives will lead to success of knowledge management.

Hence, the quantitative and real-world analysis of the meeting of a robust KM with an emerging big data technology forms the framework of this study. By identifying the correlation between these big data and KM, a connection that links both together could be discovered that shed light on its different functionalities. The results of this convergence for a sole entity can be termed as personal knowledge management (PKM), expanding into customer knowledge management (CKM) and finally into organizational knowledge management (OKM) (Pee & Kankanhalli, 2009).

The same convergence also results in the advent of a new corporate methodology, which is increasingly being used in all forms of organizations; big and small for informed decision-making. This is basically the enhanced version of decision-making, which is infused by leveraging the added potential of existing KM practices and new big data capabilities, but the effectiveness of this intersection lies on how well these two practices mingle and co-exist in a corporate's workflow. It is this co-existence that has to be closely moderated for successful implementation and enhancement of KM by featuring big data as the prime contributor.

Deliberate, timely, and careful alignment of knowledge management with big data is the key to success in the current business scenario.

Limitation and Future Scope

Although big data is gaining attention in research but limited research exists putting big data and KM in one framework; hence, it remains underdeveloped within the KM literature. This chapter is a conceptual study that need empirical testing. The literatures found to the subject in discussion depicted association between big data and only few variables of knowledge management but not all areas have been thoroughly studied. The prevalent studies have only considered a few sectors to demonstrate the association between big data and knowledge management. Thus, there is still a huge scope of testing this linkage empirically in other sectors to understand its effectiveness on other industries. The review of the literature studied between big data and various areas of knowledge management, but its linkage in a broader framework still needs to be empirically tested, which opens up the scope for future research.

Appendix

Linkage between the variables

Title of the research paper	*Main findings*	*Source of literature*
Interrelationship Between Big Data and Knowledge Management: An Exploratory Study in the Oil and Gas Sector	• Generation of predictive knowledge • Improve KM capabilities • Effective decision-making	Sumbal et al., *Journal of Knowledge Management* (2017)
Big Data Systems: Knowledge Transfer or Intelligence Insights?	• Similarity between big data and KM structure	Rothberg and Erickson, *Journal of Knowledge Management* (2017)
Creation of Knowledge-Added Concept Maps: Time Augmentation via Pairwise Temporal Analysis	• High correlation between temporal assessment and subjective assessments	Sasson et al., *Journal of Knowledge Management* (2017)
Cognitive Big Data: Survey and Review on Big Data Research and Its Implications. What's Really New in Big Data?	• Five-trait framework for big data 1. Socio-technical system 2. Data space 3. Data richness 4. Knowledge management 5. Decision-making 6. Sensory/visualization presentation	Lugmayr, Stockleben, Scheib, and Mailaparampil, *Journal of Knowledge Management* (2017)
Facilitating Knowledge Management through Filtered Big Data: SME Competitiveness in an Agri-Food Sector	• Big data consumer analytics helps in tapping explicit and tacit knowledge • Effective implementation of KM processes needed in SMEs	O'Connor and Kelly, *Journal of Knowledge Management* (2017)
Big Data Text Analytics: An Enabler of Knowledge Management	• Empower KM	Khan and Vorley, *Journal of Knowledge Management* (2017)
Big Data and Knowledge Management; A Case of DÉJÀ VU or Back to the Future?	• Pattern analysis and prediction with application of algorithms • Data-driven decision-making • Big data as a contributor to KM	Tian, *Journal of Knowledge Management* (2017)
Information and Reformation in KM Systems: Big Data and Strategic Decision-Making	• Identified four types of data based decisions which aids KM systems in handling big data and advanced analytics	Intezari and Gressel, *Journal of Knowledge Management* (2017)

(*continued*)

(continued)

Title of the research paper	*Main findings*	*Source of literature*
The Concept of Big Data Applied in Personal Knowledge Management	• Five areas of personal knowledge management with application of big data: 1. Time management 2. Computer usage efficiency management 3. Mobile device behavior management 4. Health management 5. Browser surfing management	Liu et al., *Journal of Knowledge Management* (2017)
Big Data and Knowledge Management: Establishing a Conceptual Foundation	• Connection between KM, IC and application of big data and business analytics exists • With advance use of big data knowledge assets can be beneficial	Erickson and Rothberg, *Electronic Journal of Knowledge Management* (2014)
Big Data, Tacit Knowledge and Organizational Competitiveness	• Inhabited knowledge in big data has potential of creating economic value • Major source of competitive advantage	Kabir and Carayannis, *ICICKM* (2013)
Big Data Customer Knowledge Management	• Integration of CRM, CKM and big data as value creation strategies • Increase quality • Improves trustworthiness	Chan, *Communications of the IIMA* (2014).

References

Agnihotri, R., & Troutt, M. D. (2009). The effective use of technology in personal knowledge management: A framework of skills, tools and user context. *Online Information Review, 33*(2), 329–342.

Andreeva, T., & Kianto, A. (2012). Does knowledge management really matter? Linking knowledge management practices, competitiveness and economic performance. *Journal of Knowledge Management, 16*(4), 617–636.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120.

Benitez, E. A., & Pauleen, D. (2009). Brainfiltering: The missing link between PKM and PIM? *AMCIS 2009 Proceedings*, 13.

Bontis, N. (1999, August). Managing an organizational learning system by aligning stocks and flows of knowledge. In *Academy of Management Proceedings* (Vol. 1, pp. J1–J2). Briarcliff Manor, NY: Academy of Management.

Chan, J. O. (2014). Big data customer knowledge management. *Communications of the IIMA, 14*(3), 5.

Crane, L., & Self, R. J. (2014, September). Big Data Analytics: A threat or an opportunity for Knowledge Management? In *International Conference on Knowledge Management in Organizations* (pp. 25–34). Cham: Springer.

Darroch, J. (2005). Knowledge management, innovation and firm performance. *Journal of Knowledge Management, 9*(3), 101–115.

Davenport, T. H., & Patil, D. J. (2012). Data scientist. *Harvard Business Review, 90*(5), 70–76.

Elgendy, N., & Elragal, A. (2014, July). Big data analytics: A literature review paper. In *Industrial Conference on Data Mining* (pp. 214–227). Cham: Springer.

Erickson, S., & Rothberg, H. (2014). Big data and knowledge management: Establishing a conceptual foundation. *Electronic Journal of Knowledge Management, 12*(2), 101.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management, 35*(2), 137–144.

George, G., Haas, M. R., & Pentland, A. (2014). Big Data and management. *Academy of Management Journal, 57*(2), 321–326.

Gibbert, M., Leibold, M., & Probst, G. (2002). Five styles of customer knowledge management, and how smart companies use them to create value. *European Management Journal, 20*(5), 459–469.

Gloet, M., & Terziovski, M. (2004). Exploring the relationship between knowledge management practices and innovation performance. *Journal of Manufacturing Technology Management, 15*(5), 402–409.

Gold, A. H., Malhotra, A., & Segars, A. H. (2001). Knowledge management: An organizational capabilities perspective. *Journal of Management Information Systems, 18*(1), 185–214.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal, 17*(S2), 109–122.

Hansen, M. T., Nohria, N., & Tierney, T. (1999). What's your strategy for managing knowledge. *The Knowledge Management Yearbook 2000–2001, 77*(2), 106 –116.

Hislop, D. (2003). Linking human resource management and knowledge management via commitment: A review and research agenda. *Employee Relations, 25*(2), 182–202.

Huang, Y. M., Pauleen, D., Scahill, S., & Taskin, N. (2018). A PKM–based decision-making training program for personal healthcare: An action learning approach. *International Journal of Knowledge Management, 14*(3), 101–114.

Hunter, L., Beaumont, P., & Lee, M. (2002). Knowledge management practice in Scottish law firms. *Human Resource Management Journal, 12*(2), 4–21.

Intezari, A., & Gressel, S. (2017). Information and reformation in KM systems: Big data and strategic decision-making. *Journal of Knowledge Management, 21*(1), 71–91.

Kabir, N., & Carayannis, E. (2013, January). Big data, tacit knowledge and organizational competitiveness. In *Proceedings of the 10th International Conference on Intellectual Capital, Knowledge Management and Organisational Learning: ICICKM* (p. 220).

Khan, Z., & Vorley, T. (2017). Big data text analytics: An enabler of knowledge management. *Journal of Knowledge Management, 21*(1), 18–34.

Kianto, A. (2011). The influence of knowledge management on continuous innovation. *International Journal of Technology Management, 55*(1/2), 110–121.

Kogut, B., & Zander, U. (1992). Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science, 3*(3), 383–397.

Kudyba, S. (2014). Information creation through analytics. In *Big data, mining, and analytics. Components of strategic decision making* (pp. 17–48). Boca Raton: CRC Press, Taylor & Francis Group.

Lee, H., & Choi, B. (2003). Knowledge management enablers, processes, and organizational performance: An integrative view and empirical examination. *Journal of Management Information Systems, 20*(1), 179–228.

Liu, C. H., Wang, J. S., & Lin, C. W. (2017). The concepts of big data applied in personal knowledge management. *Journal of Knowledge Management, 21*(1), 213–230.

Lugmayr, A., Stockleben, B., Scheib, C., & Mailaparampil, M. A. (2017). Cognitive big data: Survey and review on big data research and its implications. What is really "new" in big data? *Journal of Knowledge Management, 21*(1), 197–212.

Martín-de Castro, G. (2015). Knowledge management and innovation in knowledge-based and high-tech industrial markets: The role of openness and absorptive capacity. *Industrial Marketing Management, 47*, 143–146.

McAfee, A., Brynjolfsson, E., Davenport, T. H., Patil, D. J., & Barton, D. (2012). Big data: The management revolution. *Harvard Business Review, 90*(10), 60–68.

McDermott, R., & O'dell, C. (2001). Overcoming cultural barriers to sharing knowledge. *Journal of Knowledge Management, 5*(1), 76–85.

McEvily, S. K., & Chakravarthy, B. (2002). The persistence of knowledge-based advantage: An empirical test for product performance and technological knowledge. *Strategic Management Journal, 23*(4), 285–305.

Murdoch, T. B., & Detsky, A. S. (2013). The inevitable application of big data to health care. *JAMA, 309*(13), 1351–1352.

Nonaka, I., & Takeuchi, H. (1995). *The knowledge creation company: How Japanese companies create the dynamics of innovation*. New York: Oxford University Press.

O'Connor, C., & Kelly, S. (2017). Facilitating knowledge management through filtered big data: SME competitiveness in an agri-food sector. *Journal of Knowledge Management, 21*(1), 156–179.

Ohlhorst, F. J. (2012). *Big data analytics: Turning big data into big money*. John Wiley & Sons.

Palacios Marqués, D., & José Garrigós Simón, F. (2006). The effect of knowledge management practices on firm performance. *Journal of Knowledge Management, 10*(3), 143–156.

Pan, S. L., & Scarbrough, H. (1999). Knowledge management in practice: An exploratory case study. *Technology Analysis & Strategic Management, 11*(3), 359–374.

Pauleen, D. (2009). Personal knowledge management: Putting the "person" back into the knowledge equation. *Online Information Review, 33*(2), 221–224.

Pauleen, D. J. (2017). Davenport and Prusak on KM and big data/analytics: Interview with David J. Pauleen. *Journal of Knowledge Management, 21*(1), 7–11.

Pauleen, D. J., & Wang, W. Y. (2017). Does big data mean big knowledge? KM perspectives on big data and analytics. *Journal of Knowledge Management, 21*(1), 1–6.

Pee, L. G., & Kankanhalli, A. (2009). A model of organisational knowledge management maturity based on people, process, and technology. *Journal of Information & Knowledge Management, 8*(02), 79–99.

Penrose, E. T. (1959). *The theory of the growth of the firm*. New York: John Wiley.

Rajpathak, T., & Narsingpurkar, A. (2013). Managing knowledge from Big Data analytics in product development. White Paper, Tata Consultancy Services.

Rao, Y., & Chen, S. (2013, June). Research on KAAS-based social knowledge collaboration service mechanism and application. In *The Proceeding of 2013 International Conference on Artificial Intelligence and Software Engineering*.

Razmerita, L., Kirchner, K., & Sudzina, F. (2009). Personal knowledge management: The role of Web 2.0 tools for managing knowledge at individual and organisational levels. *Online Information Review, 33*(6), 1021–1039.

Rothberg, H. N., & Erickson, G. S. (2017). Big data systems: Knowledge transfer or intelligence insights? *Journal of Knowledge Management, 21*(1), 92–112.

Sasson, E., Ravid, G., & Pliskin, N. (2017). Creation of knowledge-added concept maps: Time augmention via pairwise temporal analysis. *Journal of Knowledge Management, 21*(1), 132–155.

Singh, M. D., Shankar, R., Narain, R., & Kumar, A. (2006). Survey of knowledge management practices in Indian manufacturing industries. *Journal of Knowledge Management, 10*(6), 110–128.

Spender, J. C. (1996). Making knowledge the basis of a dynamic theory of the firm. *Strategic Management Journal, 17*(S2), 45–62.

Sumbal, M. S., Tsui, E., & See-to, E. W. (2017). Interrelationship between big data and knowledge management: An exploratory study in the oil and gas sector. *Journal of Knowledge Management, 21*(1), 180–196.

Taherparvar, N., Esmaeilpour, R., & Dostar, M. (2014). Customer knowledge management, innovation capability and business performance: A case study of the banking industry. *Journal of Knowledge Management, 18*(3), 591–610.

Teece, D. J. (1998). Capturing value from knowledge assets: The new economy, markets for know-how, and intangible assets. *California Management Review, 40*(3), 55–79.

Tian, X. (2017). Big data and knowledge management: A case of déjà vu or back to the future? *Journal of Knowledge Management, 21*(1), 113–131.

Von Krogh, G. (1998). Care in knowledge creation. *California Management Review, 40*(3), 133–153.

Waller, M. A., & Fawcett, S. E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics, 34*(2), 77–84.

Wernerfelt, B. (1984). A resource-based view of the firm. *Strategic Management Journal, 5*(2), 171–180.

Zack, M. H. (2009). *Knowledge and strategy*. Routledge.

Check for updates

CHAPTER 4

Social Media in Marketing Communications: A Synthesis of Successful Strategies for the Digital Generation

Ogechi Adeola, Robert Ebo Hinson, and Olaniyi Evans

Introduction

Numerous studies have investigated social media in marketing communications and have advanced our understanding of how it unfolds (e.g., Berthon, Pitt, Plangger, & Shapiro, 2012; Hudson & Thal, 2013; Kumar, Choi, & Greene, 2016). We currently have a mounting body of (1) evidence which shows that social media influences marketing outcomes (e.g., Agnihotri, Dingus, Hu, & Krush, 2016; Chung & Austria, 2010; Erdoğmuş & Cicek, 2012), and (2) several frameworks that offer an explanation of the complex social media marketing process (e.g., Effing & Spil, 2016; Felix, Rauschnabel, & Hinsch, 2017; Kietzmann, Hermkens, McCarthy, & Silvestre, 2011).

O. Adeola (✉) • O. Evans
Pan-Atlantic University, Lagos, Nigeria
e-mail: oadeola@lbs.edu.ng

R. E. Hinson
University of Ghana Business School, Accra, Ghana

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_4

In recognition of its critical importance, according to Gallaugher and Ransbotham (2010), many companies have established online presence and continuous conversations with their consumers (e.g., Old Spice, TEDx, Starbucks, Pampers, Xbox, Zappos, Ford Motor Company, Cisco, JetBlue, Uber). These companies have shifted relationships with consumers from dialogue to trialogue, with consumers engaging one another and the firms in meaningful relationships (Lipsman, Mudd, Rich, & Bruich, 2012; Mangold & Faulds, 2009). With the use of social media, these organisations have forged relationships with new as well as existing customers and formed virtual communities that interactively join forces to recognise problems and advance solutions (Tsimonis & Dimitriadis, 2014).

These interactions, in turn, have changed the traditional roles of both the customer and the seller in exchange relationships. The consumers, on their part, have reduced their dependence on the traditional media (i.e., television, radio, newspapers, and magazines) as a source of information to guide their purchasing decisions. They now demand more control over their media consumption and prefer on-demand and immediate access to information, at their convenience (Vollmer & Precourt, 2008). By utilising the social media to conduct their information searches and make purchasing decisions (Vollmer & Precourt, 2008), the consumers increasingly perceive social media as a more trustworthy source of information on products and services than corporate-sponsored communications (Foux, 2006). In particular, the consumers respond to this information in a manner 'that directly influence all aspects of consumer behaviour, from information acquisition to post-purchase expressions of satisfaction and dissatisfaction' (Mangold & Faulds, 2009, p. 361). They add value to the information by generating content, and they can influence purchase decisions of others in peer-to-peer interactions (Sashi, 2012).

Additionally, from giants like IBM and Starbucks to the local barber, organisations have realised the fact that vast amounts of information about their products and services are communicated by individual consumers on social media. They know that social media, today, provides one of the best opportunities available to a brand for connecting with prospective consumers (Neti, 2011). In recognition of these processes, therefore, managers, who consistently wield a high level of control over company-to-consumer messages, must learn new strategies to influence discussions in the social media space. Whether it is an individual, a startup, small business, or a large corporation, successful strategies must be put in place to leverage the power of social media.

However, despite the recognition that social media is effective in connecting with customers (Selina & Milz, 2009; Hackworth & Kunz, 2011; Tsimonis & Dimitriadis, 2014), there is limited understanding regarding the successful strategies for social media marketing. Most of the relatively few studies in the literature consider social media as a novel marketing tool (e.g., Berinato, 2010; Kirtiş & Karahan, 2011; Bowen & Bowen, 2016), which can boost marketing communication effectiveness (e.g., Trusov, Bucklin, & Pauwels, 2009; Dholakia & Durham, 2010), and very few focus on specific strategies for social media marketing (e.g., Mangold & Faulds, 2009). To the best of our knowledge, there appears to be limited research examining the successful strategies that firms use for their corporate fan pages on social media like Twitter, Facebook, or YouTube.

Considering this gap, the increasing recognition of the significance of social media in marketing communications, its rapidly expanding evidence base, and the numerous social media marketing frameworks that are emerging, this study seeks to increase the understanding of social media in marketing communications by undertaking a conceptual synthesis of the successful strategies in literature and from business practices. The study identifies the successful strategies in social media marketing and includes specific actions related to these strategies that can be utilised to achieve success. This study, therefore, differs from other studies in the literature, as it delves into the 'how-to' of social media marketing (i.e., specific procedures and strategies). Systematically identifying these action-oriented strategies can serve as practical guidance related to specific tasks to include in the planning and/or execution of social media marketing. It is believed that the explicit focus on specific strategies that can be used to operationalise social media in marketing communications would make a significant contribution to the literature.

In the next section, we present a brief historical overview of social media; and issues related to social media marketing are described. The following section further discusses the electronic word-of-mouth (eWOM) framework. In the fourth section, the social media marketing strategies are synthesised; while the last section covers the concluding thoughts of this study.

Brief History of Social Media Marketing

In the last decade, social media has enjoyed phenomenal success. As at 2019, Facebook, a social networking website, has 2.27 billion active users worldwide; Twitter, a micro blogging website, currently hosts more than

349 million active users with about 500 million tweets per day; and YouTube, a video-sharing website, has over 3 billion daily video views (Internetlivestats.com, 2019).

The origin of social media marketing can be traced back to the 1990s when scholars began to highlight the implications of the Internet's interactivity for marketing and suggesting the possibility of IT developments rendering traditional marketing practices obsolete (van Raaij, 1998). Within a few years, the huge impacts of the web on consumer behaviour can be described as similar to those of the Industrial Revolution, of the printing press, or the railroads (Sharma & Sheth, 2004). Due to the impact of the web, consumer preferences are no longer only influenced by the traditionally defined controllable and uncontrollable stimuli, but also by the web experience, or the online atmospherics consisting of online controllable factors such as website usability, interactivity, aesthetics, and online marketing mix (Constantinides, 2009; Fotis, 2015), and by the website's interface, quality, satisfaction, and experience (Darley, Blankson, & Luethge, 2010).

The Internet had become a great enabler of consumer power, causing a steady decline in the influence of traditional mass media (Constantinides, 2009; Urban, 2003). As noted by Sharma and Sheth (2004, p. 696), 'the web is fundamentally changing, and will continue to change marketing thought and practice', social media realigned marketing focus from a supplier to a customer standpoint, aiding a customer-centric rather than a mass-market approach. Therefore, it is the customer now, and not the product, that initiates marketing activities (Fotis, 2015). In other words, the key lies with the change in the consumer's mindset, leading to 'a culture of active engagement, immediate access to anyone and for anything and constant communication. This is, of course, something that the traditional advertisement and other marketing methods could not ever achieve' (Kaushik, 2012, p. 92).

The first companies to use social media for marketing were Usernet in 1979, Bulletin Board System in 1979, Online Service in early to mid-1980s, CompuServe in 1980, AOL in 1983, Genie in 1985, Internet Relay Chat in 1988, and ICQ in mid-1990. The earliest social media were dating sites and online forums, such as Six Degrees in 1997 and Asian Avenue, MiGente, and Black Planet in 1997–2001, Live Journal in 1999 and Massive Multiplayer Online and Role Playing Games in early 2000. Then came along Friendster in 2002, LinkedIn in 2003, Myspace in 2003, Photo bucket in 2003, Delicious in 2003, Digg in 2004, Flickers in 2004,

Facebook in 2004, YouTube in 2005, Twitters in 2006, Kontain in 2008, and Posterous in 2009 (Kaushik, 2012).

With these platforms and more in the cyberspace, social media evolved into eight types of applications, namely, social networking sites, blogs, wikis, microblogs, consumer review sites, content community sites, Internet forums, and location-based social media. These applications created a virtual space within which users listen, express, share, discuss, engage, and interact through user- and brand-generated content. As social media became popular and the number of users increased, the commercial value soon became apparent. Marketers started to consider this shift in the media landscape, rewrote conventional strategies, and transformed the basic campaigning structure.

In the last decade, social media has transformed 'the web from a "one to many" to a "many to many" communication channel, and simultaneously to a "one to one" channel of interaction, allowing users not only to communicate but also to interact and engage' (Fotis, 2015, p. 34). This transformation has implications for marketing in the form of consumer empowerment, greater information transparency, possibility for product co-creation and personalisation, content generation and publishing, capacity to sanction/reward brands for product performance, and ethical and social responsibilities (Buhalis & O'Connor, 2005; Rodriguez-Ardura Meseguer-Artola, 2010).

There are also unintended implications. Social media makes consumers highly individualistic, more information-intensive, more time-driven and demanding, dictating time and mode of communication, and with bigger expectations (Akehurst, 2009). Moreover, the huge volume of information on the web leads to information overload, negatively affecting users' ability to locate information tailored to their specific needs, and causing other pathologies of information such as infobesity and information anxiety (Bawden & Robinson, 2009).

Theoretical Framework: Electronic Word-of-Mouth (eWOM)

The framework we used addresses the linkage between word-of-mouth (WOM) and social media (see Fig. 4.1). WOM can be defined as 'the act of exchanging marketing information among consumers and plays an essential role in changing consumer attitudes and behaviour towards prod-

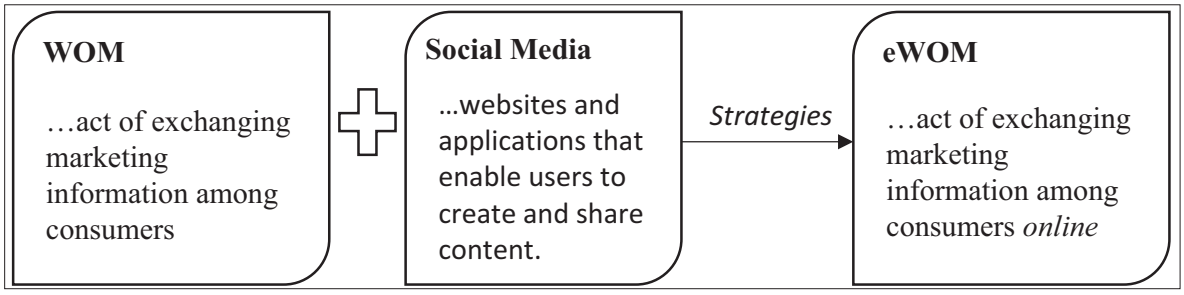


Fig. 4.1 The linkage between WOM and eWOM. Source: Authors' representation

ucts and services' (Chu & Kim, 2011, p. 47). WOM has been shown to significantly affect consumer attitudes and is seven times more effective than traditional print advertising in influencing brand-switching decisions (Engel, Kegerreis, & Blackwell, 1969; Katz & Lazarsfeld, 1955). Consumers often rely on WOM when they are in need of information to aid their purchase decisions. This is due to the fact that WOM is created and delivered by friends and family and therefore is a more trustworthy source of information than company-generated messages (Chu & Kim, 2011; Crotts 1999).

The emergence of WOM in social media has led to eWOM (electronic WOM) (see Fig. 4.1). According to Hennig-Thurau et al. (2004, p. 39), eWOM is 'any positive or negative statement made by potential, actual, or former customers about a product or company, which is made available to a multitude of people and institutions via the Internet'. eWOM occurs in a wide variety of online media, such as blogs, emails, virtual consumer communities, social networking sites, and consumer review websites and forums (Thorson & Rodgers, 2006). This eWOM acquired from friends and family is construed to 'be more credible, honest, and trustworthy than that generated from marketers, since the communicators are not compensated for the referral. Advancements in the Internet and Web 2.0 technologies now allow consumers to access personally meaningful critiques not only from friends and relatives but from strangers' (Pan & Crotts, 2012, p. 43).

In the social media marketing literature, there has been an increasing interest in understanding how eWOM affects sales, diffusion, and other marketing performance outcomes (i.e., Bruce, Foutz, & Kolsarici, 2012; Chevalier & Mayzlin, 2006; Godes & Mayzlin, 2004; Onishi & Manchanda, 2012; Trusov et al., 2009; Van den Bulte & Lilien, 2001; Villanueva, Yoo, & Hanssens, 2008). For example, Godes and Mayzlin

(2004) showed that online discussion forum activities influence television show ratings. Chevalier and Mayzlin (2006) demonstrated how user-generated online book reviews affect book sales. Trusov et al. (2009) showed that referrals to join an online social media platform affect a website's growth. Further, Onishi and Manchanda (2012) depicted that blog activity affects sales of a set of Japanese brands.

Consistent with the findings of the literature above on how eWOM affects a product's marketing performance, marketing managers combine WOM and social media to create eWOM, via effective social media marketing strategies (Fig. 4.1). Using effective strategies, the marketing managers engage the consumers in social interactions with the firm and the product and thereby get rewarded with likes, comments, or dissemination to the consumers' social connections. Consumers freely create and disseminate brand-related eWOM to friends, family, and acquaintances (Vollmer & Precourt, 2008). Through these social media interactions, consumers can show their brand preference to their connections, thus leading to eWOM (Chu & Kim, 2011). Utilising effective strategies, therefore, a firm can employ WOM on social media to generate eWOM (i.e., buzz marketing) and make its products go viral.

Successful Strategies

As more firms incorporate social media into their marketing communications, emphasis needs to be placed on creating and adopting effective strategies that can generate eWOM and make the brand go viral. The strategies outlined in this study are discrete social media practices that researchers have identified as universally beneficial for organisations to adopt. The aim here is not to generate radically new ideas for social media marketing in a discrete sense; each of the strategies is being operationalised in some organisations today. Rather, the purpose of this synthesis is to provide a measure of coherence, purpose, and success to social media marketing.

By synthesising these strategies, this study is able to aggregate the critical themes from the available literature and international experience to suggest actions that marketing practitioners can employ to ensure success. In order for a firm to achieve its goals in a social media campaign, it has to identify the successful strategies first and ensure that those strategies run in tandem with its overall marketing strategy. The strategies could serve well to nurture the creative and adaptive marketing capabilities of the organisation.

Create Good Content

Social media and content have a symbiotic relationship: without social media, the content cannot be accessible, and without great content, social media is pointless (Kaplan & Haenlein, 2011; Kietzmann et al., 2011). Studies such as Dolan (2016) have shown that social media content design and delivery influence the engagement behaviour of Facebook fans. Consumers are more likely to generate eWOM about the firm and its products when they have adequate knowledge (Mangold & Faulds, 2009). For example, Starbucks gives tips to help people use its products on its 37-million-fans Facebook page. Pages are actually a great means of connecting with audiences where they hang out.

There are three main components to social media content strategy: type of content, time of posting and frequency of posting (Chauhan & Pillai, 2013). The type of content relies on form (i.e., how the information is presented—text only, images, links, video, amongst others) and context (i.e., company voice and platform trends). Since each audience is unique, firms must identify the best type of content, time of posting, and frequency of posting in order to accrue more engagement or more likes and follows. Moreover, to get more engagement, likes and shares, firms may need to continue to inspire their audience; make their content shareable, fun and useful; use great graphics and video; and cross-promote to other channels (Berthon et al., 2012).

Get the Right People to Spread the Message

Creating successful social media marketing requires the right people to spread the message (Pfeffer, Zorbach, & Carley, 2014; Tuten & Solomon, 2014). To effectively transform an ordinary message into an eWOM phenomenon, three groups of messengers are needed: market mavens, social hubs, and salespeople. In line with the conversion theory (Moscovici, 1980), these minority group of messengers have a huge effect on the majority. According to Kaplan and Haenlein (2011, p. 256), '[c]onsistent with classical laws of concentration, 20% of messengers can be expected to carry 80% of the load; it is, therefore, especially crucial to select wisely the initial hosts' for the message.

Market mavens are individuals with access to vast marketplace information, and proactively discuss and diffuse information (Feick & Price, 1987). The market mavens first receive the message and transmit it to

their immediate social hub, who begins the eWOM phenomenon. Social hubs are people with large social connections (Goldenberg, Libai, & Muller, 2001), who facilitate immediate transmission of the message to a large number of consumers. In cases where market mavens are not particularly convincing in transmitting the message, salespeople would be needed to amplify and transmit it to the immediate social hub for further dissemination (Kaplan & Haenlein, 2011).

Create Virtual Communities of Individuals with Common Interests

The literature has shown that consumers like to network with people of similar interests and desires (e.g., Mangold & Faulds, 2009; Weber, 2009). Organisations can, therefore, leverage on this desire by creating virtual communities of like-minded individuals, centred on shared interests and values (Rolls, Hansen, Jackson, & Elliott, 2016; Weber, 2009). Virtual communities encourage interaction, sometimes focusing on a particular interest or just to communicate. Users can interact over a shared passion through various means, namely, chat rooms, message boards, virtual worlds, or social networking sites. Since a virtual community potentially crosses geographical and political boundaries, it provides unprecedented leverage to grow and maintain a firm's network, as well as create eWOM.

For example, Roadrunner Records' website, www.roadrunnerrecords.com, has a forum for fans of rock and metal music where users connect and share interests. The Obama 2008 presidential campaigns leveraged on their online presence to attract supporters and provide information. Toyota leveraged on a blog written by a team of a father and a son who embark on annual 5000-mile adventures to unusual destinations in their Toyota FJ Cruiser. Readers respond to the blog by posting their own observations, insights, and stories. For many firms, these virtual communities provide various types of information for users, with opportunities for increased interaction via forums, chat rooms, and message boards (Mangold & Faulds, 2009). Such communities provide excellent consumer engagement.

Careful Planning and Little Intervention

As with any communication exercise, it is necessary to plan social media marketing campaigns carefully prior to their launch. Once the process is set in motion, intervention may be disastrous (Berthon et al., 2012;

Kaplan & Haenlein, 2011). Specifically, a truly compelling social media marketing campaign should be able to run independently and develop its own dynamics. Firms should, therefore, develop a comprehensive and robust social media strategy which can generate eWOM dynamically. A good example is Evian, whose Roller-Skating Babies campaign was crowned by the Guinness Book of Records as 'the most viewed advertising spot', with more than 45 million online views. Evian engaged in careful planning before launching the video: it chose the right topic (10 years earlier the company had an advertising campaign based on babies), the right music (a remix of a 30-year-old rap song), and the right messengers. Once the eWOM was unleashed, the company restricted itself to reacting to the eWOM phenomenon, instead of proactively swaying the message.

Create Separate Strategies for Each Channel

Users interact with content in different ways on different social channels (Entrepreneur, 2015). While some users may find pleasure in multiple channels, some have found a single community they love and are sticking with it, even when the lures of a new social channel arise. Some may favour one network today and in six months discover that another meets their needs better. A firm must follow such customers and prospects so it can network in all the right places. It needs to adapt its strategy accordingly while maintaining the same voice across channels. A firm can also seek out inspiration by checking out its top competitors. Investigating what the competitors are doing and how they engage with their followers, can help the firm to meet its consumers' needs better (Dixon, 2000).

Design Products that Create Buzz

To stimulate conversations, products and services should be designed with consumers' desired self-images in mind (Mangold & Faulds, 2009), and must be capable of generating buzz for the business. The key is products and services that are easy to use, fun, intriguing, highly visible, and engaging are more likely to inspire eWOM (Dobele, Lindgreen, Beverland, Vanhamme, & van Wijk, 2007; Dobele, Toleman, & Beverland, 2005). The iPhone is likely to be bought and discussed among consumers who want to be perceived as 'cool'. BMW cars are likely to be discussed and purchased among auto-enthusiasts who value 'the ultimate driving experience'.

Make the Message Memorable and Interesting

The message must be sufficiently memorable and interesting to have the potential to spur eWOM. To make a message more memorable and interesting, Kaplan and Haenlein (2011, p. 257) suggested minor adjustments to the message: 'One option is to rely on true stories about real people ("My brother has a friend, John Doe …"), which are often more persuasive than corporate advertising. Another option is to use rumours, especially positive ones that reflect well on the person telling them, as they have a particularly high chance of being transmitted to others. Moreover, there are the obvious safe bets like practical short lists (e.g., "The ten best ways to lose weight"), humorous or even hilarious messages'. In general, the message must trigger the receiver's emotions (Dobele et al., 2007), often with an element of surprise, joy, or fear. Also, marketing managers must exercise caution in using messages that may be too provocative or inappropriate.

Leverage Emotional Connections

People discuss topics or subject matters in which they are emotionally connected (Sashi, 2012). Customers may be emotionally connected to finding cures for diseases such as diabetes, cancer, and heart disease. Other emotional linkages include animal rights, environmental issues, child welfare, and education. To leverage emotional connections and generate eWOM, Mangold and Faulds (2009) suggest that organisations can embrace one or more causes important to their consumers. Pepsi, for example, decided not to run any ads during the 2010 Super Bowl. Rather, it spent an estimated US$20 million of the saved budget on the Pepsi Refresh Project, an interactive online platform where users can submit and vote on ideas that have positive impacts and can help make the world a better place (Kaplan & Haenlein, 2011). Ice cream producer Ben and Jerry's supports causes that are important to the founders and its customers. These causes range from investing in the nation's children, supporting family farms, avoiding the consumption of meat from cloned animals, and promoting social and environmental concerns.

Offer Exclusivity

Everyone likes to feel special (Leung, 2013; Mangold & Faulds, 2009). By offering products, information, and special deals exclusively to the consumers, feelings of specialness can be created (Mangold & Faulds, 2009).

For example, Starbucks' stand-out feature on social media is its active online consumer engagement. Starbucks provides its customers with the option to manage and reload their coffee cards online. It also makes it possible for its consumers to locate nearby locations, check out international Starbucks spots, and apply for jobs at its stores. Likewise, Unilever gives its reader's coupons and product samples from Dove's 'Campaign for Real Beauty'. Another example is XBox's Twitter page an interesting resource, with news update, slick advertising photos, and promotions. Gamers can get instantaneous responses to their issues at its @XboxSupport handle. Also, Xbox provides phone numbers and addresses common topics online.

Harness the Power of Storytelling

Stories can be memorable; they provide the connection between facts to make them memorable (Kosara & Mackinlay, 2013). The more memorable the stories are, the more likely they are to generate eWOM (Mangold & Faulds, 2009). Refined stories that depict a real, but still discreetly created image of a business are key to humanising a brand to existing customers, and potential sales leads. For example, the Adidas campaign 'Impossible is nothing' was successful because it communicated a true story; so authentic, that it touched consumers. It left the story of Messi, the concept of the campaign, and the brand impregnated in consumers' memory for years. The phrase 'Impossible is nothing' resonated like a scream of glory in people's ears, and since then Adidas has been associated with that spirit. Such stories create vivid memories that are likely to generate eWOM, as well as traditional WOM (Mangold & Faulds, 2009). Firms must, therefore, harness the power of social media, not just to 'share' content, but to 'tell it', and to 'narrate' the brand. Social media platforms need to be seen as a conversation channel; a transmission of experiences, values, dreams, feelings, failures, successes, and truths.

Solicit Feedback

Feedbacks are important for consumer engagement. Consumers become more involved with products and organisations when they can submit feedback (Mangold & Faulds, 2009). For example, Toyota solicits and gets feedbacks from consumers in the form of accolades, criticism, and helpful suggestions on a broad range of issues through its 'Open Road

Blog'. The feedback provides a sense of community in which open communications are stimulated, and customer engagement is boosted. An organisation can leverage its network, for example, before launching products to solicit the feedback of its virtual communities (Heller Baird & Parasnis, 2011). One of the greatest benefits of social media is that a firm's network probably consists of not only existing customers, but also target customers, prospects, and industry influencers. Indeed, a firm can get a wide range of feedback on new products (i.e., whether they are worth launching in the first place; whether they are easy to use; or what should be next in the development queue) and get the all-important community buy-in for greater adoption rates at launch time.

In line with the theory of social influence (Friedkin, 2006), people are strongly influenced by how they perceive their relationship with the influencer. Likewise, consumers are more likely to communicate when they have a relationship with the product or service. In particular, some companies have been encouraging consumers to submit photos or recorded footage of their product in action. For example, a promotion for Procter and Gamble's quick-clean product, *Swiffer* challenged consumers to make short videos recounting 'how they left their old cleaning method for a new romance with Swiffer'. Then, the top ten videos were posted on YouTube and consumers were asked to vote on the winner. The contest sparked a huge amount of attention. Such online voting provides Internet users with a sense of ownership and increased engagement.

Intensify Social Media Presence to Increase Reach

Search engines are beginning to adopt signals from social media sites into their search ranking design (e.g., Bing Social Search, TripAdvisor). Social media presence can therefore greatly increase a firm's search rankings (Demers, 2015). Recent studies (i.e., Ghose, Ipeirotis, & Li, 2014) have found that a consumer utility-based ranking mechanism on product search engines, that integrates multidimensional consumer preferences and signals from social media sites leads to significant surplus gain for consumers. As with most search ranking strategies, there is one basic principle: the better experience you give your users, the higher you rank in Google. Going by the norm of reciprocity (Burger, Sanchez, Imberi, & Grande, 2009), if a firm gives its users a better experience, it has opened the widest possible channel for new potential customers to find its brand.

Include the Rest of the Marketing Communications Mix

Despite all the advantages of social media in marketing communications, it needs to be accompanied by concerted efforts with the rest of the marketing communications mix (Mangold & Faulds, 2009). The buzz surrounding social media marketing usually dies out quickly. 'Nothing tends to be more boring than yesterday's news', the old saying goes. To maintain the eWOM momentum, social media marketing must be complemented with more traditional forms of communication. Kaplan and Haenlein (2011, p. 261) illustrated this by relating the story of Wilkinson's Fight for Kisses advertisement as follows: 'To support the launch of its newly-developed disposable razor, Quattro Titanium, the company relied on a viral marketing story about a baby fighting his father for kisses from the baby's mother. This campaign consisted of an animated video and an interactive computer game, combined with a series of press announcements, radio spots, and sponsorship of the France—Ireland rugby match that took place during the same time period. Despite its limited budget of only s62,000 (about US$90,000), the campaign was a huge success and resulted in a five-percentage-point market share increase within the target group'. Social media marketing promotes satisfactory results when utilised in combination with the rest of the marketing communications mix.

The literature has also demonstrated that social media play a significant role in driving traditional media activity (Kaplan & Haenlein, 2011; Mangold & Faulds, 2009). For example, Stephen and Galak (2012) analysed 14 months of daily sales and media activity data from a microlending marketplace using a multivariate autoregressive model. The authors found that both traditional and social earned media affect sales. However, because of the higher frequency of social earned media activity after adjusting for event frequency, social earned media's sales elasticity is significantly higher than traditional earned media's. Further, social earned media tends to play a significant role in driving traditional earned media activity. Organisations should, therefore, combine social media and the traditional media in order to achieve marketing success.

Managerial Implications

The study has shown that, in social media marketing, emphasis needs to be placed on creating and adopting effective strategies that can generate eWOM and make the brand go viral. The strategies synthesised in this study will, therefore, help managers to modify their marketing input. The

budding field of social media marketing tends to borrow many of the traditional off-line techniques for managing interpersonal influence (i.e., stimulating and creating opinion leaders) while at the same time creating new techniques aided by the special characteristics of social media. Marketing managers must understand that their consumers are going online in increasing numbers and that, on the social media, they are being exposed to and influenced by their competitors. Marketing managers should, therefore, take the lead in understanding and utilising the emerging technology and strategies of social media marketing.

This study has identified a series of available strategies for harnessing the power of the social media for firms to stimulate their marketing efforts. Firstly, as with any communication exercise, it is necessary to plan social media marketing campaigns extensively before their launch. Once the eWOM is launched, the company should limit itself to reacting to the eWOM phenomenon, instead of proactively swaying the message. Additionally, in contrast to traditional marketing, which involves limited customer participation, managers need to develop new forms of advertorial content that encourage user engagement. Marketing managers are more likely to generate eWOM if they continue to inspire their audience; make their content shareable, fun and useful; use great graphics and video; and cross-promote to other channels. As well, it is necessary to create separate strategies for each channel because users interact with content differently on every social channel. A firm must follow such customers and prospects so it can network in all the right places.

Creating successful social media marketing requires the right people to spread the message. To effectively transform an ordinary message into an eWOM phenomenon, therefore, a firm must employ the three identified groups of messengers (i.e., market mavens, social hubs, and salespeople) to spread the message. While spreading the message, organisations should create virtual communities of individuals with common interests, centred on shared interests and values. Such communities provide excellent consumer engagement.

To further stimulate eWOM, products and services should be designed with consumers' desired self-images in mind to create buzz. The message must be sufficiently memorable and interesting to have the potential to spur eWOM. Marketing messages must trigger the consumer's emotions. Firms must leverage emotional connections because people discuss things to which they are emotionally connected. Likewise, people like to feel special. To leverage on this, managers must provide exclusivity.

Also, marketing managers must harness the power of stories. Stories can be memorable. The more memorable the stories are, the more likely they are to generate eWOM. Feedbacks are, as well, important for consumer engagement. Indeed, a firm can get a wide range of feedback on new and existing products and get the all-important community buy-in for all the ranges of its products. Firms must also increase social media presence in order to greatly increase their search rankings. If a firm gives its users a better experience, it has opened the widest possible channel for new potential customers to find its brand. Also, social media marketing must be complemented with the rest of the marketing communications mix. To maintain the eWOM momentum, therefore, firms should combine social media and the traditional media in their marketing campaigns.

Limitations and Implications for Future Research

Given the conceptual nature of the study and the fact that social media is an emerging field of research, the strategies should be considered exploratory. Although many global companies have successfully created eWOM with the identified strategies, surveys should be carried out to provide a better-grounded view of the social media strategies. Also, a longitudinal examination of social media strategies is necessary in order to determine how firms adjust their strategies over time.

Further, the strategies offered in this chapter suggest a number of interesting research questions. What strategies are most consistent with successful social media marketing? Some strategies have been offered, but that list is by no means exhaustive. What range of marketing strategies and associated practices flows from social media? How do the connective processes in social media contribute to the development of marketing practices?

Some social media strategies emerge over time and are not always preplanned. Investigation into those strategies means focusing qualitatively on the connective capabilities of social media and on the psychology of e-conversations that can lead to transformative change in marketing. One way to approach some of these questions would be to construct some high-level, exploratory hypotheses tying a marketing approach on social media to perceived strategic resources to organisational outcomes. One could construct a hypothesis concerning the significance of a social media strategy, and another pertaining to the degree of usage of the strategy, covering the intention and execution. For instance:

- Competitive advantage flows from appropriate marketing strategies on social media.
- Greater usage of social media marketing strategies is associated with greater relative industry performance.

A possible method for exploring these hypotheses might be to pick comparator companies, who have relatively equal strategic positions in the same industry, and gather qualitative data on their respective social media marketing schemes. Qualitative and quantitative comparisons of high-and low-performing firms in the same industry could as well be constructed to explore the hypothesis and to help in the construction of others more specific. Most significant and perhaps generalisable findings out of such a study would be not just the particular strategies themselves but the mechanism and rationale within the firm of drafting such strategies (e.g., the participants, time frames, conversations, and inputs).

Beyond the above macro-oriented questions, managers should, on a practical level, initiate their own studies focused on their specific products and services to assess the level of current eWOM and explore different strategies to increase it.

References

Agnihotri, R., Dingus, R., Hu, M. Y., & Krush, M. T. (2016). Social media: Influencing customer satisfaction in B2B sales. *Industrial Marketing Management, 53*, 172–180.

Akehurst, G. (2009). User generated content: The use of blogs for tourism organisations and tourism consumers. *Service Business, 3*(1), 51.

Bawden, D., & Robinson, L. (2009). The dark side of information: Overload, anxiety and other paradoxes and pathologies. *Journal of Information Science, 35*(2), 180–191.

Berinato, S. (2010). Six ways to find value in Twitter's noise. *Harvard Business Review, 88*(6), 34–35.

Berthon, P. R., Pitt, L. F., Plangger, K., & Shapiro, D. (2012). Marketing meets Web 2.0, social media, and creative consumers: Implications for international marketing strategy. *Business Horizons, 55*(3), 261–271.

Bowen, G., & Bowen, D. (2016). Social media: Strategic decision making tool. In *Competitive social media marketing strategies* (pp. 94–111). IGI Global.

Bruce, N. I., Foutz, N. Z., & Kolsarici, C. (2012). Dynamic effectiveness of advertising and word-of-mouth in the sequential distribution of short life cycle products. *Journal of Marketing Research, 49*(4), 469–486.

Buhalis, D., & O'Connor, P. (2005). Information communication technology revolutionizing tourism. *Tourism Recreation Research, 30*(3), 7–16.

Burger, J. M., Sanchez, J., Imberi, J. E., & Grande, L. R. (2009). The norm of reciprocity as an internalized social norm: Returning favors even when no one finds out. *Social Influence, 4*(1), 11–17.

Chauhan, K., & Pillai, A. (2013). Role of content strategy in social media brand communities: A case of higher education institutes in India. *Journal of Product & Brand Management, 22*(1), 40–51.

Chevalier, J. A., & Mayzlin, D. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research, 43*(3), 345–354.

Chu, S. C., & Kim, Y. (2011). Determinants of consumer engagement in electronic word-of-mouth (eWOM) in social networking sites. *International Journal of Advertising, 30*(1), 47–75.

Chung, C., & Austria, K. (2010). Social media gratification and attitude toward social media marketing messages: A study of the effect of social media marketing messages on online shopping value. *Proceedings of the Northeast Business & Economics Association*. (pp. 581–586).

Constantinides, E. (2009). Social Media/Web 2.0 as marketing parameter: An introduction [Online]. *International Marketing Trends Conference*, Paris 15–17 January 2009. Retrieved November 26, 2009, from http://marketing-trends-congress.com/2009_cp/Materiali/Paper/Fr/Constantinides.pdf.

Crotts, J. (1999). Consumer decision making and prepurchase information search. In *Consumer behavior in travel and tourism* (pp. 149–168). New York: The Haworth Hospitality Press.

Darley, W. K., Blankson, C., & Luethge, D. J. (2010). Toward an integrated framework for online consumer behavior and decision making process: A review. *Psychology & Marketing, 27*(2), 94–116.

Demers, J. (2015). 6 social media practices that boost SEO. Retrieved January 22, 2017, from http://www.forbes.com/sites/jaysondemers/2015/01/27/6-social-media-practices-that-boost-seo/2/#d131b5564472.

Dholakia, U. M., & Durham, E. (2010). One café chain's Facebook experiment. *Harvard Business Review, 88*(3), 26.

Dixon, N. M. (2000). *Common knowledge: How companies thrive by sharing what they know*. Harvard Business School Press.

Dobele, A., Lindgreen, A., Beverland, M., Vanhamme, J., & van Wijk, R. (2007). Why pass on viral messages? Because they connect emotionally. *Business Horizons, 50*(4), 291–304.

Dobele, A., Toleman, D., & Beverland, M. (2005). Controlled infection! Spreading the brand message through viral marketing. *Business Horizons, 48*(2), 143–149.

Dolan, R. (2016). Social media: Are Facebook fans really 'engaging' with our wine brands?: A case study of Australian wine brand Facebook pages. *Wine & Viticulture Journal, 31*(1), 67.

Effing, R., & Spil, T. A. (2016). The social strategy cone: Towards a framework for evaluating social media strategies. *International Journal of Information Management, 36*(1), 1–8.

Engel, J. F., Kegerreis, R. J., & Blackwell, R. D. (1969). Word-of-mouth communication by the innovator. *Journal of Marketing, 33*(3), 15–19.

Entrepreneur.Com. (2015, May 1). Adopting a winning social strategy. Retrieved from www.entrepreneur.com

Erdoğmuş, İ. E., & Cicek, M. (2012). The impact of social media marketing on brand loyalty. *Procedia-Social and Behavioral Sciences, 58*, 1353–1360.

Feick, L. F., & Price, L. L. (1987). The market maven: A diffuser of marketplace information. *Journal of Marketing, 51*(1), 83–97.

Felix, R., Rauschnabel, P. A., & Hinsch, C. (2017). Elements of strategic social media marketing: A holistic framework. *Journal of Business Research, 70*, 118–126.

Fotis, J. N. (2015). *The Use of social media and its impacts on consumer behaviour: The context of holiday travel.* Doctoral dissertation, Bournemouth University.

Foux, G. (2006, May 8). Consumer-generated media: Get your customers involved. *Brand Strategy*, pp. 38–39.

Friedkin, N. E. (2006). *A structural theory of social influence* (Vol. 13). Cambridge University Press.

Gallaugher, J., & Ransbotham, S. (2010). Social media and customer dialog management at Starbucks. *MIS Quarterly Executive, 9*(4), 197–212.

Ghose, A., Ipeirotis, P. G., & Li, B. (2014). Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science, 60*(7), 1632–1654.

Godes, D., & Mayzlin, D. (2004). Using online conversations to study word-of-mouth communication. *Marketing Science, 23*(4), 545–560.

Goldenberg, J., Libai, B., & Muller, E. (2001). Talk of the network: A complex systems look at the underlying process of word-of-mouth. *Marketing Letters, 12*(3), 211–223.

Hackworth, B. A., & Kunz, M. B. (2011). Health care and social media: Building relationships via social networks. *Academy of Health Care Management Journal, 7*(2), 1.

Heller Baird, C., & Parasnis, G. (2011). From social media to social customer relationship management. *Strategy & Leadership, 39*(5), 30–37.

Hennig-Thurau, T., Gwinner, K. P., Walsh, G., & Gremler, D. D. (2004). Electronic word-of-mouth via consumer-opinion platforms: What motivates consumers to articulate themselves on the internet? *Journal of Interactive Marketing, 18*(1), 38–52.

Hudson, S., & Thal, K. (2013). The impact of social media on the consumer decision process: Implications for tourism marketing. *Journal of Travel & Tourism Marketing, 30*(1–2), 156–160.

Internet Live Stats. (2019, July 18). Internet usage and social media statistics. Retrieved July 18, 2019, from https://www.internetlivestats.com.

Kaplan, A. M., & Haenlein, M. (2011). Two hearts in three-quarter time: How to waltz the social media/viral marketing dance. *Business Horizons, 54*(3), 253–263.

Katz, E., & Lazarsfeld, P. F. (1955). *Personal influence: The part played by people in the flow of mass communications.* Glencoe, IL: The Free Press.

Kaushik, R. (2012, March). Impact of social media on marketing. *IJCEM International Journal of Computational Engineering & Management, 15*(2), 2230–7893.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons, 54*, 241–251.

Kirtiş, A. K., & Karahan, F. (2011). To be or not to be in social media arena as the most cost-efficient marketing strategy after the global recession. *Procedia-Social and Behavioral Sciences, 24*, 260–268.

Kosara, R., & Mackinlay, J. (2013). Storytelling: The next step for visualization. *Computer, 46*(5), 44–50.

Kumar, V., Choi, J. B., & Greene, M. (2016). Synergistic effects of social media and traditional marketing on brand sales: Capturing the time-varying effects. *Journal of the Academy of Marketing Science, 45*(2), 1–21.

Leung, L. (2013). Generational differences in content generation in social media: The roles of the gratifications sought and of narcissism. *Computers in Human Behavior, 29*(3), 997–1006.

Lipsman, A., Mudd, G., Rich, M., & Bruich, S. (2012). The power of 'like'. How brands reach (and influence) fans through social-media marketing. *Journal of Advertising Research, 52*(1), 40–52.

Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. *Business Horizons, 52*(4), 357–365.

Moscovici, S. (1980). Toward a theory of conversion behavior. *Advances in Experimental Social Psychology, 13*, 209–239.

Neti, S. (2011). Social media and its role in marketing. *International Journal of Enterprise Computing and Business Systems, 1*(2), 1–15.

Onishi, H., & Manchanda, P. (2012). Marketing activity, blogging and sales. *International Journal of Research in Marketing, 29*(3), 221–234.

Pan, B., & Crotts, J. C. (2012). Theoretical models of social media, marketing implications, and future research directions. In *Social media in travel, tourism and hospitality: Theory, practice and cases* (pp. 73–85). Surrey, UK: Ashgate.

Pfeffer, J., Zorbach, T., & Carley, K. M. (2014). Understanding online firestorms: Negative word-of-mouth dynamics in social media networks. *Journal of Marketing Communications, 20*(1–2), 117–128.

Rodríguez-Ardura, I., & Meseguer-Artola, A. (2010). Toward a longitudinal model of e-commerce: Environmental, technological, and organizational drivers of B2C adoption. *The Information Society, 26*(3), 209–227.

Rolls, K., Hansen, M., Jackson, D., & Elliott, D. (2016). How health care professionals use social media to create virtual communities: An integrative review. *Journal of Medical Internet Research, 18*(6), e166.

Sashi, C. M. (2012). Customer engagement, buyer-seller relationships, and social media. *Management Decision, 50*(2), 253–272.

Selina, D., & Milz, T. (2009). Social media will be a driving force for relationship development. *Credit Union Journal, 13*(32), 16.

Sharma, A., & Sheth, J. N. (2004). Web-based marketing: The coming revolution in marketing thought and strategy. *Journal of Business Research, 57*(7), 696–702.

Stephen, A. T., & Galak, J. (2012). The effects of traditional and social earned media on sales: A study of a microlending marketplace. *Journal of Marketing Research, 49*(5), 624–639.

Thorson, K. S., & Rodgers, S. (2006). Relationships between blogs as eWOM and interactivity, perceived interactivity, and parasocial interaction. *Journal of Interactive Advertising, 6*(2), 5–44.

Trusov, M., Bucklin, R., & Pauwels, K. (2009). Effects of word-of-mouth versus traditional marketing: Findings from an internet social networking site. *Journal of Marketing, 73*(5), 90–102.

Tsimonis, G., & Dimitriadis, S. (2014). Brand strategies in social media. *Marketing Intelligence & Planning, 32*(3), 328–344.

Tuten, T. L., & Solomon, M. R. (2014). *Social media marketing*. Sage.

Urban, G. (2003). Customer advocacy: Is it for you? *Center for Ebusiness at MIT* [Online]. Retrieved March 25, 2011, from http://ebusiness.mit.edu/research/papers/175_Urban_Trust.pdf

Van den Bulte, C., & Lilien, G. L. (2001). Medical innovation revisited: Social contagion versus marketing effort. *American Journal of Sociology, 106*(5), 1409–1435.

van Raaij, W. F. (1998). Interactive communication: Consumer power and initiative. *Journal of Marketing Communications, 4*(1), 1–8.

Villanueva, J., Yoo, S., & Hanssens, D. M. (2008). The impact of marketing induced versus word-of-mouth customer acquisition on customer equity growth. *Journal of Marketing Research, 45*(1), 48–59.

Vollmer, C., & Precourt, G. (2008). *Always on: Advertising, marketing, and media in an era of consumer control*. New York: McGraw-Hill.

Weber, L. (2009). *Marketing to the social web: How digital customer communities build your business*. John Wiley & Sons.

CHAPTER 5

Social Media Technologies and Export Marketing

Frederick Avornyo, Mahmoud Abdulai Mahmoud, Ogechi Adeola, Robert Ebo Hinson, and Richard Boateng

Introduction

Social media marketing is a major enabler of small- and medium-scale enterprises (SMEs) venturing into the international markets (Berthon, Pitt, Plangger, & Shapiro, 2012; Okazaki & Taylor, 2013). Social media, as a resource to drive marketing, enables small firms with limited resources to improve their international communication, be more efficient with market transactions with customers abroad, achieve foreign customer satisfaction and loyalty, and generally develop international network relationships (Alarcón-del-Amo, Rialp, & Rialp, 2015). These qualities, defined as international social media competencies by Alarcón-del-Amo et al. (2015), lead to the rapid internationalisation of SMEs through exports

F. Avornyo (✉) • M. A. Mahmoud • R. E. Hinson • R. Boateng
University of Ghana Business School, Accra, Ghana
e-mail: mamahmoud@ug.edu.gh; richboateng@ug.edu.gh

O. Adeola
Pan-Atlantic University, Lagos, Nigeria
e-mail: oadeola@lbs.edu.ng

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_5

(Alarcón-del-Amo, Rialp, & Rialp, 2017). SMEs generally struggle to internationalise due to inadequate resources (Paul, Parthasarathy, & Gupta, 2017). A key challenge is the ability to communicate effectively with foreign customers (Paul et al., 2017), but this appears to be changing with the advent of social media, justifying the position by scholars that social media is changing the conduct of business, and internet technology, in general, has reduced the cost of international business (Galati, Crescimanno, Tinervian, & Fagnani, 2017; Kaplan & Haenlein, 2010; Paniagua, Korzynski, & Mas-Tur, 2016; Sinkovics, Sinkovics, & Jean, 2013). The changes in the conduct of business triggered by social media, therefore, require that more scholarly attention should be directed towards the phenomenon.

The ability of the internet technology to facilitate export has been attested to and gained much scholarly attention (Gregory, Ngo, & Karavdic, 2017; Hinson, 2010; Sinkovics et al., 2013). However, high cost and reliability of internet technology (emails and websites) have been found to be impairing its successful adoption by SMEs in developing countries (Boateng, Heeks, Molla, & Hinson, 2011; Hinson, 2010). This narrative could also be changing with the emergence of social media, the revolutionary internet-based technology, which is offering firms new ways of interacting with their stakeholders (Alarcón-del-Amo et al., 2015; Kaplan & Haenlein, 2010). Importantly, social media is cheaper and easier to use, provides global reach, breaks time and distance barriers, makes communication immediate and engenders networking between the firm and the foreign consumer (Alarcón-del-Amo, Rialp, & Rialp, 2016; Berthon et al., 2012; Genc & Oksuz, 2015; Odoom, Anning-Dorson, & Acheampong, 2017; Okazaki & Taylor, 2013). The ability to access social media via the mobile phone, which has high penetration rate, is considered an added advantage, making it an important and easy-to-use resource (Packer, 2011) for owner-managers of SMEs.

The impact of social media resource on exports is indisputable (Alarcón-del-Amo et al., 2017), as affirmed by recent findings by a few scholars. The global logistics company UPS also confirms that 'social media is having a profound effect on exporters, and in particular, on small businesses, that often find their goods in demand from overseas customers long before they have thought about breaking into international markets' (compass.ups.com, n.d.). This claim by UPS is consistent with international marketing literature which suggests that social media is enabling

otherwise cash-strapped firms or SMEs, which could not afford physical presence on the international market, to be in a position to have their products on the international market (Alarcón-del-Amo et al., 2017).

However, some authors argue that firms' mere presence on social media cannot be the reason for the rapid internationalisation of SMEs (Alarcón-del-Amo et al., 2017), rather success is achieved when firms are able to align the social media resource and other complementary resources, to the needs of consumers in foreign markets (Seggie & Griffith, 2008). This assertion is further buttressed by the argument that technology on its own cannot lead to performance, while firm resources alone can neither produce significant gains in firm performance (Bharadwaj, 2000; Trainor, Andzulis, Rapp, & Agnihotri, 2014). Scholars posit that firm performance is possible when technological resources such as social media is combined with firm's strategic resources to achieve distinctive capabilities (Alarcón-del-Amo et al., 2017; Trainor et al., 2014). This is consistent with the assertion by the Resource-Advantage (R-A) theorists, who argue that firms combine resources, both tangible and intangible, to achieve superior performance (Hunt & Arnett, 2003; Hunt, & Morgan, 1995).

Scholars have empirically verified that social media utilisation leads to firm performance (Luo, Zhang, & Duan, 2013; Paniagua & Sapena, 2014) and export performance specifically (Alarcón-del-Amo et al., 2017). However, the issue as to what resource combination SMEs on social media are aligning to the needs of foreign consumers, to achieve this export performance, appears not to have been adequately addressed by extant literature (Alarcón-del-Amo et al., 2017). Social media is considered nascent (Killian & McManus, 2015) and social media application to international marketing is still considered emerging (Okazaki & Taylor, 2013), and therefore not surprising that research into social media marketing and export is scanty (Paniagua et al., 2016). This chapter, therefore, seeks to understand which resources firms are integrating with social media to achieve superior performance. The study is linked to the central focus of the resource-advantage theory (R-A) theory and is, therefore, premised on the approach. The rest of the chapter looks at the theoretical foundation of the study, a review of the literature on social media marketing, a discussion on social media export marketing capabilities, social media and export performance measurement, proposes a framework and presents its conclusions.

Theoretical Foundation

Social media is a resource (Paniagua et al., 2016; Wang & Kim 2017) which firms efficiently and effectively combine with other resources to achieve superior performance (Alarcón-del-Amo et al., 2016; Wang & Kim, 2017). This assertion is in sync with the R-A theory that firms are combiners of dynamic and heterogeneous resources to create a comparative advantage leading to competitive advantage and superior financial performance (Hu & Wang, 2009; Hunt, 2011; Hunt & Morgan, 1995; Thoeni, Marshall, & Campbell, 2016). The theory places emphasis on innovation, market segments, heterogeneous firm resources, a comparative market advantage/disadvantage in resources and a marketplace position of competitive advantage/disadvantage (Hunt, 2001, 2012). R-A theory enumerates seven resources which it argues are available to firms. They include financial, physical, legal, human, organisational, informational and relational (Barua & Michie, 2014; Hunt 2001, 2012). The theory posits that firms combine both tangible and intangible resources to deliver value for a segment of the market which is akin to what exporting SMEs present on social media do. They combine social media marketing with other firm resources to deliver value to customers in the international segment of the market.

R-A theory was propounded in 1995 by Shelby Hunt and Robert Morgan in an article titled 'The Comparative Advantage Theory of Competition'. It is an interdisciplinary and evolutionary general theory of competition that evolved from ten different research traditions (Seggie & Griffith, 2008) and has 12 foundational premises and a structure (Hunt & Arnett, 2003). Hunt and Arnett posit that R-A is

> an evolutionary, disequilibrium-provoking, process theory of competition, in which innovation and organisational learning are endogenous, firms and consumers have imperfect information, and in which entrepreneurship, institutions, and public policy affect economic performance.

This suggests that firms will avail themselves of innovations (such as social media), and as they combine resources available to them to achieve superior performance (such as in export), they will learn internally (since organisational learning is endogenous) which resources would deliver better outcomes, so they can determine, in which resources more investment commitment should be made (Rest & Roper, 2013).

The theory also argues that firms combine an assortment of resources available to them in order to have a comparative advantage in resources

(Hunt 2012; Hunt and Morgan, 1995, 1997). For firms, this will result in a marketplace position of competitive advantage and hence superior performance (Hunt, 2012; Hunt & Arnett, 2003). On the other hand, a comparative disadvantage in resources will lead to a marketplace position of competitive disadvantage and inferior performance, with a third position being where there is parity (Hunt, 2011). It, therefore, becomes relevant for firms to appreciate what resource combinations deliver comparative advantage and ultimately superior performance and which ones deliver the opposite outcomes.

Since its seminal paper in 1995, several other researchers have contributed to the theory by using it as basis to explain various phenomena including resource matching and globalisation of services (Seggie & Griffith, 2008), to demonstrate that Corporate Social Responsibility is core to firm operations (Hu & Wang, 2009), while others have extended the theory into firm's pricing strategy (van der Rest & Roper, 2013). However, a review of the extant literature on R-A theory did not reveal its usage to explain the social media marketing and export performance phenomenon. This is in spite of the fact that R-A theory is considered a marketing strategy theory, while social media marketing has become an integral part of marketing strategy for firms (Akar & Topçu, 2011; Barua & Michie, 2014; Schlegelmilch, 2011). Additionally, the list of seven resources as provided for by the R-A theory cannot be exhaustive as marketing practice is dynamic (Grönroos, 1994). The list of resources could, therefore, be extended. The chapter also argues that not all the seven resources can be of equal importance to a firm. The Competitive Position Matrix of R-A theory demonstrates that some resources will lead to marketplace position of competitive advantage, while others may not (Hunt, 2001). On that basis, it is argued that some resources may be of higher importance than others. The most relevant resources can be classified as core and the rest secondary. This qualification of the resources is considered relevant, as the complementarity or otherwise of the resources could influence how they are converted into capabilities and eventually to superior performance (Trainor et al., 2014).

Literature Review

Social media marketing is one of the contemporary marketing concepts, and business practitioners are seeking to understand how to apply it effectively in order to achieve superior performance (Packer, 2011). It has unique features which could explain the special traction it has received in

recent times in preference to the older internet-based technologies (Elena, 2016; Mangold & Faulds, 2009). These features include the easy and quick unfettered access to the global market, content co-creation, real-time interactivity and engagement between the firm and the consumer as well as among consumers (Martin, Goncalves, Oliveira, Cota, & Branco, 2016; Wang & Kim, 2017). Social media marketing emerged following the advent of social media platforms such as Facebook, Twitter, Instagram and several other forms such as Pinterest, YouTube, WhatsApp, LinkedIn, Google+, Snapchat (Paniagua et al., 2016; Socialpilot.co, n.d.). It consists of text, pictures, videos and networks (Berthon et al., 2012). Attempts have been made to categorise the various social media platforms into

1. collaborative projects such as Wikipedia,
2. blogs and user-generated content communities such as YouTube,
3. social networking sites such as Facebook,
4. microblogs such as Twitter,
5. virtual game worlds such as EverQuest and
6. virtual social worlds such as Second Life (Berthon et al., 2012; Kaplan & Haenlein, 2010; Okazaki & Taylor, 2013).

Social media is a Web 2.0 application which allows for two-way and many-to-many ways communication (Ananda, Hernández-García, & Lamberti, 2016). Scholars have explained Web 2.0 to be the interactive component of Web 1.0 (Yadav & Rahman, 2017), as Web 1.0 is only a one-way application such as websites (Constantinides, 2014). One of the scholars cited as a pioneer, Tim O'Reilly (Berthon et al., 2012) sees social media in the light of Web 2.0 and defines it in a technical sense as 'a set of principles and practices that tie together a veritable solar system of sites that demonstrate some or all of those principles, at a varying distance from that core' (O'Reilly, 2007). Subsequently, Kaplan and Haenlein (2010) simplify it as an internet-based application that allows for the creation and exchange of user-generated content, but they also draw the linkage with Web 2.0.

Though the definitions found in extant literature are varied, there are commonalities which include the fact that social media is an internet-based technology (Kaplan & Haenlein, 2010), modelled on Web 2.0 (Constantinides, 2014; Wang & Kim, 2017) that allows users to interact, share information and co-create content (Berthon et al. 2012). Based on these various definitions, and others in extant literature, this chapter

conceptualises social media as a social Web 2.0 platform that allows for real-time interactivity and content co-creation among users, in the form of text, picture, video, audio and network (Berthon et al., 2012; Constantinides, 2014; Kaplan & Haenlein, 2010).

Social media marketing on the other hand involves the use of social media technology for firms' promotional activities (Akar & Topçu, 2011; Kaplan & Haenlein, 2010) to achieve a firm's goal by creating value for its various stakeholders (Felix, Rauschnabel, & Hinsch, 2017; Tuten & Solomon, 2015) and for relationship building (Chang et al., 2015). Based on these cited expert views (i.e. Chang et al., 2015; Felix et al., 2017), we define social media marketing as the use of social media for marketing activities, including research, promotion, distribution and engagement with and among stakeholders, for value creation.

Thematic Areas

We reviewed 133 articles to address various issues which were synthesised into ten thematic areas for the purpose of this chapter. A significant number of articles (27.8%) examined factors that influence social media adoption in order to determine how firms can be successful with its application (Chikandiwa, Contogiannis, & Jembere, 2013; Galati et al., 2017). Some examined social media and branding (Bruhn, Schoenmueller, & Schäfer, 2012; Nguyen, Yu, Melewar, & Chen, 2015), and how social media could predict consumer behaviour (Godey et al., 2016; Kim & Drumwright, 2016). Other scholars delved into how social media influences electronic word of mouth and its impact on the operations of firms (Erkan & Evans, 2016; Teng, Khong, Chong, & Lin, 2017). Other thematic areas identified include social customer relationship management (Malthouse, Haenlein, Skiera, Wege, & Zhang, 2013; Trainor et al., 2014; Wang & Kim, 2017), social media and marketing communications (Mangold & Faulds, 2009; Wang, Pauleen, & Zhang, 2016), and social media marketing strategy (Ananda et al., 2016; Felix et al., 2017). There are also articles that addressed how social media impact firm performance (Luo et al., 2013; Moro, Rita, & Vala, 2016), social media challenges (Kaplan & Haenlein, 2010; Schultz & Peltier, 2013) and finally social media and internationalisation (Alarcón-del-Amo et al., 2017; Berthon et al., 2012; Paniagua et al., 2016).

Three thematic areas namely social media adoption, social media and branding as well as social media and consumer behaviour jointly accounted

for 62.4% of the literature examined, while social media marketing strategy, social media challenges and social media and internationalisation, which included export, were among the least researched areas. This finding is consistent with the suggestion by Paniagua et al. (2016) and Alarcón-del-Amo et al. (2017) that the theme, social media and export is under-researched and requires more scientific research attention.

Social Media Adoption and Exporting SMES

Businesses have adopted social media to enhance their marketing operations. Some have achieved success (Felix et al., 2017) others have not witnessed the corresponding improvement in firm performance (Zhu & Chen, 2015). These two outcomes have engendered scholars to research into factors that influence a successful adoption of social media (Martin et al. 2016; Packer, 2011), and the various adoption models and how they impact success (Chikandiwa et al., 2013; Felix et al., 2017).

Chikandiwa et al. (2013) identify two social media adoption models. The first is based on how social media is managed within the organisation; whether it is decentralised or not. They list centralised, distributed, 'coordinated hub and spoke', as well as 'multiple hub and spoke', as forms of adoption frameworks. According to the authors, the 'coordinated hub and spoke' is the adoption model where tactical decision considering a single higher department takes the adoption and usage of social media, but the implementation is decentralised at the various departments within the firm. In the case of 'multiple hub and spoke', again the tactical decision is a top management action, but the implementation is further decentralised beyond departments to smaller units such as product lines. However, in the case of the centralised model, both tactical decision and implementation is the sole responsibility of a dedicated department or unit, and the reverse holds for the distributed model. The second category of social media adoption model by Chikandiwa et al. (2013) examines the various social media adoption stages. They include the social media strategy learning curve, social media adoption curve, social engagement journey and social media maturity model. The authors then explicate the intricacies of each stage and how firms behave at those stages of maturity in the adoption of social media.

Felix et al. (2017) also propose a four-dimensional model which looks at the scope, the culture, the structure and the governance. The scope is concerned with whether firms use social media to engage a few stakehold-

ers or it is an imbibed and holistic philosophy that ensures that both employees and external stakeholders are engaged using social media. In terms of culture, the scholars seek to address the question whether a firm adopts a conservative or modern approach towards the use of social media. Structure, on the other hand, is similar to Chikandiwa et al.'s (2013) first category, which looks at the centralisation or otherwise. Felix et al. (2017) however use the term hierarchies to denote centralisation, and networks to represent a decentralised system of social media usage in a firm. The last dimension is governance. Where there are strict regulations as to who is permitted to use social media within the firm, it is referred to as autocracy, but where there is no strict compliance, it is termed anarchy. Felix et al. (2017) however suggest that in adopting social media, firms should be flexible and have a blend of the dimensions and also use influencers within the firm to lead the social media adoption model.

It should be noted that the studies conducted by Chikandiwa et al. (2013) and Felix et al. (2017) are not specific to SMEs, particularly exporting SMEs. SME-biased researches, however, are extremely important because they dominate most economies (Opoku, Abratt, Bendixen, & Pitt, 2007; Stoian, Rialp, & Rialp, 2011) and more importantly exhibit unique characteristics such as the dominance of owner-managers, inadequate resources to have a full-fledged organisational set up, but quick and flexible in decision making (Lu & Julian, 2008; Mehrtens, Cragg, & Mills, 2001; Paul et al., 2017). This chapter posits that the decision as to which social media adoption model to opt for is constrained by the social media competence of the owner-manager, the strength and quality of the human resource available to the firm, and financial resources to outsource the role of managing the firm's social media marketing implementation. A study conducted on SMEs' social media adoption model, should, therefore, bring out these perspectives unique to SMEs. This chapter, therefore, takes the position that the unique characteristics of SMEs would influence their social media adoption model. A resource-strapped but social media savvy owner-manager of an exporting SME, for instance, may also be constrained by time availability, to continually update its page, respond and regularly engage with online customers who may be of different time zones.

The argument that SMEs could have different social media adoption model is supported by Genc and Oksuz (2015) who clearly demonstrate the large-scale and small-scale business dichotomy by indicating that small businesses have a more liberal approach to social media compared to their

large-scale counterparts. Where there is no difference between large and smaller firms is the compelling reason for social media adoption which is the attainment of superior performance (Genc & Oksuz, 2015; Martin et al., 2016), and this can be achieved if firms succeed in effectively and efficiently combining resources available to them. The mere presence of exporting SMEs on any of the social media platforms would not guarantee them superior performance but rather their ability to combine it with the right set of resources (Bharadwaj, 2000; Gregory et al., 2017; Hunt & Morgan, 1997; Trainor et al., 2014). This assertion makes the ability to identify the set of resources that integrates well with social media for exporting SMEs an important scientific enquiry.

Social Media Export Marketing Capabilities

Marketing capabilities have been studied by previous researchers (Barney, 1991; Day, 1994; Hunt, 2011) and also the issue of technology capabilities (Bharadwaj, Bharadwaj, & Konsynski, 1999; Rasiah & Myint, 2013; Wignaraja, 2002). Capabilities have been defined as 'an organisation's ability to assemble, integrate and deplore resources in combination to achieve a competitive advantage' (Trainor et al., 2014). In social media marketing, Alarcón-del-Amo et al. (2016) define social media capabilities 'as an integrated set of strategic resources and capabilities that can create, through the use of social media applications, sound competitive advantages and superior firm performance based upon a more effective information management'. This is distinct from 'an integrated set of technological resources' which is termed as social media competence (Alarcón-del-Amo et al., 2015). Some scholars have tried to explain social media internationalisation capabilities. For instance, Paniagua et al. (2016) list identity, conversation, sharing, presence, relationship, reputation and group as social media resources. They, therefore, argue that converting these social media resources into capabilities that enable a firm to internationalise will constitute social media internationalisation capabilities. The assertion, however, does not explain how the social media resource is converted into capabilities.

Based on various explanations, this chapter attempts to define social media export marketing capabilities as a firm's ability to effectively and efficiently combine social media marketing resources with key strategic firm resources to achieve superior export performance (Alarcón-del-Amo

et al., 2017; Hunt & Morgan, 1997; Trainor et al., 2014). The more significant concern though is the ability to determine which firm resources combined with social media marketing can create the export capabilities. This chapter agrees with R-A theorists that resources available to firms are financial, physical, legal, human, organisational, informational and relational. It however opines that since SMEs are dominated by owner-managers (Paul et al., 2017) and are constrained financially, the owner-manager who is social media savvy is likely to be the key or even sole person updating the social media platforms, responding to customer enquiries and regularly engaging with customers, some of whom may be living in places with different time zones. As a result, an own-manager will require enough time to effectively and efficiently utilise the social media resource. Time availability, therefore, becomes an important factor in the operations of exporting SMEs. Scholars such as Klein (2007) have suggested that though some researchers consider time as a commodity, time also exhibits features that qualify it as a resource. In corroboration with the argument by Klein (2007), this chapter considers time as a resource that exporting SMEs present on social media require. Premised on the R-A theory, it is advocated that exporting SMEs would need to effectively and efficiently combine social media marketing with financial, physical, organisational, legal, human, relational, informational (market intelligence) and time to develop capabilities for export performance. This study, therefore, extends the R-A theory by proposing time as an 8th resource to the existing seven. It further argues that resources could produce different outcomes as held by the R-A competitive position matrix (Hunt, 2001; Hunt & Arnett, 2003). Some will generate a marketplace position of competitive advantage and others competitive disadvantage (Hunt, 2012). It also postulates that some resources may be more complementary compared to others. The eight resources can, therefore, be categorised into core and secondary resources. These set of firm resources can, therefore, be empirically verified to determine which ones can be considered as core, secondary or irrelevant to exporting SMEs.

Social Media and Export Relationship Performance

Export performance has been measured using financial and accounting ratios such as return on investment, return on equity and the various profitability ratios (Lages & Lages, 2004) or through the use of other absolute values such as export sales volume or export market share (Beleska-

Spasova, 2014). It has also been measured using non-financial performance indicators which are considered subjective but relevant in determining export performance (Morgan, Vorhies, & Mason, 2009). The inclusion of non-financial export performance measurements has been largely necessitated by the difficulty in gathering financial data on SMEs in particular (Ibeh & Wheeler, 2005). However social media is first of all a social online technology (Schultz & Peltier, 2013) and thrives on networking and relationship which are also essential ingredients for SME success (Felix et al., 2017; Kaplan & Haenlein, 2010; Kietzmann, Hermkens, McCarthy, & Silvestre, 2011; Nguyen et al., 2015; Trainor et al., 2014). This chapter, therefore, argues that the measurement of export performance of SMEs on social media platforms cannot ignore the firm's success at maintaining relationships with foreign customers, as one of the performance indicators. An export performance tripod is proposed with the third leg categorised relationship performance.

Although export relationship performance can be categorised as a non-financial measurement, the centrality of networking and relationship to social media marketing, requires that export relationship should be treated as the third leg of export performance measurement. Trainor et al. (2014) conceptualised relationship performance within the social media context, developed a scale and tested it. This study adopts this and terms it 'export relationship performance'. It is defined in this chapter as the level of relationship between exporting firms and their foreign customers (Trainor et al., 2014). Adapting from Trainor et al. (2014), export relationship performance measures

1. how long foreign customers have worked with firms,
2. how long new foreign customers tend to stay,
3. how loyal they perceive foreign customers to be,
4. how satisfied they perceive foreign customers to be and
5. how important customer retention is to the firm.

These five measurable variables will help to determine how successfully a firm on social media has effectively utilised the social media technology to engage customers in the export market segment. Introducing the third measurement, which has not received adequate attention in the area of export performance, is considered an important contribution to the literature on export performance.

Conceptual Framework

Following from the discourse, it is argued that SMEs can achieve export performance because they are successful at efficiently and effectively integrating the social media (platforms) resource with complementary strategic resources and aligning them to the interest of the foreign customer (Seggie & Griffith, 2008; Trainor et al., 2014). This study, therefore, proposes a framework that encapsulates the various social media platforms that exporting SMEs are leveraging on and the resources they combine to achieve export performance which includes relationship performance as demonstrated in Fig. 5.1.

Notably, WhatsApp, Facebook, YouTube, Facebook Messenger and Instagram are amongst the topmost social media platforms in the world (Genc & Oksuz, 2015). However, since Facebook messenger is an application on Facebook, this research merges it with Facebook as one platform. The study then adopts the seven resources listed by R-A theory: financial, physical, legal, human, organisational, informational and relational (Hunt, 2012), and includes 'time' as the eighth resource. The chapter posits that exporting SMEs should integrate these resources with social media, effectively and efficiently to achieve financial, non-financial and export relationship performances (Fig. 5.2).

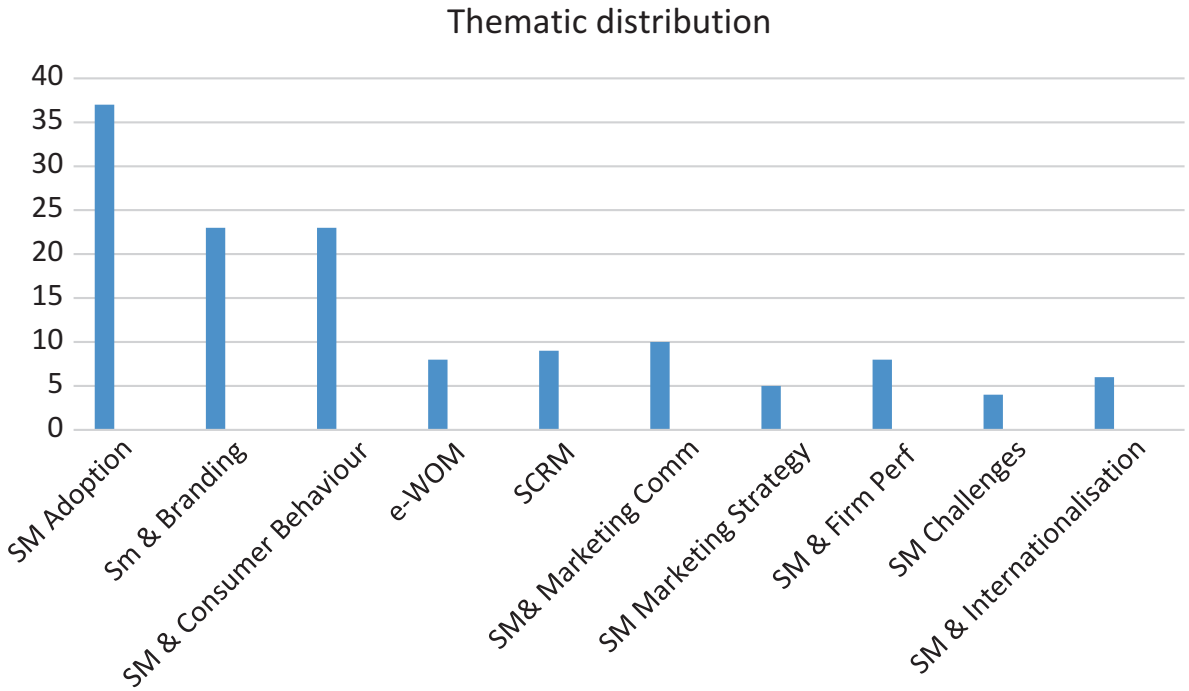


Fig. 5.1 Thematic distribution of included literature

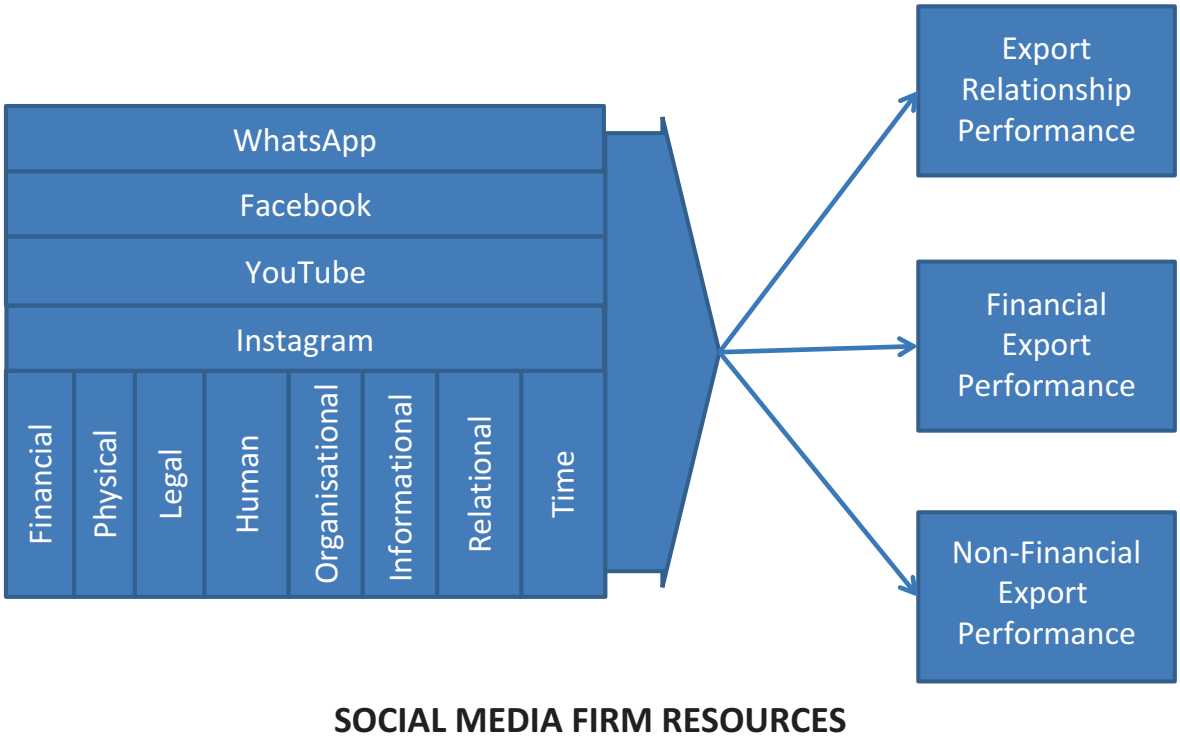


Fig. 5.2 Social media export marketing performance capabilities framework

Conclusion

Social media, a relatively cheap resource available to otherwise cash-strapped SMEs (Alarcón-del-Amo et al., 2016) is creating a new era for businesses, with the competency to improve international communication, efficiency in the international market transaction, loyalty with foreign customers and development of international network (Alarcón-del-Amo et al., 2015). It facilitates firm internationalisation by enabling firms to reach out to customers in other countries, while avoiding the cost and hurdles of physically setting up in the foreign market (Berthon et al., 2012; Okazaki & Taylor, 2013). However, research on social media marketing and export is scanty (Paniagua et al., 2016). Although few studies in the area have affirmed the positive relationship between social media marketing and export performance, there is limited focus on the combination of firm resources that enables performance.

This chapter therefore posits that since firms combine resources (Hunt, 2012), and firm performance is enhanced by the conversion of combined resources into capabilities (Alarcón-del-Amo et al., 2016; Bharadwaj, 2000; Bharadwaj et al., 1999; Trainor et al., 2014), exporting SMEs are

able to achieve export performance by efficiently and effectively combining social media marketing with financial, physical, legal, organisational, human, relational and informational, and the new addition, time, resource. The chapter further suggests that firm resources are not of equal importance and therefore advocates a further qualification of the resources into core and secondary. The study also posits that to investigate social media marketing and export, export performance should not be assessed solely based on the traditional financial and non-financial indicators but also on an additional factor, export relationship performance. However, the arguments made can only be generalised if they are empirically validated. Further research is therefore proposed to test the assertions made.

References

Akar, E., & Topçu, B. (2011). An examination of the factors influencing consumers' attitudes toward social media marketing. *Journal of Internet Commerce, 10*(1), 35–67.

Alarcón-del-Amo, M. C., Rialp, A., & Rialp, J. (2015). The effect of social media adoption on exporting firms' performance. *Advances in International Marketing, 25*, 161–186.

Alarcón-del-Amo, M. C., Rialp, A., & Rialp, J. (2016). Social media adoption by exporters: The export-dependence moderating role. *Spanish Journal of Marketing—ESIC, 20*, 81–92.

Alarcón-del-Amo, M. C., Rialp, A., & Rialp, J. (2017). Examining the impact of managerial involvement with social media on Exporting firm performance. *International Business Review*. https://doi.org/10.1016/j.ibusrev.2017.09.003

Ananda, A. S., Hernández-García, A., & Lamberti, L. (2016). N-REL: A comprehensive framework of social media marketing strategic actions for marketing organizations. *Journal of Innovation & Knowledge, 1*(3), 170–180.

Barney, J. (1991). Firm resources and sustained competitive advantage. *Journal of Management, 17*(1), 99–120.

Barua, M. E., & Michie, D. A. (2014, November 4–8). The role of opportunism within the context of resource advantage theory: A conceptual framework. In Meter, R. V., & Weiser, J. (Eds.), *Advances in marketing transformational marketing: Proceedings of 2014 Society for Marketing Advances, New Orleans.* Houston: Society for Marketing Advances. Retrieved from https://www.societyformarketingadvances.org/resources/Documents/Resources/Conference%20Proceedings/2014_SMA_Proceedings.pdf#page=412.

Beleska-Spasova, E. (2014). Determinants and measures of export performance—Comprehensive literature review. *JCEBI, 1*(1), 63–74.

Berthon, P. R., Pitt, L. F., Plangger, K., & Shapiro, D. (2012). Marketing meets Web 2.0, social media, and creative consumers: Implications for international marketing strategy. *Business Horizons, 55*, 261–271.

Bharadwaj, A. S. (2000). A resource-based perspective on information technology capability and firm performance: An empirical investigation. *MIS Quarterly, 24*(1), 169–196.

Bharadwaj, A., Bharadwaj, S., & Konsynski, B. (1999). Information technology effects on firm performance as measured by Tobin's q. *Management Science, 45*(7), 1008–1024.

Boateng, R., Heeks, R., Molla, A., & Hinson, R. (2011). Advancing e-commerce beyond readiness in a developing country: Experiences of Ghanaian firms. *Journal of Electronic Commerce in Organizations, 9*(1), 1–16.

Bruhn, M., Schoenmueller, V., & Schäfer, D. B. (2012). Are social media replacing traditional media in terms of brand equity creation? *Management Research Review, 35*(9), 770–790.

Chikandiwa, S. T., Contogiannis, E., & Jembere, E. (2013). The adoption of social media marketing in South African banks. *European Business Review, 25*(4), 365–381.

Chang, Y. T., Yu, H., & Lu, H. P. (2015). Persuasive messages, popularity cohesion, and message diffusion in social media marketing. *Journal of Business Research, 68*(4), 777–782.

Compass-UPS. (n.d.). Retrieved from https://compass.ups.com/how-social-media-can-boost-exports/

Constantinides, E. (2014). Foundations of social media marketing. *Procedia—Social and Behavioral Sciences, 14*, 40–57.

Day, G. S. (1994). The capabilities of market-driven organizations. *Journal of Marketing, 58*(4), 37–52.

Elena, C. A. (2016). Social media—A strategy in developing customer relationship management. *Procedia Economics and Finance, 39*, 785–790.

Erkan, I., & Evans, C. (2016). The influence of eWOM in social media on consumers' purchase intentions: An extended approach to information adoption. *Computers in Human Behavior, 61*, 47–55.

Felix, R., Rauschnabel, P. A., & Hinsch, C. (2017). Elements of strategic social media marketing: A holistic framework. *Journal of Business Research, 70*, 118–126.

Galati, A., Crescimanno, M., Tinervian, S., & Fagnani, F. (2017). Social media as a strategic marketing tool in the Sicilian wine industry: Evidence from Facebook. *Wine Economics and Policy, 6*, 40–47.

Genc, M., & Oksuz, B. (2015). A fact or an illusion: Effective social media usage of female entrepreneurs. *Procedia—Social and Behavioral Sciences, 195*, 293–300.

Godey, B., Manthiou, A., Pederzoli, D., Rokka, J., Aiello, G., Donvito, R., & Singh, R. (2016). Social media marketing efforts of luxury brands: Influence on brand equity and consumer behavior. *Journal of Business Research, 69*(12), 5833–5841.

Gregory, G. D., Ngo, L. V., & Karavdic, M. (2017). Developing e-commerce marketing capabilities and efficiencies for enhanced performance in business-to-business export ventures. *Industrial Marketing Management*. https://doi.org/10.1016/j.indmarman.2017.03.002

Grönroos, C. (1994). From marketing mix to relationship marketing: Towards a paradigm shift in marketing. *Management Decision, 32*(2), 4–20.

Hinson, R. (2010). The value chain and e-business in exporting: Case studies from Ghana's non-traditional export (NTE) sector. *Telematics and Informatics, 27*, 323–340.

Hu, Y.-C., & Wang, F. C.-C. (2009). Collectivism, corporate social responsibility, and resource advantages in retailing. *Journal of Business Ethics, 86*(1), 1–13.

Hunt, S. D. (2001). Commentary—A general theory of competition: Issues, answers and an invitation. *European Journal of Marketing, 35*(5/6), 524–548.

Hunt, S. D. (2011). Developing successful theories in marketing: Insights from resource-advantage theory. *AMS Review, 1*, 72–84.

Hunt, S. D. (2012). The evolution of resource-advantage theory: Six events, six realizations, six contributions. *Journal of Historical Research in Marketing, 4*(1), 7–29.

Hunt, S. D., & Arnett, D. B. (2003). Resource-advantage theory and embeddedness: Explaining R-A theory's explanatory success. *Journal of Marketing Theory and Practice, 11*(1), 1–17.

Hunt, S. D., & Morgan, R. M. (1995). The comparative advantage theory of competition. *Journal of Marketing, 59*, 1–15.

Hunt, S. D., & Morgan, R. M. (1997). Resource-advantage theory: A snake swallowing its tail or a general theory of competition? *Journal of Marketing, 61*, 74–82.

Ibeh, K. I. N., & Wheeler, C. N. (2005). A resource-centred interpretation of export performance. *International Entrepreneurship and Management Journal, 1*, 539–556.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons, 53*(1), 59–68.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons, 54*, 241–251.

Killian, G., & McManus, K. (2015). A marketing communications approach for the digital era: Managerial guidelines for social media integration. *Business Horizons, 58*, 539–549.

Kim, E., & Drumwright, M. (2016). Engaging consumers and building relationships in social media: How social relatedness influences intrinsic vs. extrinsic consumer motivation. *Computer Human Behaviour, 63*, 970–979.

Klein, C. C. (2007). The Economics of Time as a resource. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.329.7498&rep=rep1&type=pdf.

Lages, L. F., & Lages, C. R. (2004). A measure of short-term export performance improvement. *Journal of International Marketing, 12*(1), 36–56.

Lu, V. N., & Julian, C. C. (2008). The internet, strategy and performance: A study of Australian export market ventures. *Journal of Global Marketing, 21*(3), 231–240.

Luo, X., Zhang, J., & Duan, W. (2013). Social media and firm equity value. *Information Systems Research, 24*(1), 146–163.

Malthouse, E. C., Haenlein, M., Skiera, B., Wege, E., & Zhang, M. (2013). Managing customer relationships in the social media era: Introducing the social CRM house. *Journal of Interactive Marketing, 27*(4), 270–280.

Mangold, W. G., & Faulds, D. J. (2009). Social media: The new hybrid element of the promotion mix. *Business Horizons, 52*(4), 357–365.

Martin, J., Goncalves, R., Oliveira, T., Cota, M., & Branco, F. (2016). Understanding the determinants of social network sites adoption at firm level: A mixed methodology approach. *Electronic Commerce Research and Applications, 18*, 10–26.

Mehrtens, J., Cragg, P. B., & Mills, A. M. (2001). A model of Internet adoption by SMEs. *Information & Management, 39*, 165–176.

Morgan, N. A., Vorhies, D. W., & Mason, C. H. (2009). Market orientation, marketing capabilities, and firm performance. *Strategic Management Journal, 30*, 909–920.

Moro, S., Rita, P., & Vala, B. (2016). Predicting social media performance metrics and evaluation of the impact on brand building: A data mining approach. *Journal of Business Research, 69*(9), 3341–3351.

Nguyen, B., Yu, X., Melewar, T. C., & Chen, J. (2015). Brand innovation and social media: Knowledge acquisition from social media, market orientation, and the moderating role of social media strategic capability. *Industrial Marketing Management, 51*, 11–25.

Odoom, R., Anning-Dorson, T., & Acheampong, G. (2017). Antecedents of social media usage and performance benefits in small- and medium-sized enterprises (SMEs). *Journal of Enterprise Information Management, 30*(3), 383–399.

Okazaki, S., & Taylor, C. R. (2013). Social media and international advertising: Theoretical challenges and future directions. *International Marketing Review, 30*(1), 56–71.

Opoku, R. A., Abratt, R., Bendixen, M., & Pitt, L. (2007). Communicating brand personality: Are the web sites doing the talking for food SMEs? *Qualitative Market Research: An International Journal, 10*(4), 362–374.

O'Reilly, T. (2007). *What is Web 2.0: Design patterns and business models for the next generation of software*. Retrieved from http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html.

Packer, R. (2011). Social media marketing: The art of conversational sales. Retrieved from https://wsimarketbuilders.com/socialmediamarketing-whitepaper.pdf.

Paniagua, J., Korzynski, P., & Mas-Tur, A. (2016). Crossing borders with social media: Online social networks and FDI. *European Management Journal, 35*(3), 314–326.

Paniagua, J., & Sapena, J. (2014). Business performance and social media: Love or hate? *Business Horizons, 57*, 719–728.

Paul, J., Parthasarathy, S., & Gupta, P. (2017). Exporting challenges of SMEs: A review and future research agenda. *Journal of World Business, 52*, 327–342.

Rasiah, R., & Myint, M. M. (2013). Ownership, technological capabilities and exports of garment firms in Myanmar. *Technological and Economic Development of Economy, 19*(1), S22–S42.

Rest, J.-P., & Roper, A. (2013). A resource-advantage perspective on pricing: Shifting the focus from ends to means-end in pricing research? *Journal of Strategic Marketing, 21*(6), 484–498.

Schlegelmilch, B. B. (2011). Commentary on developing successful theories in marketing: Insights from resource-advantage theory. *Academy of Marketing Science Review, 1*, 85–89.

Schultz, D. E., & Peltier, J. (2013). Social media's slippery slope: Challenges, opportunities and future research directions. *Journal of Research in Interactive Marketing, 7*(2), 86–99.

Seggie, S. H., & Griffith, D. A. (2008). The resource matching foundations of competitive advantage: An alternative perspective on the globalization of service firms. *International Marketing Review, 25*(3), 262–275.

Sinkovics, N., Sinkovics, R. R., & Jean, R. B. (2013). The internet as an alternative path to internationalization? *International Marketing Review, 30*(2), 130–155.

Socialpilot.co. (n.d.). Retrieved May 3, 2017, from https://socialpilot.co/blog/125-amazing-social-media-statistics-know-2016/.

Stoian, M. C., Rialp, A., & Rialp, J. (2011). Export performance under the microscope: A glance through Spanish lenses. *International Business Review, 20*(2), 117–135.

Teng, S., Khong, K. W., Chong, A. Y. L., & Lin, B. (2017). Persuasive electronic word-of-mouth messages in social media. *Journal of Computer. Information Systems, 57*(1), 76–88.

Thoeni, A. T., Marshall, G. W., & Campbell, S. M. (2016). A resource-advantage theory typology of strategic segmentation. *European Journal of Marketing, 50*(12), 2192–2215.

Trainor, K. J., Andzulis, J. M., Rapp, A., & Agnihotri, R. (2014). Social media technology usage and customer relationship performance: A capabilities-based examination of social CRM. *Journal of Business Research, 67*(6), 1201–1208.

Tuten, T. L., & Solomon, M. R. (2015). *Social media marketing* (2nd ed.). Sage.

Wang, W. Y. C., Pauleen, D. J., & Zhang, T. (2016). How social media applications affect B2B communication and improve business performance in SMEs. *Industrial Marketing Management, 54*, 4–14.

Wang, Z., & Kim, H. G. (2017). Can social media marketing improve customer relationship capabilities and firm performance? Dynamic capability perspective. *Journal of Interactive Marketing, 39*, 15–26.

Wignaraja, G. (2002). Firm size, technological capabilities and market-oriented policies in Mauritius. *Oxford Development Studies, 30*(1), 87–104.

Yadav, M., & Rahman, Z. (2017). Measuring consumer perception of social media marketing activities in e-commerce industry: Scale development & validation. *Telematics and Informatics, 34*, 1294–1307.

Zhu, Y-Q., & Chen, H-G. (2015). Social media and human need satisfaction: Implications for social media marketing. *Business Horizons, 58*, 335–345.

CHAPTER 6

Data Analytics and Predictive Analytics: How Technology Fits into the Equation

Brian J. Galli and Gabrielle Muniz

Introduction

Throughout the years, technology has changed considerably. In the mid-1960s, mainframe computers were state of the art in the information-processing kingdom. Although IBM has existed since 1952, behemoths were produced by the 1960s and 1970s. The term "mainframe computer" was largely synonymous with IBM because of the company's dominant market share at the time. These things were room-filling metal contraptions of vacuum tubes and wires that had to be air-cooled to avoid overheating. The cost was upwards of $3 million in 1960s dollars (Klimberg, 2016)! Also, memory was about 5121 KB, and hard drive storage was about 350 MB, while processing speed was a scorching 1 MIPS (million instructions per second). This, of course, mirrors Bill Gates's (in)famous 1981 statement: "640 KB ought to be enough for anybody."

However, there was quite a reversal by 2016 with the iPhone 6. This phone featured 32 GB (+9300%) of memory and was pocket sized. Author and futurist Ray Kurtzweil notes that cell phones are much cheaper, more

B. J. Galli (✉) • G. Muniz
Long Island University, Brookville, NY, USA
e-mail: Brian.Galli@liu.edu

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_6

powerful, and more compact than a standard computer. Compared to the computer at Massachusetts Institute of Technology (MIT) in 1965, the cell phone is a billion times more capable per dollar/euro (Klimberg, 2016). Though most corporate executives would suggest that their two most valuable corporate assets are their people (employees) and their products, a closer look at the products and their composition reveals that data is often listed as a primary component. That would include customer data and the actual product data required for creating output. Many businesses have now come to embrace this concept and have built analytics or intelligence units to better understand and utilize the data within their control (Buchanan, 2017).

According to IBM, there are many advanced analytics capabilities that come from predictive analytics, for example ad hoc statistical analysis, real-time scoring, predictive modeling, text analytics, and data mining, to name only a few. With these tools, patterns can be found in data to help businesses to predict the future. The field of "big data" has also expanded to include activities around data mining, machine learning, and artificial intelligence to better optimize the "10,000-foot view" of a company's business. The term "predictive analytics" can be thought of as a catch-all for disciplines, which includes regression analysis: a technique that looks at values of correlated variables to prove or disprove a specific set of assumptions (Marvin, 2016).

Furthermore, predictive analytics is about recognizing patterns in data to project probability, according to Allison Snow, senior analyst of business-to-business (B2B) marketing at Forrester. One must realize that analytics is centered on probabilities, rather than facts. It is impossible to know what data is significant with predictive analytics, while it is usually the opposite case for traditional analytics. With predictive analytics, one can indicate the predictive data for the particular outcome that is being predicted (Marvin, 2016). For example, a company may have a model for predicting how much a particular customer may purchase over time. The model could also be based on purchases that the customer has already made, which predicts the next product that they are likely to buy. Sales forecasting and the use of marketing models to plan where best to place advertising buys are examples of predictive analytics (Davenport, 2014).

Thomas H. Davenport, a distinguished professor in management and information technology at Babson College, is a research fellow at the MIT

Initiative on the Digital Economy. He observes that the three components of supporting predictive analysis—the data, the statistics, and the assumptions—are all critical contributors to the final product. They are compared to the legs of a three-legged stool, each of which can cause a collapse if it is not functioning correctly. Davenport points out that the assumptions underlying a data set or statistical evaluation process are equally important to the data and the statistics. The primary assumption in a predictive analysis is often that the future always mirrors the past. For instance, the financial crisis between 2008 and 2009 was the result of invalid models that predicted the likelihood of mortgage customers repaying their loans, since the decline of house prices was not considered. As house prices fell, the models were useless in predicting mortgage repayment, given that a hidden assumption was that housing prices would always rise (Davenport, 2014). Clearly, the data and the statistics provide a necessary balance for conclusions drawn from the assumptions.

Moreover, both data analytics and predictive analytics are potent tools for extrapolating conclusions from data. Data analytics and predictive analytics are usually associated with being effective means of information analysis that exchange privacy for results. Also, there are staggering numbers of applications for this kind of analysis, especially with the various methods of recording information. Many companies are already applying these analytical techniques, and they will continue to refine them as the field expands. As more technology is being incorporated into products, those products provide more data for the producing companies. When data is made more readily available, then companies can find correlations between trends based on that data. Thus, they can draw different conclusions from the premises of their research.

Though data analytics and predictive analytics can be used to learn many things, they are not without their downsides. To reveal these downsides, we will explore both current and future applications of data analytics and predictive analytics, as well as how they function in different environments. We will explore how the flexible nature of the fields allows for data analytics and predictive analytics to apply to many purposes and contexts. Afterwards, we will explore the various benefits and drawbacks of using data analytics. Understanding how this type of analytics will be used, what is required, and the results that it provides will help to reveal which fields will most benefit from this kind of analytics.

The Value of Analytics

Essentially, there are a vast number of opinions about the usefulness and reliability of data and predictive analytics, since data analytics is applicable to many different fields. The uses of data analytics are innumerable, but more often than not they will relate back to theory testing. For example, someone will generate a theory based on existing data, a hypothesis, or even a whim, and will then create a model to process data to find correlations between that theory and the accumulated data. There is a great amount of data within scientific fields (i.e., medicine, biology, physics, etc.), so big data analytics can open the door for many scientific breakthroughs (Zhao, Gao, & Liu, 2018).

Regarding how the data is being processed, it is somewhat different for comprehensive data analytics and predictive analytics. Data analytics requires the manipulation of a data set for information, either about the evidence itself or a specific conclusion. Information about evidence is typically referred to as metadata, and it helps to observe trends, patterns, and so on. Predictive analytics works on a different level, while also working from a base data set. It uses both data and metadata to determine the likelihood of an event. Furthermore, predictive analytics is an excellent technique for determining risk management, but its accuracy is determined both by the data size and by the accuracy of existing data. Subsequently, we will be observing the uses of data and predictive analytics from medicinal, governmental, business, personal, scholastic, and financial perspectives.

The highest concentration of predictive analytics appears to be located in the pharmaceutical world, at least from the pool of information that can be observed. Healthcare professionals can make use of these techniques to determine which medications are right for individual patients, as models can forecast the risks and benefits of using medications (Flynn & Stevenson, 2018). This would require large companies to share data and existing predictive models, such as with the US Food and Drugs Administration's (FDA) Sentinel Initiative. Flynn and Stevenson also discuss the potential of wearable technology to give medication providers valuable data, as well as the tradeoff that many users may not want their information to be shared with those companies (Flynn & Stevenson, 2018). Trading privacy for better-tailored products, whether it is medicine or something else, is a common theme with these analytics (Manisha & Lathwal, 2016).

On the other hand, government institutions, such as the US Internal Revenue Service (IRS), make use of data analysis, but it is done in a somewhat dubious manner. While data mining often goes hand in hand with

government bodies, such as the National Security Agency (NSA) or the Department of Defense (DOD), it can be used to determine a variety of different things for other than security purposes. The IRS has moved on to creating profiles of taxpayers that are constructed from both public and commercial data sets, with the data processed using analytics. However, Houser and Sanders argue that these practices work against fair information practices and are a substantial threat to privacy. While these practices are very cost efficient for the IRS, they are invasive of taxpayer privacy and can lead to further government overreach (Houser & Sanders, 2017).

Additionally, data and predictive analytics can be applied in the workplace environment. Tasks such as hiring, training, and employee performance can also be enhanced by data analytics. Many companies use data analytics to make observations about their environment, but rarely do they use them to improve their training and business development. A business can analyze its data to figure out more cost-effective means of operating, as well as which hires are more likely to be successful (Giacumo & Breman, 2016).

Another newer method of data collection is coming from automobiles, specifically from electric vehicles. Analysts are referring to the information that is collected from these cars to make future improvements to their design without using any personal information about the consumer. Various aspects are recorded, such as charging time, charger style, driving distances, daily drives, charging frequencies, and other information about product use. By observing the typical usage patterns for their vehicles, designers can design better and newer versions of those products to cater to the data collected (Ahmed & Kapadia, 2017).

Outside of a business environment, predictive analytics can be applied to personal uses: identifying and preventing spam e-mails, for instance. Due to the variable nature of spam, it is tough to filter out completely. A system that can analyze e-mail logs and make use of predictive analytics can build trust criteria to determine which senders are considered spam (Vesely, 2017). Examining previous instances of spam would allow for the system to recognize spam behavior and to act accordingly when e-mails that fit the behavior profile are sent. Such a system could update its behavioral patterns based on newly formed spam (Yuksel, Cankaya, & Uncu, 2017).

In the scholastic setting, predictive analytics used analytical techniques to measure how likely students were to succeed. The documented system was referred to as Predictive Analytics Reporting or PAR. The method developed could be applied to show a student's likelihood of passing a

class, but Wagner and Longanecker make clear that the framework was designed with both conventional and non-conventional students in mind. Most federal data collection only focused on full-time students, which led to an incomplete data set. PAR was created to analyze the unchanging rates of postsecondary education success, as they have not changed for the past 40 years, despite attempts being made to improve them (Wagner & Longanecker, 2016).

Lastly, predictive analysis has various applications in a financial setting. John Blossom calls the multiple devices that collect and analyze data the "signal economy." More specifically, this is an economy that is driven by analysis of the data collected by those numerous different devices. In other words, it is significant data that has been processed to pinpoint trends. By subjecting economic data to the same processes that one may run to find the likeliness of an event's occurrence, one can figure out all sorts of potential areas of growth. Due to the staggering amount of data available from all kinds of different sources, markets, news outlets, universities, and even social media, software is required to process all of that data. Such data is continuously being collected and handled by software for hedge funds, the government, or some other private institution. Clearly, the information that comes from this kind of inquiry can predict how markets will act if enough data is being appropriately processed (Blossom, 2014).

As one can see, there are many applications for data analytics and predictive analytics that span pretty much every field. Many institutions, both public and private, make use of these analytical practices, and these processes are improving over time. In its current state, predictive analytics is more or less a means of trying to go as close as possible to determine the likelihood of a future event. The more data that is used, the more probability there will be of precision. With the advent of technological convergence, more devices can record information which will make it possible to map out the conditions for any event. From that point, analyzing the states that surround events will lead to the development of more accurate predictive analysis tools.

How Technology Impacts Analytics

Evidently, these analytics have proven to be beneficial in many aspects of human life. Predicting the future is a coveted tool in one's lifetime, especially within organizations. No one could dislike the ability to foresee success when creating a product, as investments could be much more easily

made with little to no error. With these tools, many types of success could be guaranteed.

However, there can be a cost to such powerful tools if they are not used responsibly. In Thomas Davenport's foreword for Eric Siegel's book *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie or Die*, he observes that the developing science of predictive analytics has, by and large, been for the good of humanity. Davenport asserts that predictive analytics can save lives, such as with healthcare, crime, and terrorism. He also addresses how these approaches can be detrimental, as one must exercise responsibility so as not to abuse this power. Large amounts of data would be required from people and companies, so predictive models can be compared to other powerful technologies or disruptive human innovations (Davenport, 2014). Davenport further suggests that Siegel counterpoints Nassim Nicholas Taleb's *The Black Swan* and other works. Taleb posits that most predictive efforts must be doomed to fail from the inherent unpredictability of things, but Siegel is the first to acknowledge that black swan events occur beyond prediction (Siegel, 2013). Furthermore, potential business applications continue to multiply. Models range from how predictive analytics is changing the retail industry to fintech start-ups using predictive modeling on fraud analysis and financial transaction risk. In essence, we have not fully tapped into the ways for industries to integrate data analysis and how predictive analytics tools can redefine business methods. After the breakthrough of mapping an artificial brain through creating a computer, there are infinite possibilities (Marvin, 2016).

Applications of Technology in Analytics

There are numerous examples of predictive analytics applications in multiple industries, especially involving technology. Primarily, marketing intersects with many sectors, as marketing campaigns have become much more scientific through the use of predictive analytics and big data. The ability to predict and recommend future purchases and actions is the principal benefit for many companies, and predictive analytics and big data are gateways to such a benefit.

With these tools, customer satisfaction and network safety can both be improved for some major corporations. For example, one of the largest users of predictive analytics is Amazon. The website uses past purchases and known "likes" to identify, predict, and offer other items to a customer. Similarly, fraud detection is a field where analyzing past patterns is key in

detecting future threats or criminal activity. Cybersecurity has become a huge budget item at many large corporations, such as banks, insurance companies, investment firms, credit card issuers, social media companies, and even law enforcement agencies (Moghaddass, Zuo, Liu, & Huang, 2015). PayPal avoids network breaches by using historical payment data, device type, country of origin, and user profiles to develop algorithms that detect fraud during any transaction. Even insurance companies, such as Liberty Mutual, use predictive analytics techniques to help evaluate lifestyle, health, and other factors to calculate life expectancy and related premiums. Furthermore, airlines use providers to analyze dynamic fare pricing that is based on the predicted number of travelers over a given period, for instance Virgin Atlantic utilizes Amadeus. Hotels do the same with room rates based on anticipated future travel volume (Farooq, 2016). Clearly, predictive analytics can help many corporations to gain a competitive edge.

By studying the company called Endor, one can find that *social physics* can also provide a massive amount of data in a matter of minutes. The science of social physics uses machine learning and algorithmic models to understand and predict crowd behaviors. It is comparable to Google, as it provides simple and fast results for any question (Matheson, 2017). Endor was recently spun out of the MIT Media Lab incubator, and it has developed a new platform that allows businesses to ask real-time questions: "Where should we open our next store?" or "Who is likely to try product X?" In one recent exercise, Endor analyzed Twitter data for an unnamed US defense agency to detect potential terrorist activity. Endor gathered 15 million data points with examples of 50 Twitter accounts of known ISIS activists that were based on identifiers in the metadata. The challenge was to detect new accounts with identifiers extremely well hidden in the metadata. Someone at Endor completed the task on a laptop in 24 minutes, detecting 80 "lookalike" ISIS accounts, 45 of which were from the pool of well-hidden accounts that were suspected by the agency (Ghofrani, He, Goverde, & Liu, 2018). As one can see, social physics was able to provide helpful results in a matter of minutes.

Endor is becoming a well-known company that is making great strides with predictive analytics. It was co-founded by Alex "Sandy" Pentland, Toshiba Professor of Media Arts and Sciences, and Yaniv Altshuler, a former MIT postdoctoral researcher. Early testing on the platform was done with the US Defense Advanced Research Project Agency (DARPA). It analyzed mobile data from individual cities in times of civil unrest to show how emerging patterns can help predict future riots. Lastly, Endor recently

won "Cool Vendor" status from Gartner, reserved for industry disrupters, and was acknowledged as a "Technological Pioneer" by the World Economic Forum (Matheson, 2017). It is evident that the merging of analytics and technology can guarantee more opportunities than one could have ever imagined over the history of industrialization.

How Technology and Analytics Come Together

The two types of analytics considered in this chapter, data and predictive, are powerful means for analyzing data in any number of different fields. Evidently, the results of the analyses can provide information to keep people safe, to make people money, to save people time, and so forth. Unfortunately, there is a cost in relation to privacy, as large amounts of data have to be analyzed to attain such results. With smart technology and smart environments, more data can be collected and processed, which will increase the results from data analytics and predictive analytics. These are great tools that are sure to grow and evolve with their related industries because they are derived from data. The immense applicability of these analytical methods ensures that they will be around for quite some time. Overall, it will be exciting to see how these analytical methods are applied to both new and existing fields, as it is inevitable that even more areas will jump on the bandwagon.

References

Ahmed, N. K., & Kapadia, J. (2017). Big data analytics: How big data is shaping our understanding of electrified vehicle customers. *SAE International Journal of Materials & Manufacturing, 12*, 99–107.

Blossom, J. (2014). The Signal Economy: How to target new revenues through predictive analytics. *Information Services & Use, 34*, 17–25.

Buchanan, E. (2017). Considering the ethics of significant data research: A case of Twitter and ISIS/ISIL. *PLoS ONE, 12*, 1–6.

Davenport, T. H. (2014). A predictive analytics primer. *Harvard Business Review*.

Farooq, M. (2016). Applications of predictive analytics in various industries. *BigData-MadeSimple.com*. Retrieved from http://bigdata-madesimple.com/applications-of-predictive-analytics-in-various-industries-2/

Flynn, A. J., & Stevenson, J. G. (2018). The future of data, analytics, and information technology. *Pharmacy Forecast*, pp. 31–34.

Ghofrani, F., He, Q., Goverde, R. M., & Liu, X. (2018). Recent applications of big data analytics in railway transportation systems: A survey. *Transportation Research: Part C, 90*, 226–246.

Giacumo, L. A., & Breman, J. (2016). Emerging evidence on the use of big data and analytics in workplace learning. *Quarterly Review of Distance Education, 17*, 21–38.

Houser, K. A., & Sanders, D. (2017). The use of big data analytics by the IRS: Efficient solution or the end of piracy as we know it? *Vanderbilt Journal of Entertainment & Technology Law, 14*, 817–872.

Klimberg, R. K. (2016). *Fundamentals of predictive analytics with JMP, Second edition*. Cary, NC: SAS Institute.

Manisha, A., & Lathwal, P. (2016). Exploring classification & clustering techniques for predictive analytics. *International Journal of Recent Research Aspects, 6*, 76–78.

Marvin, R. (2016). Predictive analytics, big data, and how to make them work for you. *PC Magazine*.

Matheson, R. (2017, December 19). Inventing the "Google" for predictive analytics. *MIT News*. Retrieved from http://news.mit.edu/2017/endor-inventing-google-predictive-analytics-1220

Moghaddass, R., Zuo, M., Liu, Y., & Huang, H.-Z. (2015). Predictive analytics using a nonhomogenous semi-Markov model and inspection data. *IIE Transactions, 47*, 505–520.

Siegel, E. (2013). *Predictive analytics: The power to predict who will click, buy, lie or die*. Hoboken, NJ: John Wiley & Sons, Inc.

Vesely, R. (2017). Predictive analytics: IU Health knows the patient in Room 103 is at high-risk for CLABSI (cover story). *H&HN: Hospitals & Health Networks, 91*, 20–25.

Wagner, E., & Longanecker, D. (2016). Scaling student success with predictive analytics: Reflections after four years in the data trenches. *Change, 48*, 52–59.

Yuksel, A. S., Cankaya, S. F., & Uncu, I. S. (2017). Design of a machine learning predictive analytics system for spam Problem. *Acta Physica Polonica, A, 132*, 500–504.

Zhao, W., Gao, L., & Liu, A. (2018). Programming foundations for scientific big data analytics. *Scientific Programming, 5*, 1–2.

Check for updates

CHAPTER 7

Effective Engagement of Digital Natives in the Ever-Transforming Digital World

Anju Varghese Philip and Zakkariya K. A.

Introduction

According to the latest estimates, more than 4.2 billion people worldwide have access to the Internet (Internet World Stats, 2018). That number is more than half of the world's present population. The digital divide is being demolished every second of every day and we are fast moving into a hyper-connected networked world where past boundaries of geographical distances, time zones, and cultural differences are ceasing to matter. The average person uses more than four connected devices every day, and digital technology has arrived as the invisible but all-knowing, all-controlling dimension of our everyday lives.

Modern-day businesses have to thrive in a data-driven culture where the only thing that does not change is change itself. And the speed at which the changes happen is increasing so exponentially that it feels like everyone is on a constant learning loop. Jumping off this loop is equivalent to committing professional suicide as you can no longer refuse to embrace constant learning and change. How can businesses negotiate a world where everything is digital and connected and your customers are at

A. V. Philip (✉) • Zakkariya K. A.
Cochin University of Science and Technology, Kochi, India

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_7

times more informed than you are? The best solution is by letting go of the scary connotations of the phrase 'digital disruption' and embracing the possibilities and opportunities of 'digital transformation.'

Digital transformation can be defined as the acceleration of business activities, processes, competencies, and models to fully leverage the changes and opportunities of digital technologies and their impact in a strategic and prioritized way (Edmead, 2016). It can also be called digital business transformation, since digital transformation is in essence transformation of the underlying framework of your business too. Technology innovation, customer behavior and demand, and external environmental factors are some of the reasons why a business might want to undergo a digital business transformation (Edmead, 2016). For this we need to build 'digital fluency' among our employees as well as increase the 'digital readiness' of our organizations (Briggs & Makice, 2012).

Digital fluency is an ability to reliably achieve desired outcomes through use of technology. It is the maximum individual potential to achieve desired outcomes through the use of digital technology (Briggs & Makice, 2012). It is not enough to know how to use the technological tools but one has to understand how to use them to achieve the desired outcomes. Therefore, being digitally fluent is a step up from mere digital literacy. A digitally literate person will have the basic skills required to participate in the digitally enabled social sphere, but, in contrast, a digitally fluent person will be able to navigate the same with an ease that comes from constant practice and immersion in the technology. Companies can be described as being 'digitally ready' when they are able to support the digital fluency of their employees by aligning their resources, culture, and purpose with the tasks they are asked to undertake (Briggs & Makice, 2012). Digital fluency is not a static state of being. It is a state of constant learning, up-gradation, and up-skilling because of the speed with which digital technologies change.

Like for any language, being digitally fluent implies that you have a level of comfort with the digital world that implies that you were born speaking the digital language. And that is where the digital natives or millennials enter the discussion. Who are the digital natives? Why are they important in a discussion about digital transformation of business? Digital natives (Prensky, 2001) is the terminology used to describe that part of the world's population who were born after 1980 and have been immersed in a technological environment all their lives, thereby making them especially adept at handling changing technology with an almost effortless flair.

Digital immigrants, on the other hand, lived part of their lives in a simpler time where technology was not the driving force of their lives and hence have a more difficult time navigating in and adapting to today's technology led milieu. It is expected that by 2025, digital natives will comprise 75% of the workforce, thereby making them the most economically powerful demographic group in history (dmresourcecenter.com, 2018).

In this scenario, it is important for businesses to understand who the digital natives are so that they can better leverage them as employees and also understand what makes them tick as customers. An organization that has to manage a successful digital business transformation will require a workforce that consists of both digital natives and digital immigrants to cooperate with the process and successfully make the transition so that no one is left feeling overwhelmed and falls behind in the process. This chapter attempts to throw some light on how to manage digital natives both as internal and as external stakeholders of the business.

Getting Familiar with the Digital Natives

To quote Mark Prensky (2001), digital natives are '"native speakers" of the digital language of computers, video games and the Internet.' 'Digital immigrants' are those who need to learn the digital language, they are not born speaking it, so they carry an 'accent' no matter how proficient they become at using digital tools. Digital immigrants prefer to use e-mail for online communication, whereas digital natives prefer the more synchronous forms of instant messaging. With mobile phones, digital immigrants favor speaking to people, whereas digital natives prefer speed texting. Digital natives prefer to share differently as well. Blogging is increasingly becoming popular among both digital immigrants and digital natives, but once again for different reasons. Digital natives use blogging to share personal experiences and treat personal blogging websites as forms of online journals. By contrast, digital immigrants tend to use blogging sites more as intellectual tools to share and discuss ideas with their peers (Prensky, 2001).

Not only do digital natives use technology differently—their lives are also being impacted by technology differently to the extent that their self-presentation, self-image, self-esteem, and wellbeing are influenced (Colbert, Yee, & George, 2016). They are often described as digitally literate, highly connected, experiential, social, and in need of immediate gratification (McMahon & Pospisil, 2005). Digital natives consider the

digital world to be part of their personalities. They expect to be connected at all times (Vodanovich, Sundaram, & Myers, 2010).

Digital is almost coded into the genes of the millennial generation; it is that natural for them. They are digital by default. Research has shown that they spend 41% of their free time on their devices (Morrison, 2015). They own an average of 2.4 devices, which might include smartphones, laptops, notebooks, iPads, and music streaming gadgets. Prensky (2001) says that brain scans show that their increased use of digital devices has actually caused changes to their brain structure. They relate to technology in a different way and are naturals at interacting and understanding with technology.

Among the adverse consequences of this constant connectedness is what is called FOMO in urban slang—the Fear Of Missing Out. It is a compulsive need to be up to date in what is going on within their social circles as well as in the outer world, which can even lead to disruption of healthy sleep patterns as people take their smartphones to bed with them and check messages and other updates during the nights and first thing in the morning.

At the same time, many among digital natives say they feel a sense of social fatigue, which is 'a feeling that being connected all the time is too much. A desire to disconnect and be offline and to regulate and bring down the time spent on social media' (techopedia.com). As a result, the more conscientious among them try to limit their time spent on social media by prioritizing their social media use and limiting themselves to being active on only a handful of social media sites rather than trying to be everywhere all at once and getting overwhelmed by it.

Many digital natives have grown up playing video games and have acquired a set of unique skills such as the ability to take risks and overcome obstacles, to learn from failures, to manage diverse teams, solve conflict, distribute rewards equitably, and so on as a result (Colbert et al., 2016). These skills make them invaluable assets in the workplace. Similarly, businesses that are looking to implement 'design thinking' for their innovation process will do well to inculcate the gaming skills of learning by trial and error (Beck & Wade, 2004; Glen, Suciu, & Baughn, 2014).

The digital workforce will likely be comfortable with technology-based instruction (Kraiger & Ford, 2006), giving organizations a low-cost, replicable solution for helping employees develop the skills that they need. Some tools that are currently in use are Google Drive for collaborative writing, Trello for collaborative project management, and Yammer or

Slack for communication and enterprise-level social networking. With more organizations using firm-generated social media content to connect with customers and build their brands, employees who understand and can leverage the power of social media will also be valuable to organizations (Kumar, Bezawada, Rishika, Janakiraman, & Kannan, 2016).

On the other hand, many millennials experience a decrease in face-to-face communication skills and nervousness during telephone calls because most communication is conducted via texting which is seen as more efficient and gives the opportunity to self-present in a favorable manner. Also they display decreasing levels of empathy probably because of increasing levels of technology immersion (Colbert et al., 2016).

Technology has blurred the lines between work and nonwork domains (Ramarajan & Reid, 2013). While email, the Internet, and even social media are integral tools for doing work, they also provide easy access to family, friends, online shopping, and other nonwork purposes while at work. Inability to disconnect from technology during nonwork hours can lead to higher levels of stress and work-life conflict (Ramarajan & Reid, 2013) adversely affecting employee wellbeing (Colbert et al., 2016).

What Can Companies Do for Digital Natives?

As Employees

Digital natives are highly cause driven and will love to work for companies that measure performance not just in financial terms but also in terms of contribution to social good. They are also aware of and support brands that have a social cause (dmresourcecenter.com, 2018).

While digital natives boast a strong work ethic, how, when, and where they accomplish the work should be left to them. Companies that expect a 9–5, 5-day workweek from them can expect a dissatisfied bunch of employees (Lovelock, 2018). Work-life balance is very important to them, especially because they have the technological tools that make this balance easier to achieve. While they want to earn enough to facilitate their life goals, they also want the free time to pursue their interests outside of work (Ramarajan & Reid, 2013).

They have grown up in a culture where collaboration is important (Commonplaces, 2014) and companies should provide the necessary internal knowledge management tools to accomplish this. They seek opinions and inputs from various quarters and believe that this gives them better solutions

to problems. Companies should be open to setting up collaborative teams not limited by location or department. Virtual teams are a great way to accomplish this (Gilson, Maynard, Jones Young, Vartiainen, & Hakonen, 2015).

Digital natives are not intimidated by titles, seniority, or positions. They see their superiors as mentors rather than bosses and will expect to be treated as equals. They are highly ambitious and expect to make managerial positions in two years of joining the work force (Ashridge and MSLGroup, 2014).

Loyalty to a particular organization is not important to them and they will switch jobs if they get bored. Millennials will not stay in a company and wait for a promotion every 3–5 years (Ashridge and MSLGroup, 2014). Companies should function as corporate lattices where employees are moved horizontally across business functions for different projects rather than as corporate ladders where the only movement is vertical (Benko, Anderson, & Vickberg, 2011).

Mobility is an important factor for millennials (Gilson et al., 2015). Companies should provide the necessary technological support to enable this. A combination of cloud computing and use of light technological solutions to enable mobility will increase efficiency in the workplace too. Also an option to bring your own technology (BYOT) will be a boon to this workforce (Puybaraud, 2012).

As Customers

For marketers and brands seeking to engage their millennial audience, it is first of all imperative to realize that as a result of living all their lives in an environment where information is available at their fingertips, they are impossible to fool and incredibly impatient with slow responses and campaigns that don't really speak to them. They carry their idealism proudly and will happily purchase brands and products that espouse a social cause. They are also vociferous in speaking out about anything that they dislike, so companies need to be very careful about how they engage with this audience online (dmresourcecenter.com, 2018). Be quick at apologizing for any mistakes and try to rectify them at the earliest. Best strategy will be to not make any mistakes at all!

Millennials are also characterized by their focus on accumulating life experiences rather than owning property. So they might opt to use their income to travel rather than settling down into a typical middle-class lifestyle buying homes, cars, and other large immovable investments

(Niewiadomski & Anderson, 2015). This demography is moving from a lifestyle of ownership to a lifestyle of easy access, instead of buying a car they will use a ride aggregator such as Uber, instead of investing on office space, they might prefer to work from shared office spaces. This type of thinking gives them all the conveniences of modern life that money can provide access to without having to bother with the hassles that come with ownership such as maintenance and upkeep.

They are constantly digitally networked at all times (Sørensen, 2016), which means they have easy access to information at all times and this means that they are attuned to an on demand lifestyle and have a growing sense of entitlement. Millennials handle half their ecommerce transactions from a mobile device and are the largest online audience and have more buying power than any other previous generation (dmresourcecenter.com, 2018).

Therefore, given their high buying power, their influence on the online market, their active social media presence, and most of all their impatience with delay, it is imperative that brands make all effort to understand and engage them. This is especially pertinent in a country like India, which is expected to become the youngest country in the world by 2022, with the average age of 29 years (Financial Express, 2017).

Firms should wake up to the advantages of using social media as a business tool to engage their digitally savvy audience (Lovelock, 2018). An effective social media marketing strategy can help increase an organization's brand awareness, improve customer engagement and satisfaction, and allow the company to establish their voice. Social media provides an accessible platform for direct communication between a company and a client and allows for a more human-centric approach to customer success and support. Firm-generated content (FGC) has a greater effect on customers with a longer customer–firm relationship, who are technologically savvy, and who are more prone to social networking. It is social media's ability to give 'voice' to customers in the form of likes and comments that makes FGC more effective. FGC can play a key role in strengthening customers' relationship with the firm by encouraging them to buy across several product categories (Kumar et al., 2016).

Statistics show that desktop sales are falling rapidly, while smartphones are being used for everything ranging from work to leisure. So the way to engage the consumer of today is by going mobile and social and being entertaining while doing so. That is spend more on mobile and digital

campaigns. Give your target consumers a chance to connect with their friends and other like-minded people through your campaign.

Use the trails of digital data that is left on the Internet by your consumers to really understand them. Social Media Analytics and Big Data should be seen as an opportunity to understand consumers and customize and individualize your offerings rather than as an insurmountable challenge best left to the geeks. While designing platforms for customer engagement remember to personalize, interact, be intuitive, attractive, and social (Vodanovich et al., 2010). Companies can work on enhancing the customer experience through understanding customers by using digital trails, analytics-driven customer segmentation, and digitally enhanced selling.

Skills That Are Needed for the Digital Workplace

Many times the obstacle toward implementing a successful digital transformation is people and not technology (Bonnet & Nandan, 2011). This is because digitalization changes the ways of working and increases the rate of change of the organization. It will require the manpower to rapidly acquire new skills and competencies, and also usher in a new organizational culture. Leadership also has to change to new forms to accommodate and enable this transformation to digital. Digital transformation requires the continuous alignment of people, technologies, organizational structures, and ultimately organizational culture (Briggs & Makice, 2012). Effective change management is required to manage this process well and adopting a design thinking style will help with the implementation (Gruber, De Leon, George, & Thompson, 2015). Design thinking can be defined as 'a human-centered approach to innovation that puts observation and discovery of often highly nuanced, even tacit human needs right at the forefront of the innovation process. It considers not just the technological system constraints but the socio-cultural system context' (Glen et al., 2014).

The speed and complexity of digitalization will change how we handle information (find, filter, and use information from many sources), interact with people (choosing an identity or identities to engage with others through multiple channels), inspire, and involve people to act in support of organizational goals, and empower them to move toward a future reality. To succeed in the digital age, a person must be able to appropriately weight the practical risks and future benefits of new tech. New

levels of skilling required to make the transition to digital fluency transcend from basic computer skills such as conducting a basic search for information, creating a profile on different social media apps to performing basic troubleshooting tasks to more advanced ones such as editing and uploading photos, tagging content, crafting semi-public content for various contexts, navigating complex privacy controls, and editing a wiki page. Some of the skills exhibited by an email-able person remain useful: good typing skills, writing a persuasive argument, and finding an optimal time to communicate.

Adaptation is what allows us to use our abilities effectively when shifting between contexts, usually as the world changes around us. Adaptation is an *ability of abilities*, drawing from the knowledge and relevant skills one accumulated from previous experience. Adaptation requires a mindset that sees all of this change as an opportunity to learn.

Briggs and Makice (2012) listed the skills and abilities needed for employees to make a successful digital transition as enumerated by Jenkins (2006) and summarized them as given below:

1. *Networking*: The ability to search for, synthesize, and disseminate information
2. *Play*: The capacity to experiment with one's surroundings as a form of problem-solving
3. *Performance*: The ability to adopt alternative identities for the purpose of improvisation and discovery
4. *Simulation*: The ability to interpret and construct dynamic models for real-world processes
5. *Appropriation*: The ability to meaningfully sample and remix media content
6. *Multitasking*: The ability to scan one's environment and shift focus as needed to salient details
7. *Distributed cognition*: The ability to interact meaningfully with tools that expand mental capacities
8. *Collective intelligence*: The ability to pool knowledge and compare notes with others toward a common goal
9. *Judgement*: The ability to evaluate the reliability and credibility of different information sources
10. *Transmedia navigation*: The ability to follow the flow of stories and information across multiple modalities

11. *Negotiation*: The ability to travel across diverse communities, discerning and respecting multiple perspectives, and grasping and following alternative norms
12. *Visualization*: The ability to interpret and create data representations for the purposes of expressing ideas, finding patterns, and identifying trends

A number of tools are now available to measure the new skills needed. These include the iSkills assessment (Katz, 2007), New Literacies (Lankshear & Knobel, 2006), and the New Media Literacies (NML) framework (Literat, 2014).

Discussion and Conclusion

Digital natives are a great mascot for organizational transformation programs, as they have the ability to come into a company unhindered by previous experience with legacy technologies and provide new solutions to modernize business processes and drive a truly digital workplace. As digital natives join the workforce, they bring with them a willingness to embrace technology as a driver of not just productivity, but also profitability. Organizations must adopt this mindset and plan for digital disruption in the modern workplace, or risk being left behind.

Digital natives can help their older colleagues use technology more efficiently and effectively, thereby increasing productivity. The technology gap between generations will lessen as elder generations learn to embrace and understand new technologies like their digital native peers. Digital natives' higher level of intuition when it comes to technology, coupled with the fact that new products are easier to use than their predecessors, means that integration with technology will be less of an issue for all generations (Gilson et al., 2015). In addition, digital natives have a healthy level of respect for the older generations' work ethic and values and will want to imbibe these from their seniors at work.

Ultimately, digital natives seek the same things from their employer as any other generation that came before them. A sense of belonging and value and an opportunity to contribute to the success of their employer. They seek to be respected and nurtured and to work for companies that work for social causes and innovation and have a sense of social responsibility. They want their work and lives to be an extension of their own

identities and will want to work in and buy from companies that provide them this avenue for self-expression.

As customers, digital natives are highly tech savvy, value conscious, and cause driven. They desire instant gratification and quick responses for their issues. Companies should set in place efficient teams to monitor, engage, and address their millennial customers, as this is considered the most economically powerful demographic that the world has ever seen. The most effective customer touch point for this audience would be some form of social media outreach in addition to the usual sales and promotion activities. Companies that go the extra mile to really understand their digital native market will be able to target their offerings in a way that truly speaks to the identity of the millennial.

To conclude, business leaders should acknowledge the direction that technology is leading the world in, embrace the future, and ride the waves of technology into the hearts and minds of their digital native stakeholders. Since the number of digital native stakeholders are continuously increasing and leading to a world filled only with digital natives in the long run, companies should have short-term and long-term plans and strategies for attracting, engaging, and managing digital native to retain them as loyal stakeholders. Digital transformation is a never-ending process as the technology is growing and transforming every day, and this poses a challenge to organizations, as they need to cater to the ever-changing needs and behaviors of the digital native. No company would be able to continuously either engage or satisfy the digital native without having continuously upgraded technologically based plans, strategies, and actions.

References

Ashridge and MSLGroup. (2014). The Millennial Compass—Truths about the 30-and-under generation in the workplace. Retrieved from http://www.scribd.com/doc/211602632/The-Millennial-Compass-The-MillennialGeneration-In-The-Workplace.

Beck, J. C., & Wade, M. (2004). *Got game: How the gamer generation is reshaping business forever*. Boston, MA: Harvard Business School Press.

Benko, C., Anderson, M., & Vickberg, S. (2011). The corporate lattice: A strategic response to the changing world of work. *Deloitte Review*, Issue 8

Bonnet, D., & Nandan, P. (2011). *Transform to the power of digital: Digital transformation as a driver of corporate performance* (Report). Capgemini Consulting.

Briggs, C., & Makice, K. (2012). *Digital fluency: Building success in the digital age*. SociaLens.

Colbert, A., Yee, N., & George, G. (2016). The digital workforce and the workplace of the future. *Academy of Management Journal, 59*(3), 731–739.

Commonplaces. (2014, January 21). Digital natives and the future of the workplace [Blog post]. Retrieved October 10, 2018, from https://www.commonplaces.com/blog/digital-natives-and-the-future-of-the-workplace/

Edmead, M. (2016, May 2). *Digital transformation: Why it's important to your organization*. Retrieved October 13, 2018, from https://www.cio.com/article/3063620/it-strategy/digital-transformation-why-its-important-to-your-organization.html

Gilson, L. L., Maynard, M. T., Jones Young, N. C., Vartiainen, M., & Hakonen, M. (2015). Virtual teams research: 10 years, 10 themes, and 10 opportunities. *Journal of Management, 41*(5), 1313–1337.

Glen, R., Suciu, C., & Baughn, C. (2014). The need for design thinking in business schools. *Academy of Management Learning & Education, 13*(4), 653–667.

Gruber, M., De Leon, N., George, G., & Thompson, P. (2015). From the editors: Managing by design. *Academy of Management Journal, 58*(1), 1–7.

Internet World Stats: Usage and population statistics. (2018, June 30). Retrieved October 13, 2018, from https://www.internetworldstats.com/stats.htm.

Jenkins, H. (2006). *Confronting the challenges of participatory culture: Media education for the 21st century*. White paper for MacArthur Foundation.

Katz, I. R. (2007). Testing information literacy in digital environments: ETS's iSkills assessment. *Information Technology and Libraries, 26*(3), 3–12.

Kraiger, K., & Ford, J. K. (2006). The expanding role of workplace training: Themes and trends influencing training research and practice. In L. L. Koppes (Ed.), *Historical perspectives in industrial and organizational psychology* (pp. 281–309). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.

Kumar, A., Bezawada, R., Rishika, R., Janakiraman, R., & Kannan, P. K. (2016). From social to sale: The effects of firm-generated content in social media on customer behavior. *Journal of Marketing, 80*(1), 7–25.

Lankshear, C., & Knobel, M. (2006). *New literacies: Everyday practices and classroom learning*. Open University Press.

Literat, I. (2014). Measuring new media literacies: Towards the development of a comprehensive assessment tool. *Journal of Media Literacy Education, 6*(1), 15–27.

Lovelock, D. (2018, August 14). *How digital natives are changing the modern workplace*. Retrieved October 13, 2018, from https://www.nintex.com/blog/how-digital-natives-are-changing-the-modern-workplace/

McMahon, M., & Pospisil, R. (2005). Laptops for a digital lifestyle: Millennial students and wireless mobile technologies. *Proceedings of the Australasian Society for Computers in Learning in Tertiary Education, 2*, 421–431.

Morrison, K. (2015, August 11). *Report: Digital natives do everything from mobile devices*. Retrieved October 11, 2018, from https://www.adweek.com/digital/report-digital-natives-do-everything-from-mobile-devices/

Niewiadomski, R., & Anderson, D. (2015). *Digital Generation: Is this the beginning of paradigm shift in ownership?* Retrieved October 13, 2018, from https://cie.acm.org/articles/digital-generation-beginning-paradigm-shift-ownership/.

Prensky, M. (2001). Digital natives, digital immigrants part 1. *On the Horizon, 9*(5), 1–6.

Puybaraud, M. (2012, February 1). *Digital natives: A tech-savvy generation enters the workplace*. Retrieved October 14, 2018, from https://workdesign.com/2012/02/digital-natives-a-tech-savvy-generation-enters-the-workplace/

Ramarajan, L., & Reid, E. (2013). Shattering the myth of separate worlds: Negotiating nonwork identities at work. *Academy of Management Review, 38*(4), 621–644.

Sørensen, C. (2016). The curse of the smart machine? Digitalisation and the children of the mainframe. *Scandinavian Journal of Information Systems, 28*(2), 57–68.

The impact of digital natives on the future of business. (2018). Retrieved October 11, 2018, from http://dmresourcecenter.com/unit-1/digital-natives/

Vodanovich, S., Sundaram, D., & Myers, M. (2010). Research commentary—Digital natives and ubiquitous information systems. *Information Systems Research, 21*(4), 711–723.

'*With an average age of 29, India will be the world's youngest country by 2020*'. (2017, March 26). Retrieved October 14, 2018, from https://www.financialexpress.com/india-news/with-an-average-age-of-29-india-will-be-the-worlds-youngest-country-by-2020/603435/

CHAPTER 8

Robotics, Artificial Intelligence, and the Evolving Nature of Work

Craig Webster and Stanislav Ivanov

Introduction

In their thought-provoking paper Brynjolffson and McAfee (2015) ask the question 'Will humans go the way of horses?' and they have good ground to do so. Human labour has been largely replaced by machines in ways that consumers can understand and this has been going on for decades. For example, many will have a historical memory of going to a bank during the bank's hours of operation to make a withdrawal of cash. Now, consumers regularly go to machines to make cash withdrawals from the accounts and the labour for making the transaction is largely mechanized. While much work in the service industry is still based upon human labour, manufacturing is much more automated. While humans may still be needed to perform many commercial operations, human labour is much less critical than it had been at any time before in history.

C. Webster (✉)
Ball State University, Muncie, IN, USA
e-mail: cwebster3@bsu.edu

S. Ivanov
Varna University of Management, Varna, Bulgaria
e-mail: stanislav.ivanov@vumk.eu

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_8

Recent advances in robotics, artificial intelligence, and automation technologies (RAIA) (Bhaumik, 2018; Miller & Miller, 2017; Neapolitan & Jiang, 2013; Russell & Norvig, 2016; Warwick, 2012) have allowed companies from various sectors of the economy to adopt RAIA in search of lower costs, faster production time, provision of consistent product quality, management of supply chain operations, etc. While initially it was the manufacturing that used industrial robots (Colestock, 2005; Cubero, 2007; Low, 2007; Pires, 2007), currently all other sectors of the economy and society intensively adopt RAIA technologies: from supply chain management (Min, 2010), agriculture (Driessen & Heutinck, 2015), autonomous vehicles (Maurer, Gerdes, Lenz, & Winner, 2016), and warfare (Crootof, 2015; Sparrow, 2007), through travel, tourism, and hospitality (Ivanov & Webster, 2018; Ivanov, Webster, & Berezina, 2017; Murphy, Hofacker, & Gretzel, 2017), education (Ivanov, 2016; Timms, 2016), journalism (Clerwall, 2014), provision of legal (Remus & Levy, 2015), and other services (Huang & Rust, 2018; Wirtz et al., 2018), to trading on the financial markets (Dunis, Middleton, Karathanasopolous, & Theofilatos, 2017), and implementing medical operations (Kaur, 2012; Mirheydar & Parsons, 2013; Schommer, Patel, Mouraviev, Thomas, & Thiel, 2017). Both large and small companies use chatbots to communicate and maintain their relationship with customers (Hill, Ford, & Farreras, 2015; Xu, Liu, Guo, Sinha, & Akkiraju, 2017). Social robots enter actively our lives (Agah, Cabibihan, Howard, Salichs, & He, 2016; Nørskov, 2016; Royakkers & van Est, 2016), and already redefine our understanding of 'sex' (Cheok, Devlin, & Levy, 2017; Lee, 2017). These brief examples reveal the overwhelming intertwining of RAIA technologies into the fabric of society, leading to huge transformations of the ways humans live, work, and do business (Makridakis, 2017; Talwar, 2015; Talwar, Wells, Whittington, Koury, & Romero, 2017). Some authors even state that economists and marketers need to broaden their definition of a 'consumer' to include robots in it (Ivanov & Webster, 2017).

The advances in RAIA technologies are expected to accelerate in the future and some researchers consider the possibility that artificial intelligence might at one point surpass human intelligence—a point of time usually referred to as 'technological singularity' (Callaghan, Miller, Yampolskiy, & Armstrong, 2017; Kurzweil, 2005; Shanahan, 2015). While we do not know whether silicon will surpass carbon in terms of intelligences and whether transhumanism and human augmentation through microchip brain implants is a viable future and not science fiction,

we are sure that technological advances in RAIA, Industry 4.0 (Andelfinger & Hänisch, 2017; Schwab, 2016; Skilton & Hovsepian, 2018), and the Internet of things (Sendler, 2018) are challenging the role of human labour as a production factor to a point when all or an overwhelming share of goods and services are delivered by RAIA technologies instead of human employees. Such an economic system is called 'robonomics' (Crews, 2016; Ivanov, 2017). Experts' opinions about that automated robot-based economy and society are not uniform. One group of researchers expresses explicit fear of AI technologies (Barrat, 2013; Bostrum, 2014; Crews, 2016; Leonhard, 2016), considers the self-aware AI as 'our final invention' (Barrat, 2013) and paints dystopian dooms-day Terminator-like scenarios for the human species. Others seem overly optimistic and perceive technology as the ultimate utopian solution to all human problems, including the merger between humans and machines (Kurzweil, 2005). The majority of authors, however, take a more pragmatic approach and see the opportunities that AI and robonomics as an economic system would create like extended life expectancy, improved health and quality of life, more time for activities people value, to name just a few (Brynjolfsson & McAfee, 2014; Frank, Roehring, & Pring, 2017; LaGrandeur & Hughes, 2017; Talwar, 2015; Talwar et al., 2017), although the expected radical abundance of goods and services that technology experts promise (Drexler, 2013) may not benefit all stakeholders.

Within this domain of thought, the future of human's work and the impacts of digitalization and automation on the labour market have received a great deal of attention by researchers (e.g., Alasoini, 2017; Association for Advancing Automation, 2017; Daugherty & Wilson, 2018; Ford, 2009, 2015; Frank et al., 2017; Harteis, 2018; Kaplan, 2015; Kauhanen, 2016; West, 2018). As a matter of fact, in a recent highly cited research, Frey and Osborne (2017) assess that 47% of US jobs are already susceptible to computerization. Some jobs are very obviously on the brink of extinction. For example, the most common job in 29 of the 50 states in the United States is the 'truck driver,' according to 2014 data (NPR, 2015) and the self-driving car/truck (a technology on the brink of being viable on the road) will almost immediately lead to the unemployment of these workers (Solon, 2016).

This raises justified concerns among researchers, politicians, and industry representatives how people, companies, economies, governments, and societies as a whole would need to adapt to the new technological, economic, social, and political realities that robonomics would create. For

example, in 2016, the Executive Office of the President of the United States, National Science and Technology Council, and the Committee on Technology published three documents intended to prepare the US economy for the age of automation: *The National Artificial Intelligence Research and Development Strategic Plan* (2016), *Preparing for The Future of Artificial Intelligence* (2016), and *Artificial Intelligence, Automation, and the Economy* (2016). In the same manner, in 2017, OECD published an extensive report dedicated to the implications of automation for businesses and governments (OECD, 2017), while the World Economic Forum published annual reports on the future of jobs (WEF, 2016, 2018).

This chapter contributes to the body of knowledge by elaborating on the impacts of RAIA technologies on the labour market and the nature of work. This chapter looks into such things as the employment dilemma (Giarini & Liedtke, 1996), an issue related to the human need to have productive work and the modern industrial capability for productivity. It also looks into the issue of the human Luddite possibilities, as humans should have a fear of being replaced by robots (Lehman, 2015; McClure, 2017). In addition, the chapter also looks into the issue of the ethical use of robots (Lin, Abney, & Bekey, 2014), as robots are not merely replacements of human labour but also entails a real impact upon the human population.

The Economics of Robotics and AI

The first element to look into is the issue of why RAIA technologies are being implemented. The obvious reasons for the implementation are the overlap between the technological abilities that now exist and the need for efficiency and cost-effective production. While the concept of the robot was invented in 1920 (NPR, 2011), the robot (and more importantly the intelligent robot) was a thing of the future and not entirely practical. By the 1980s, robots were being utilized widely in manufacturing, although service industries have generally lagged behind in the use of robots.

At its heart, the current robotic economy undermines the economic principle of the labour theory of value was ascribed to by such thinkers as Karl Marx, Adam Smith, and David Ricardo. While traditional economic theory identifies land, labour, and capital (and sometimes also 'enterprise' or 'entrepreneurship' as the factors of production), we see a shift in the stress on which of these factors is most critical. Traditional economic theory holds that human labour is a necessary factor for the production of

wealth, meaning that human labour is critical in the productive process, something that any casual reading of Marx makes very clear. However, with the rapid advances in robotics and artificial intelligence, human labour is much less stressed as a necessary input, and this will be increasingly true once robots are used to produce robots.

The widespread implementation of a robotic economy will lead to a new economic and social landscape in which there will be winners and losers. Since the technological capability to replace human labour with increasingly sophisticated technologies exists, those humans with labour to sell that can be replaced easily and cheaply with new technologies will be replaced, leading to a new political economy of robots (Kiggins, 2018). This will be a new world in which most labour in terms of production of goods and services is created by machines, meaning that humans will be increasingly marginalized from the productive process, although there is the enduring question of whether robots will largely be used as complements to human labour or substitutes to human labour (DeCanio, 2016).

Another political and economic trend that will impact upon the new economy is a by-product of the globalized economy. Since economies are now interconnected as never before and there are massive advantages to those who conquer a market early, since markets tend to be big and the ability to penetrate such markets. What we can expect in markets is the markets will increasingly gravitate towards 'winner-take-all' markets (Brynjolfsson & McAfee, 2014). There are a number of reasons for this, including the digitalization, improvements in communications (and transportation), and the increased importance of networks and standards. In other words, the economy is now different, as global markets enable consumers to have access to information with regards to goods and services and information costs with regards to products and transportation of products is dropped, and networks enable the spread of information and good and services as never before. Economies of scale develop, enabling the establishment of near monopolies, as penetrating markets outside of prevailing channels and networks becomes increasingly difficult for start-ups.

The value of labour and the way that labour is treated can be understood by its value in the market. The relationship between increased labour productivity and wages in the United States tell an interesting story. While the data in the United States show massive increases in productivity per worker since the 1950s, by the mid-1970s, increases in salaries have not kept up with increases in productivity (Watson, 2018). It seems that companies see

no need to compensate workers for their productivity, resulting in an outcome that has a massive impact upon the economy and society, people who are employed and productive but do not experience salary/wage increases as a result of their labour. It is also suggestive that workers are so productive that there are underemployed populations driving the real wages down. What is most interesting is that this is occurring in economies in which there are demographic declines, meaning that the available labour pool is generally shrinking and will continue to shrink. Another issue is the question of the sustainability of such economies, as machines do not consume many goods nor services, unlike humans (Ford, 2015, p. 196). This means that economies may suffer in the future, as machines do not consume the way that humans do. Manufacturers may be producing high-quality goods at a low price per unit, but the question remains who will be able to purchase the goods and services, when salaries remain relatively flat in many countries.

Directions of RAIA Impacts on the Labour Market

For us human mortals living today, we will have to contend with the massive changes that robotics and artificial intelligence will do to the labour market in the next 10, 20, or more years. While we understand that in 100 years, the production and service landscape will change dramatically because the capability of technologies that we now are impressed with will seem very primitive, we have to prepare for the nearer future. In the nearer future, we can expect some major impacts that are almost predictable, such as the automation of a great deal of the transportation with self-driving trucks and cars. At any rate, there will be an impact upon different countries in different ways based upon the country's current level of development and social/technological factors. There will also be some impact upon social structures and the mobility of populations, as well as the mobility of production. At any rate, legislation will have to change a great deal and the legal/political elements as well.

The first and major impact of technologies will be the elimination of jobs via automation (Frank et al., 2017), although there will be many jobs lost, there will also be the creation of new jobs and the changing face of other jobs. Some publications have discussed various solutions to how technological unemployment can be dealt with (Stevens & Marchant, 2017; Swan, 2017). Many of the jobs we now see exist in the market will be changed a great deal with many jobs being de-skilled, re-skilled, or

upskilled in terms of their tasks associated with them. There is also a geographical element to this, as well. There is an indication that automation of many tasks may enable production to take place in places in which labour costs are notoriously high such as Germany, where shoe production is now taking place (Wiener, 2017) presumably because the other factors of production are equal to or lower than in Asia and labour costs approach zero because of automation. Production may return to the Western economies, as labour costs will largely be bypassed and production may now be done in ways that lower transportation costs.

Another major impact that will occur is the quality of the jobs that persist. Humans will likely continue to do some of the work that robots and computers will not yet be able to do and some of the jobs are rather surprising. For example, while we can reasonably expect that artificial intelligence can handle most of service complaints and inquiries from humans, most robots have a hard time doing some activities that many humans find very simple, such as opening doors. It seems that there will be a need for humans to do many of the physical tasks that machines cannot do or at least do well, although it is possible that social changes may work in ways to bypass this. For example, while currently no robots can change the sheets on a bed well, it is possible that hotels in the future could hand guests clean sheets and insist that the guests make their own beds, bypassing the need to pay for human labour to change the sheets on the bed. Such a social/cultural shift may be accepted by hotel guests, if there are cost advantages.

Because of the disruption of advanced robotics and artificial intelligence into the production of goods and services, humans will likely have to find a niche in the world's gig economy. Since human labour will be increasingly relegated to physical or mental tasks that humans can still perform better than machines, humans will have to hustle, doing jobs that machines cannot yet do well nor as effectively as machines. One of the problems that will be a political outcome of this is that many people will be reliant upon short-term tasks and jobs that may exist below the radar of tax authorities and do not supply the human with the legal and social protections that a person would have with a full-time job that exists above the board. How males and females will work under these conditions and survive under such conditions will be interesting, but it is so far unknown how the sexes will either react or flourish under such unstable and shifting work environments (Schwab, 2016).

The class differences will increase, as has been the case in nearly all developed countries in recent decades, since labour will consist of skilled knowledge-based workers and less-skilled workers. Those that are less skilled will constantly run into competition with robots and artificial intelligence that will undermine the value of labour (Economic Times, 2018). There will be some specializations that remain difficult to replace with robots or artificial intelligence, and these positions will continue to function, although there will be downward pressures upon such jobs, as the demand remains equal, the supply of labour for such jobs will decrease the returns for those working such jobs. However, those who feel secure with skills and are knowledge workers should not feel too secure, as there are technological changes that can make their labour redundant. For example, university professors now teach many online courses. Such courses are generally done for the convenience of the students, but there are risks for the teachers, as many courses may rely on exams/tests that are multiple choice and can already be graded with very basic computer technology, and it may not be long before computers can grade and give intelligent feedback to students on essays (see also Ivanov, 2016).

We can also expect that the disruptions that will take place will have a massive reaction among the public and there should be political responses. Since humans may be less key to the production of goods and services and many humans will be replaced with robotic/AI labour, what should be done to sustain the economy and sustain the political economy? A likely answer spoken about is guaranteeing humans a universal basic income (Santens, 2017; Sheahen, 2012). This income would have the stabilizing impact by ensuring that human consumers can purchase the goods and services that a largely automated economy would produce. Paying for such a plan would require a different form of taxation from how taxation is now considered, as a great deal of taxation in economies stems from the taxation of the wages of working people. So with a reduction in the number of humans working and paying taxes on their incomes and the reduction of jobs, entire new systems of taxation will have to be created. The current economy and its laws presume human labourers paying taxes upon their salaries and presumes that such human labour is necessary to sustain economies. Laws also presume that people are engaged in full-time employment in which people have stable incomes and benefits, meaning that many of the laws that now exist regulating labour presume an old-fashioned industrial model of labour. But the new economy, to a large extent, depends upon gigs and short-term contracts and laws have not kept up with these changes.

Staying Competitive on the Labour Market

The future, at least the near future, is one that requires a skilled workforce. Table 8.1 lists the top ten job skills the World Economic Forum (2018) feels the workforce should have in 2020. The list of skills seems to be uniquely intellectual and emotional in nature. There is no mention of physical strength or physical agility. According to the World Economic Forum (2018), the skills needed are intellectual and emotional in nature, including solving complex problems, critical thinking, and coordinating with others.

These key skills identified by the World Economic Forum presumes that the competitive advantage that humans have over robots and artificial intelligence is not based upon physicality but upon intellectual capabilities and the ability to interpret the emotional landscape of the workplace. The development of such skills would reasonably thought to be developed by people through education/training, experience, and naturally occurring intellectual ability. What this implies is that education will have to work in ways to cultivate those talents that only humans at this point in time can develop in ways that are useful in the economy (creativity, complex communication skills, emotional intelligence, etc.). The need to further develop these very human capabilities will have to depend upon an educational system that is based upon an old industrial model from a previous century, presuming a tiered system of education that would build skills that are permanent. The new economy will likely need a new approach to education, one that is based upon the development of human skills rather than building of knowledge and will have to retool to consider the needs

Table 8.1 Top ten job skills needed in 2020

Rank	*Skill*
1.	Complex Problem Solving
2.	Critical Thinking
3.	Creativity
4.	People Management
5.	Coordinating with Others
6.	Emotional Intelligence
7.	Judgement and Decision Making
8.	Service Orientation
9.	Negotiation
10.	Cognitive Flexibility

Source: World Economic Forum (2018)

of people in the workforce who will have to pick up new skills to be relevant in a flexible and changing economy. Since many employees will change jobs and industries, they will need relevant education that will enable them to have relevant skills in the new positions that they take in different companies and fields. Companies will also have to pick up some of the responsibility for the training and retraining of employees and such responsibility should not be the sole responsibility of universities, entities comfortably insulated from the economy in many respects.

Apart from how individuals and how the educational systems will have to change, industry will also have to change. In the short run, for companies to stay relevant and competitive, companies will have to invest in robots and artificial intelligence. While such investments may require investments upfront, such investments will assist in the transition to the new economy, despite the risks and hiccups in the system. The changeover for most industries is a sure thing. While adopting some technologies later may offer some advantages, adopting early may enable the customers and workforce to adjust to the new realities and work in ways that are more effective and efficient.

Conclusion

Humans have been on the Earth for quite some time and we hope that the species will persist. While humans probably will not go the way of the horses (Brynjolffson & McAfee, 2015), there are massive changes caused by technological advances that humans need to be aware of and must adjust to. First, we expect that humans will be removed increasingly from the production of goods and services. Human labour will remain one the factors of production but is decreasingly the key component in the production process. The qualities that the human labourer has and brings to the table in the factors of production will remain the ability to create, to interpret, and to work with others. The human will also be prized in the near future for the physical ability to do things that many robots have a hard time doing well and quickly (opening doors, walking, and other physical things), the human will remain a great multi-purpose tool that can negotiate the geography of a house or office space. But the flooding of labour markets with labour displaced by robots and artificial intelligence will mean that such physical jobs will be flooded.

For humans, those lacking the skills to compete in the new economy will be competing against each other. Skills that so far are unique to

humans will have to be developed for many workers to remain relevant. Problem solving, emotional intelligence, interpersonal communications, and other skills that humans can do well if they are sophisticated enough to do it will enable humans to remain relevant in the new economy. But the traditional establishments that train humans will have to adjust to the needs of the new economy and the demands upon workers.

Second, the political and legal system has to take account of the new economy. As labour becomes more flexible and concentrated upon gigs and contracts, governments and taxation systems have to keep up with the new economic reality. From the perspective of the government, the traditional way of taxing working people will be less prevalent and the persistence of the gig economy/sharing economy will mean that taxation systems will have to be amended to remain relevant in an economy that has changed. Taxation will change and keeping the new economy stable and healthy will have to rely on different thinking. With a large number of workers who are not needed, guaranteeing that people have a reliable way of living, especially if they are not needed in the workforce, is a necessity. One easy way to adjust for this is the guarantee of a minimal income for all, an idea that may seem radical but has a long history and has become quite popular to speak about in recent years (see, for example, Caputo, 2012; Stern, 2016). The stability of the political economy rests upon the need to have consumers who use products and services and the lack of the need for much of the labour available in the economy means that much of the workforce will be competing for de-skilled jobs and with people who have lost their jobs due to increasing robotization of the workplace.

Finally, there is the major concern about the assumptions of the human's ability to be able to be relevant in the new economy. While the World Economic Forum (2018) can identify skills needed in the workplace, it is a question as to whether all humans have the ability to develop skills that are needed by industry. While there are some who may contend that IQ is a meaningful concept (Gould, 1996), most agree that there is variation in each human's ability to absorb, retain, and process information. Unless it is found that all humans have roughly the same intellectual capacity to develop the skills needed for the new economy, it would seem that many would be excluded from the new economy, simply because of their incapacity to develop the skills needed. It could be, then, that a substantial portion of humanity could be excluded from the new economy, simply because of a lack of the ability to develop the skills needed to compete as a human in the economy.

In conclusion, humans are entering a new age, an age of mechanical reproduction that is unlike anything envisioned before. While humans may remain a factor of production, they may be decreasingly critical in terms of the terms of physicality of work and even more critical in terms of the intellectual needs of the workplace. At any rate, the role of the human will be different. As a result, many of the institutions that are part of the economy, such as government and education, will have to change to be relevant and consistent with the needs of a society that has an increasingly automated productive base.

References

Agah, A., Cabibihan, J. J., Howard, A., Salichs, M. A., & He, H. (Eds.). (2016). *Proceedings of the 8th International Conference on Social Robotics (ICSR) 2016*, Kansas City, MO, USA, November 1–3, 2016. (Vol. 9979). Springer.

Alasoini, T. (Ed.). (2017). *Nordic Future of Work Conference I: The Future of Work and new Forms of Work from the Global and the Nordic Perspectives*. TemaNord 2017:551. Nordic Council of Ministers. https://doi.org/10.6027/09086692

Andelfinger, V. P., & Hänisch, T. (Eds.). (2017). *Industrie 4.0: Wie cyber-physische Systeme die Arbeitswelt verändern*. Wiesbaden: Springer Gabler.

Artificial Intelligence, Automation, and the Economy. (2016). Published by the Executive Office of the President of the USA. Retrieved from https://obamawhitehouse.archives.gov/sites/whitehouse.gov/files/documents/Artificial-Intelligence-Automation-Economy.PDF

Association for Advancing Automation. (2017). *Work in the automation age: Sustainable careers today and into the future*. Retrieved from http://www.a3automate.org/docs/Work-in-the-Automation-Age-White-Paper.pdf

Barrat, J. (2013). *Our final invention: Artificial intelligence and the end of the human era*. New York: Macmillan.

Bhaumik, A. (2018). *From AI to robotics: Mobile, social, and sentient robots*. Boca Raton, FL: CRC Press.

Bostrum, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. W. W. Norton & Company.

Brynjolffson, E., & McAfee, A. (2015). Will humans go the way of horses? Labor in the second machine age. *Foreign Affairs, 94*(4), 8–14.

Callaghan, V., Miller, J., Yampolskiy, R., & Armstrong, S. (Eds.). (2017). *The technological singularity: Managing the journey*. Berlin: Springer.

Caputo, R. (Ed.). (2012). *Basic income guarantee and politics: International experiences and perspectives on the viability of income guarantee*. Springer.

Cheok, A. D., Devlin, K., & Levy, D. (Eds.) (2017). Revised selected papers of the *Second International Conference on Love and Sex with Robots (LSR) 2016*, London, UK, December 19–20, 2016. Springer.

Clerwall, C. (2014). Enter the robot journalist: Users' perceptions of automated content. *Journalism Practice, 8*(5), 519–531.

Colestock, H. (2005). *Industrial robotics: Selection, design, and maintenance*. New York: McGraw-Hill.

Crews, J. (2016). *Robonomics: Prepare today for the jobless economy of tomorrow*. CreateSpace Independent Publishing Platform.

Crootof, R. (2015). War, responsibility, and killer robots. *North Carolina Journal of International Law and Commercial Regulation, 40*(4), 909–932.

Cubero, S. (Ed.). (2007). *Industrial robotics: Theory, modelling and control*. Mammendorf: Pro Literature Verlag Robert Mayer-Scholz.

Daugherty, P. R., & Wilson, H. J. (2018). *Human + machine: Reimagining work in the age of AI*. Boston, MA: Harvard Business Review Press.

DeCanio, S. J. (2016). Robots and humans—Complements or substitutes? *Journal of Macroeconomics, 49*, 280–291.

Drexler, K. E. (2013). *Radical abundance: How a revolution in nanotechnology will change civilization*. New York: Public Affairs.

Driessen, C., & Heutinck, L. F. M. (2015). Cows desiring to be milked? Milking robots and the co-evolution of ethics and technology on Dutch dairy farms. *Agriculture and Human Values, 32*(1), 3–20.

Dunis, C. L., Middleton, P. W., Karathanasopolous, A., & Theofilatos, K. A. (Eds.). (2017). *Artificial intelligence in financial markets: Cutting edge applications for risk management, portfolio optimization and economics*. London: Palgrave Macmillan.

Economic Times. (2018). Low skilled workers to face disadvantage as robotics, AI take over: ADB. *The Economic Times*, April 11. https://economictimes.indiatimes.com/jobs/low-skilled-workers-to-face-disadvantage-as-robotics-ai-take-over-adb/articleshow/63715526.cms

Ford, M. R. (2009). *The lights in the tunnel: Automation, accelerating technology and the economy of the future*. Acculant Publishing.

Ford, M. R. (2015). *Rise of the robots: Technology and the threat of a jobless future*. New York: Basic Books.

Frank, M., Roehring, P., & Pring, B. (2017). *What to do when machines do everything: How to get ahead in a world of AI, algorithms, bots and big data*. Hoboken, NJ: John Wiley & Sons, Inc.

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change, 114*, 254–280.

Giarini, O., & Liedtke, P. M. (1996). *The employment dilemma: The future of work*. The Club of Rome.

Gould, S. J. (1996). *The mismeasure of man* (Revised & Expanded). W. W. Norton & Company.

Harteis, C. (Ed.). (2018). *The impact of digitalization in the workplace: An educational view* (Vol. 21). Cham: Springer.

Hill, J., Ford, W. R., & Farreras, I. G. (2015). Real conversations with artificial intelligence: A comparison between human–human online conversations and human–chatbot conversations. *Computers in Human Behavior, 49*, 245–250.

Huang, M. H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research, 21*(2), 155–172.

Ivanov, S. (2016). Will robots substitute teachers? *Yearbook of Varna University of Management, 9*, 42–47.

Ivanov, S. (2017). Robonomics—Principles, benefits, challenges, solutions. *Yearbook of Varna University of Management, 10*, 283–293.

Ivanov, S., & Webster, C. (2017). *The robot as a consumer: A research agenda*. Paper presented at the "Marketing: Experience and perspectives" Conference, 29–30 June 2017, University of Economics, Varna, Bulgaria, pp. 71–79.

Ivanov, S., & Webster, C. (2018). Adoption of robots, artificial intelligence and service automation by travel, tourism and hospitality companies—A cost-benefit analysis. In V. Marinov, M. Vodenska, M. Assenova, & E. Dogramadjieva (Eds.), *Traditions and innovations in contemporary tourism* (pp. 190–203). Newcastle upon Tyne: Cambridge Scholars Publishing.

Ivanov, S., Webster, C., & Berezina, K. (2017). Adoption of robots and service automation by tourism and hospitality companies. *Revista Turismo & Desenvolvimento, 27/28*, 1501–1517.

Kaplan, J. (2015). *Humans need not apply. A guide to wealth and work in the age of artificial intelligence*. Yale University Press.

Kauhanen, A. (2016). *The future of work: Challenges for men and women*. ETLA Brief No. 50. Retrieved from http://pub.etla.fi/ETLA-Muistio-Brief-50.pdf

Kaur, S. (2012). How medical robots are going to affect our lives. *IETE Technical Review, 29*(3), 184–187.

Kiggins, R. (2018). *The political economy of robots: Prospects for prosperity and peace in the automated 21st century*. Cham: Palgrave Macmillan.

Kurzweil, R. (2005). *The singularity is near. When humans transcend biology*. London: Duckworth Overlook.

LaGrandeur, K., & Hughes, J. J. (Eds.). (2017). *Surviving the machine age. Intelligent technology and the transformation of human work*. London: Palgrave Macmillan.

Lee, J. (2017). *Sex robots: The future of desire*. London: Palgrave Macmillan.

Lehman, T. (2015). Countering the modern Luddite impulse. *The Independent Review, 20*(2), 265–283.

Leonhard, G. (2016). *Technology vs. humanity*. Fast Future Publishing.

Lin, P., Abney, K., & Bekey, G. A. (2014). *Robot ethics: The ethical and social implications of robotics*. Cambridge, MA: The MIT Press.

Low, K.-H. (Ed.). (2007). *Industrial robotics: Programming, simulation and applications*. Mammendorf: Pro Literature Verlag Robert Mayer-Scholz.

Makridakis, S. (2017). The forthcoming Artificial Intelligence (AI) revolution: Its impact on society and firms. *Futures, 90*, 46–60.

Maurer, M., Gerdes, J. C., Lenz, B., & Winner, H. (Eds.). (2016). *Autonomous driving: Technical, legal and social aspects*. Berlin and Heidelberg: Springer Open.

McClure, P. K. (2017). "You're fired," says the robot: The rise of automation in the workplace, technophobes, and fears of unemployment. *Social Science Computer Review* (forthcoming). https://doi.org/10.1177/0894439317698637

Miller, M. R., & Miller, R. (2017). *Robots and robotics: Principles, systems, and industrial applications*. McGraw-Hill Education.

Min, H. (2010). Artificial intelligence in supply chain management: Theory and applications. *International Journal of Logistics Research and Applications, 13*(1), 13–39.

Mirheydar, H. S., & Parsons, J. K. (2013). Diffusion of robotics into clinical practice in the United States: Process, patient safety, learning curves, and the public health. *World Journal of Urology, 31*(3), 455–461.

Murphy, J., Hofacker, C., & Gretzel, U. (2017). Dawning of the age of robots in hospitality and tourism: Challenges for teaching and research. *European Journal of Tourism Research, 15*, 104–111.

Neapolitan, R. E., & Jiang, X. (2013). *Contemporary artificial intelligence*. Boca Raton, FL: CRC Press.

Nørskov, M. (2016). *Social robots: Boundaries, potential, challenges*. London and New York: Routledge.

NPR. (2011). Science diction: The origin of the word 'robot'. Retrieved from https://www.npr.org/2011/04/22/135634400/science-diction-the-origin-of-the-word-robot

NPR. (2015). Map: The most common job in every state. Retrieved from https://www.npr.org/sections/money/2015/02/05/382664837/map-the-most-common-job-in-every-state

Organisation for Economic Co-operation and Development/OECD/. (2017). *The next production revolution: Implications for governments and business*. Paris: OECD Publishing.

Pires, J. N. (2007). *Industrial robots programming: Building applications for the factories of the future*. New York: Springer US.

Preparing for The Future of Artificial Intelligence. (2016). Published by the Executive Office of the President of the USA, National Science and Technology Council and the Committee on Technology, 2nd October 2016. Retrieved October 14, 2016, from https://www.whitehouse.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf

Remus, D. & Levy, F. (2015) Can robots be lawyers? Computers, lawyers, and the practice of law. *SSRN Working paper*. Retrieved from http://ssrn.com/abstract=2701092

Royakkers, L., & van Est, R. (2016). *Just ordinary robots: Automation from love to war*. Boca Raton: CRC Press.

Russell, S. J., & Norvig, P. (2016). *Artificial intelligence: A modern approach*. Harlow: Pearson Education Limited.

Santens, S. (2017). Unconditional basic income as a solution to technological unemployment. In K. LaGrandeur & J. J. Hughes (Eds.), *Surviving the machine age. Intelligent technology and the transformation of human work* (pp. 107–116). London: Palgrave Macmillan.

Schommer, E., Patel, V. R., Mouraviev, V., Thomas, C., & Thiel, D. D. (2017). Diffusion of robotic technology into urologic practice has led to improved resident physician robotic skills. *Journal of Surgical Education, 74*(1), 55–60.

Schwab, K. (2016). *The fourth industrial revolution*. Cologny and Geneva: World Economic Forum.

Sendler, U. (Ed.). (2018). *The internet of things: Industrie 4.0 unleashed*. Berlin: Springer Vieweg.

Shanahan, M. (2015). *The technological singularity*. Cambridge, MA: The MIT Press.

Sheahen, A. (2012). *Basic income guarantee: Your right to economic security*. New York: Palgrave Macmillan.

Skilton, M., & Hovsepian, F. (2018). *The 4th industrial revolution: Responding to the impact of artificial intelligence on business*. Cham: Palgrave Macmillan.

Solon, O. (2016, June 17). Self-driving trucks: What's the future for America's 3.5 million truckers? *The Guardian*. Retrieved from https://www.theguardian.com/technology/2016/jun/17/self-driving-trucks-impact-on-drivers-jobs-us

Sparrow, R. (2007). Killer robots. *Journal of Applied Philosophy, 24*(1), 62–77.

Stern, A. (2016). *Raising the floor: How a universal basic income can renew our economy and rebuild the American dream*. Public Affairs.

Stevens, Y. A., & Marchant, G. E. (2017). Policy solutions to technological unemployment. In K. LaGrandeur & J. J. Hughes (Eds.), *Surviving the machine age. Intelligent technology and the transformation of human work* (pp. 117–130). London: Palgrave Macmillan.

Swan, M. (2017). Is technological unemployment real? An assessment and a plea for abundance economics. In K. LaGrandeur & J. J. Hughes (Eds.), *Surviving the machine age. Intelligent technology and the transformation of human work* (pp. 19–33). London: Palgrave Macmillan.

Talwar, R. (Ed.). (2015). *The future of business*. Fast Future Publishing.

Talwar, R., Wells, S., Whittington, A., Koury, A., & Romero, M. (2017). *The future reinvented. Reimagining life, society, and business*. Fast Future Publishing.

The National Artificial Intelligence Research and Development Strategic Plan. (2016). Published by the Executive Office of the President of the USA, National Science and Technology Council and the Committee on Technology,

2nd October 2016. Retrieved October 14, 2016, from https://www.whitehouse.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/national_ai_rd_strategic_plan.pdf

Timms, M. J. (2016). Letting artificial intelligence in education out of the box: Educational cobots and smart classrooms. *International Journal of Artificial Intelligence in Education, 26*(2), 701–712.

Warwick, K. (2012). *Artificial intelligence: The basics*. Oxon: Routledge.

Watson, P. (2018, August 23). There's only one way to solve our stagnant growth problem. *Forbes*. Retrieved from https://www.forbes.com/sites/patrickwwatson/2018/08/23/theres-only-one-way-to-solve-our-stagnant-growth-problem/#60cb57462e1d

West, D. M. (2018). *The future of work: Robots, AI, and automation*. Washington: Brookings Institution Press.

Wiener, A. (2017, November 29). Inside Adidas' robot-powered, on-demand sneaker factory. *Wired*. https://www.wired.com/story/inside-speedfactory-adidas-robot-powered-sneaker-factory/

Wirtz, J., Patterson, P., Kunz, W., Gruber, T., Lu, V. N., Paluch, S., & Martins, A. (2018). Brave new world: Service robots in the frontline. *Journal of Service Management, 29*(5), 907–931. in press.

World Economic Forum. (2016). *The future of jobs. Employment, skills and workforce strategy for the fourth industrial revolution*. World Economic Forum report. Retrieved from http://www3.weforum.org/docs/WEF_Future_of_Jobs.pdf

World Economic Forum. (2018). *The future of jobs report 2018*. World Economic Forum report. Retrieved from http://www3.weforum.org/docs/WEF_Future_of_Jobs_2018.pdf

Xu, A., Liu, Z., Guo, Y., Sinha, V., & Akkiraju, R. (2017). A new chatbot for customer service on social media. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3506–3510). New York: ACM.

CHAPTER 9

The Dark Side of Big Data: Personal Privacy, Data Security, and Price Discrimination

Yang Liu and Connor Greene

Introduction

According to Furman and Simcoe (2015), the Council of Economic Advisers released a study, "Big Data and Differential Pricing", in February 2015. In terms of this study (2015), big data is defined as the capability to collect a huge amount of data from various sources for investigating, analyzing, and understanding consumers, and forecasting market demand and affordable prices for consumers. In turn, the way firms do business has been changed by this new technology. Consequently, the whole of society has been changed, with this new technology both bringing benefits and generating challenges.

Y. Liu (✉)
Fitchburg State University, Fitchburg, MA, USA
e-mail: yliu4@fitchburgstate.edu

C. Greene
Southern New Hampshire University, Hooksett, NH, USA
e-mail: connor.greene@snhu.edu

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_9

Benefits and Challenges to Business and Society since the Big Data Era

Personal Privacy and Data Security

When the big data era finally arrives, firms will cheer up because of this smart technology for which they have been waiting for years. Web browsers or the owners of websites normally generate a "cookie", which is a file on a user's computer. This file saves all the information on the user. Therefore, web browsers or the owners of websites can draw a picture from the details of any user based on the information they accessed. TRUSTe, a technology compliance and security company, reported in 2014 that the top 100 websites browsed the most are monitored by more than 1300 firms. Some of the data collected by these firms is "shared" (also known as "piggybacking") with outsiders. Normally this monitoring process is unknown to web users. For example, an advertising technology company, RadiumOne, puts cookies on users without notifying them, when they click on a link sent by a friend (Big Data and Differential Pricing, 2015; Chen, Choe, & Matsushima, 2018; The Economist, 2014).

New information technology platforms offer applications (or apps) for maps, search engines, social media, and so on to enrich people's daily lives. In turn, the omnipresent online search, payment transactions, and social interactions leave a trail of personal information, containing emotions, tastes, behaviors, locations, affordable prices, and so on. These apps do not support cookies, but they can access and record the unique device identification (UDID) number of the mobile device to recognize individual users. PubMatic is a company helping publishers increase sales. It provides around 50 to 70 points of user information on desktops and around 100 on mobiles, which includes the mobile device user's location (Chen et al., 2018; The Economist, 2014).

New Business Pattern

According to Kamenica, Mullainathan, and Thaler (2011), the more information cell-phone providers and data collectors gather and analyze, the more they understand consumer behavior. Given enough time to collect enough data, data analytics can understand a consumer more than him- or herself. Calo (2014) explains personalized advertising as the ability of firms to detect when is the best time to deliver a commercial or how

to build an emotional relationship with consumers based on their personal experience. Duhigg (2012) observes that Target can distinguish pregnant consumers based on their shopping record and directly deliver related advertisements regarding baby products to them.

These big data enable firms to reach their target consumers directly and save millions of dollars that they are wasting on people who will not purchase their goods or services no matter how hard they try, because the firms are able to analyze consumers' current situation, such as pregnancy, divorce, relocation, and so on. Before the big data era, marketing researchers would design, deliver, collect, and analyze surveys to understand their potential consumers. This process took months, even years, to provide a rough understanding based on a sample group that was never big enough. All business plans and marketing designs would be built on that information, until the day that big data ended it.

Big data change the way firms do business, make marketing plans, and deliver advertisements. With big data, firms can directly locate a potential consumer, deliver an advertisement at an appropriate time, and charge an affordable price to make a deal. For example, when consumers search for a business suit online, advertisements do not pop up randomly but in a logical layout regarding what they purchased before. This technology matches consumers with the goods or services they are looking for. Consumers can benefit from this personalized advertising as well, since it saves them time. They can find what they need more quickly.

Price Discrimination

On the other hand, big data bring new challenges to this world as well. There are some side effects accompanying big data. Nowadays we face new situations that we could not imagine years ago, such as personal privacy, data security, and price discrimination.

Big Data and Differential Pricing (2015) points out that one main concern with big data is price discrimination, which means charging different consumers different prices for similar products. Price discrimination can be divided into three types. First-degree price discrimination is personalized pricing, meaning there is no fixed sales price. Every consumer can negotiate the price, and they eventually pay different prices. Second-degree price discrimination is quantity discounts. In other words, buy more, save more. Third-degree price discrimination is a demographic discount, which means a specific group of people can get specific treat-

ment. For example, universities provide students who serve in the military with a specific scholarship or free tuition (Big Data and Differential Pricing, 2015).

In some cases, even though price discrimination sounds unfair, it is not, due to the factual situation. For example, children normally can get a specific discount, not only because they are in a specific age range but also because they increase the expenditure with their entire family. For example, parents have to accompany their children to the Disneyland, the cinema, the arcade, and so on. Therefore, children stimulate demand for certain goods and services that would not otherwise be consumed by adults. An airplane ticket is more expensive when the booking day is closer to the departure day, due to the increase in overall demand on that given date, as well as the finite number of available seats.

In some other cases, price discrimination may be really unfair. According to Townley, Morrison, and Yeung (2017), Amazon.com sold DVDs to different people for different prices in 2000. After a public protest, Amazon refunded the consumers who were charged a higher price and claimed that this was a test. However, a similar issue happened again in 2012. A consumer found that the price of a set of mahjong tiles in her shopping basket suddenly jumped from $54.99 to $79.99; when she emptied the basket and added the same goods into it again, the price went back to $59.99. The *Wall Street Journal* reported in 2012 that Staples Inc. was displaying different prices to people based on their location. Staples.com showed a lower price to consumers who lived within 20 miles of a rival store (Townley et al., 2017). According to Michael and Mohammed (2015), in international travel bookings for US hotel and car rental, the situation is the same.

According to Newcomer (2017), Uber Technologies Inc. uses a "route-based pricing" system, which charges consumers differently depending on the expectation of how much they would like to pay. This system is different from the traditional pricing model, which charges consumers only based on the quantity of a service they purchase. The pricing system Uber used before calculates fares by adding up time, mileage, and multipliers based on geographic demand. The new "route-based pricing" system charges some consumers a higher price not because they ride to a farther location, but because it is considered they would pay a higher price.

A similar situation was reported by Consumer Intelligence Research Partners (CIRP) in 2013: that among all smartphone consumers, iPhone consumers were charged the highest average monthly carrier bill in the

USA. CIRP co-founder Michael Levin believes this situation happens not because consumer behaviors are different among smartphone users, but because of their data plans and carriers (Keller, 2013; Paczkowski, 2013). In other words, iPhone users are charged a higher price for their carrier bills in the USA, not because the same apps cost more in data streaming than on an Android phone, nor because iPhone users spend longer on their phone than Android users, but possibly because they are considered richer, since the price of an iPhone is higher than that of an Android phone.

More evidence can be observed internationally. For example, according to Wu (2018), among 2008 participants, 51.3% suffered price discrimination when they purchased tickets, booked hotels, and subscribed for memberships online in China. There are several different types of price discrimination. (1) More refreshes, higher price. When booking flight tickets and hotels, the more times consumers refresh the website, the higher the price they will see. Once consumers clean the website's cache, cookies, and history, the price decreases. (2) Members pay more. Consumers may purchase a membership for movie ticket discounting, but some of them find that their price is higher than for consumers who do not have a membership. (3) Consumers pay a higher price on iPhone than on Android. Many iPhone consumers are charged a higher price when they subscribe to a membership or pay a taxi. Some firms explain this situation by saying that mobile payment services (Apple Pay, Alipay, etc.) offer convenience to their clients and users while charging a higher service fee on the iPhone than on Android.

Machlup (1955) may not be the first economist who has noticed that the definition of price discrimination ("the practice of a firm selling a homogeneous commodity at the same time to different purchasers at different prices") is somewhat vague, but he is the first to clarify this definition as "the practice of a firm or group of firms of selling (leasing) at prices disproportionate to the marginal costs of the products sold (leased) or of buying (hiring) at prices disproportionate to the marginal productivities of the factors bought (hired)". Moreover, according to Thomas (2012), in the insurance context the concept of price discrimination, which is selling products with similar marginal costs at different prices to different clients, is "non-risk" price discrimination. However, insurance is a risk-related product. Various price segments of it exhibit different marginal costs. This is why the data on (potential) consumers' health is so important for both insurance companies and consumers. With the data, the insurance company can charge a consumer a personalized price which may be signifi-

cantly different from the price without the data. Consumers may even do not know that their price is personalized and higher than that for others. Big Data and Differential Pricing (2015) explains the risk-based pricing policy of insurance companies as finally helping the whole of society by charging drivers with risky behaviors or a history of traffic accidents a higher price. Therefore, these drivers will change their behaviors and society will become safer for everyone.

It is not hard to understand that firms have plenty of reasons to treat their clients differently. This is considered necessary and motivation for them to pay for big data, especially service firms. Different from goods companies, service companies have a long-term relationship with their clients. They provide services constantly, which means the relationship between a service company and a consumer keeps changing with time. For instance, insurance companies service their clients from cradle to grave. Some clients have good health while some others do not. Therefore, insurance companies could separate them into different segments (high risk, moderate, low risk) and foresee the future revenue compared to the expected underwriting expense. Thus, insurance companies charge different prices to match the health situation of their clients.

Discussion and Conclusion

Podesta, Pritzker, Moniz, Holdren, and Zients (2014) believe that it deserves more serious concern when data impact decisions on individuals' credit, housing, education, healthcare, or employment. Moreover, personal privacy is another critical concern (Big Data and Differential Pricing, 2015). According to Payton, Schmidt, and Claypoole (2014), the core of privacy is not about information but about choice. According to Hacker and Petkova (2017), both greater social equality and inequality have been generated with big data. Because of big data, technology, education, and business are more efficient than before. However, consumer protection and discrimination bring greater inequality to society. Big data provide firms with opportunities for personalizing services. Before the new big data era, consumers were exposed to the same advertisements. They had an equal chance with the same offers at the same price. However, big data collection and analytics are changing this situation. With the new technology, firms know exactly what consumers want and how much they can afford. On the other hand, consumers do not know that they are being treated differently from one another. In other words, with detailed data,

companies can manipulate their consumers individually. Consumers are guided by an invisible hand to a place they think will satisfy their needs. However, they may pay a higher price or make a suboptimal decision because of firms obtaining their data and using them inappropriately. Big Data and Differential Pricing (2015) suggests that enhancing the transparency of how firms collect, use, and trade consumers' data would promote competition in the market and provide consumers with choices. In addition, Hacker and Petkova (2017) point out that transparency is important during this new big data era. Citron and Pasquale (2014) state that people have the right to access and correct their information and they have the right to know how they are rated and ranked as well. In other words, consumers have the right to keep their privacy, especially when they do not buy insurance from certain insurance companies. In addition, even though it may be necessary for insurance companies to obtain these data to lower their risks and costs, do they have the right to do so?

First of all, who owns the data? Or we may consider this from another angle: where is the line drawn when collecting data without notifying people that their data are collected, retained, and will be used to manipulate them later? Before the big data era, when companies had to collect data by sending out surveys, they had to send an obvious notification to the participants regarding this concern. They had to tell the participants it was voluntary. In some cases, they paid participants to collect their data. Moreover, normally, the participants were anonymous. They did not provide their personal information, such as name, date of birth, and address. Therefore, they did not need to worry about how these data would be used. However, the new technology changed this. The new technology enables web browsers and apps to collect and retain personal data on a continual basis. Because of this process, eventually, consumers get personalized advertisements. In some cases, this is harmless. On the contrary, it is convenient. On the other hand, it is a trap. Under the manipulation of a marketing strategy provided to companies based on specific data analytics, consumers purchase something that they do not need or are charged a higher price just because they are (considered) richer than others.

Consumers may not know that their data are being collected and their behaviors are being manipulated. They may enjoy the convenience. For example, a pregnant woman does need to prepare for the upcoming baby, and advertisements and coupons for baby products are useful to her. Or for a resident who has just moved to a new apartment, information about furniture discounts and internet services is helpful.

However, price discrimination is not a convenient experience. For example, ticket selling or hotel booking websites manipulate consumers by increasing prices according to their refresh frequency. This manipulates consumers to close the deal as soon as possible by implying that the price will keep increasing or that tickets are selling out. Consumers waste money as a result of their actions being recorded and analyzed by data analytics.

In conclusion, we believe that transparency of data collection is important and necessary. Individuals have the right to access data held on them and to have greater control over their own information. They should be notified when their information is collected and traded. They should have the right to say no to certain firms whose business is to dig, gather, collect, and analyze data. Furthermore, firms should be restricted by regulations regarding their operations. On the other hand, consumers who would like to enjoy personalized services, such as advertisements or coupons, can choose to allow data-gathering firms to collect their data with their permission.

References

Big Data and Differential Pricing. (2015). *The White House President Barack Obama*. Retrieved from https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/docs/Big_Data_Report_Nonembargo_v2.pdf

Calo, R. (2014). Digital market manipulation. *George Washington Law Review, 82*(4), 995–1051. 57p. Database: Business Source Ultimate.

Chen, Z., Choe, C., & Matsushima, N. (2018). Competitive personalized pricing. *ISER Discussion Paper*. Retrieved from https://www.monash.edu/business/economics/research/publications/publications2/0218Copetitivechenchoe.pdf

Citron, D. K., & Pasquale, F. A. (2014). The scored society: Due process for automated predictions. *Washington Law Review, 89*(1), 33. Washington Law Review Association.

Duhigg, C. (2012). How companies learn your secrets. *New York Times Magazine*, February 16. Retrieved from https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html

Furman, J., & Simcoe, T. (2015). The economics of big data and differential pricing. *The White House President Barack Obama*. Retrieved from https://obamawhitehouse.archives.gov/blog/2015/02/06/economics-big-data-and-differential-pricing

Hacker, P., & Petkova, B. (2017). Reining in the big promise of big data: Transparency, inequality, and new regulatory frontiers. *Northwestern Journal of Technology & Intellectual Property, 15*(1), 1–42. Northwestern University.

Kamenica, E., Mullainathan, S., & Thaler, R. (2011). Helping consumers know themselves. *The American Economic Review, 101*(3), 417–422. American Economic Association.

Keller, J. (2013). iPhone users pay higher cell phone bills than any other smartphone user. *iMore*. Retrieved from https://www.imore.com/iphone-users-biggest-cash-cows-carriers

Machlup, F. (1955). Characteristics and types of price discrimination. NBER chapters. In *Business concentration and price policy* (pp. 397–440). National Bureau of Economic Research, Inc.

Michael, R., & Mohammed, R. (2015). Who's paying more to tour these United States? Price differences in international travel bookings. *Technology Science*.

Newcomer, E. (2017, May 19). Uber starts charging what it thinks you're willing to pay. *Bloomberg*. Retrieved from https://www.bloomberg.com/news/articles/2017-05-19/uber-s-future-may-rely-on-predicting-how-much-you-re-willing-to-pay

Paczkowski, J. (2013). iPhone users rack up the highest carrier bills. *All Things D*. Retrieved from http://allthingsd.com/20130130/wireless_bills_by_os_android_ios/

Payton, T., Schmidt, H. A., & Claypoole, T. (2014). *Privacy in the age of big data: Recognizing threats, defending your rights, and protecting your family*. Lanham, MD: Rowman & Littlefield Publishers.

Podesta, J., Pritzker, P., Moniz, E. J., Holdren, J., & Zients, J. (2014). Big data: Seizing opportunities, preserving values. *Executive Office of the President*. Retrieved from https://obamawhitehouse.archives.gov/sites/default/files/docs/20150204_Big_Data_Seizing_Opportunities_Preserving_Values_Memo.pdf

The Economist. (2014, September 11). Getting to know you. Retrieved from https://www.economist.com/special-report/2014/09/11/getting-to-know-you

Thomas, R. (2012). Non-risk price discrimination in insurance: Market outcomes and public policy. *The Geneva Papers on Risk and Insurance. Issues and Practice, 37*(1), 27–46. Palgrave Macmillan.

Townley, C., Morrison, E., & Yeung, K. (2017). Big data and personalized price discrimination in EU competition law. *Yearbook of European Law, 36*(1), 683–748. 66p. Publisher: Oxford University Press/USA.

Wu, Y. Q. (2018). 用苹果手机打车、看电影、买机票比安卓贵?记者亲测被大数据"杀熟". *Shen Zhen Shang Bao*. Retrieved from http://www.sohu.com/a/226569266_355791

CHAPTER 10

Digital Finance for Financial Inclusion and Inclusive Growth

Md. Nur Alam Siddik and Sajal Kabiraj

Introduction

The notion of financial inclusion emerged in early 2000 and thus was given policy priority by governments and international organizations to reduce poverty and achieve inclusive growth. Consequently, governments and policy makers formulated policies to foster financial inclusion in developing and emerging economies (GPFI, 2010). Nowadays, digital finance has gained much attention from policy makers as a mechanism for reducing poverty and achieving financial stability, because digital finance expands financial inclusion, which ultimately promotes financial stability.

Digital finance has numerous benefits to all stakeholders, including financial service users, digital finance service providers, governments, policy makers and the economy. For example, it expands access to finance among the poor and disadvantaged, it can reduce the cost of service for

Md. N. A. Siddik (✉)
Begum Rokeya University, Rangpur, Bangladesh
e-mail: nasiddik@brur.ac.bd

S. Kabiraj
Dongbei University of Finance and Economics, Dalian, China
e-mail: skabiraj@dufe.edu.cn

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_10

banks and fintech suppliers by employing a branchless technology-based banking system, and it can lead to a rise in total expenditure for governments. The World Bank Group President Jim Yong Kim indicated that by 2020, digital financial instruments such as e-money cards, debit cards and low-cost bank accounts will significantly expand the scope of financial access for excluded people.[1] Notwithstanding its benefits, the issue of digital finance, financial inclusion and financial stability has not been addressed in the literature. This study aims to fill this gap by addressing the link among digital finance, financial inclusion and financial stability.

This research contributes to the financial inclusion campaigns initiated by the World Bank as an operational solution for poverty eradication in developing and emerging economies. The findings of this study can help policy makers to understand the issues related to the prompt expansion of digital financial services, strategies for its effective delivery to the poor and the risks involved in digital financial inclusion. Secondly, this study adds to the existing literatures on emerging digital finance and financial inclusion, which will enable academics and researchers to attend to the effects of digital financial services on financial inclusion.

Conceptualizing Digital Finance

Manyika, Lund, Singer, White, and Berry (2016) defined digital finance as financial services delivered through a digital infrastructure such as a mobile phone and internet network, which eventually encourage less usage of cash and the traditional brick-and-mortar banking system.

According to Gomber, Koch, and Siering (2017), digital finance comprises a magnitude of new financial products, financial services, financial management software and new ways to communicate and interact with customers. Digital finance implies a procedure of a certain blend of related money management tools that are conveyed using a portable or Web-based system (Peake, 2012).

Klapper, El-Zoghbi, and Hess (2016) demonstrate that digital finance permits people to access finance from distant places and at any time, which enables them to manage their funds well at times of need and thus shrinks the possibility that they will fall into poverty.

While there is no universal description of digital finance, there is a consensus that it incorporates financial products, financial services, technolo-

[1] https://www.worldbank.org/en/news/speech/2016/04/14/remarks-by-world-bank-group-president-jim-yong-kim-at-the-global-connect-initiative

gies and infrastructure that empower people to have access to payments, savings and credit facilities along with other required financial services through online means and without a physical visit to a bank branch.

Conceptualizing Financial Inclusion

Financial inclusion aims to ensure access to and use of financial services such as accounts, savings and credit for everyone, including the poor and disadvantaged. It targets the unbanked, people who have not been offered financial services, bringing them under the umbrella of the formal financial system, and the under-banked, people who have access to but have not used financial services. The overriding goal is to create an environment in which all people will have easy access to financial services to ensure the participation of everyone in the course of economic growth and receiving the benefits of economic growth accordingly.

Siddik, Sun, and Kabiraj (2015) describe financial inclusion as a means of providing easier and timely access to finance to all sections of society, including disadvantaged groups, which in turn will promote inclusive growth.

Hannig and Jansen (2010) state that financial inclusion programs aim to offer greater opportunity with easier access to proper financial services to people, including disadvantaged groups, which in turn fosters the economic development of a country. Sarma (2015) defines financial inclusion as a procedure which warrants easier access, readiness and usage of financial services for all people of a country. This explanation strictly emphasizes access to and usage of formal financial services for entire segments of society, which is one of the most vital factors in the economic growth of a nation. Following the definition of Sarma (2015), we focus on the access and usage dimensions of financial inclusion and define it as a process which aims to expand the financial network among people including the unbanked and the under-banked. In developing economies, the banking sector is reflected as the key pillar and the most contributive sector to economic growth. Thus, we focus on financial inclusion by commercial banks.

Effect of Digital Finance on Financial Inclusion

The impact of digital finance on financial inclusion is now evident. Digital financial services have the substantial prospect of magnifying the delivery of financial services to a large segment of society, including the poor, through innovative technologies such as mobile banking solutions,

e-money systems and digital payment infrastructure at a reasonable cost, in appropriate modes and in a safe environment. According to the 2016 annual report of the International Telecommunication Union (ITU), 95% of the world population lives in areas which are covered by a mobile cellular signal and in 2016, 47% of the world population was online, which indicates progress, but at the same time over half of the world's population, or roughly 3.9 billion people, was not online. This suggests that there is huge potential to settle financial matters digitally.

Digital finance adds new stratums to financial inclusion, proposing new methods of expanding the inclusion of groups profiled as poor and disadvantaged in mainstream financial activities (Gabor & Brooks, 2017).

Manyika et al. (2016) point out that the wide-ranging usage of digital finance could increase the annual gross domestic product (GDP) of all economies by 6% by 2025. Arguing the same, Ghosh (2016) employs data from 2001–2012 on Indian states and observes that mobile telephony as a means of digital financial services has significant effects on economic growth. Proposing similar contributive effects of digital finance on financial inclusion, we first examine world data and then provide case studies on India.

Methodology

To determine whether digital finance promotes financial inclusion, in this section we present the effects of digital finance on financial inclusion. In this study financial inclusion is measured by outstanding loans with commercial banks (% of GDP; log of outstanding loans or LOL) and digital finance is measured by three widely used proxies: the number of automated teller machines (ATMs) per 100,000 adults (ATM); the percentage of the population using the internet (INT); and mobile money transaction value (percentage of GDP; MMT). Based on the availability of data, we collected data from 2004 to 2016 on 189 countries. Variable measurements and data sources are presented in Table 10.1. Taking the average of each variable for each year, firstly we conducted trend analysis and secondly we produced two case studies in support of our hypothesis.

Empirical Findings

We divide this section into two parts. In the first part we conduct a trend analysis with linear regression, and in the second part we provide case studies. In Figs. 10.1, 10.2 and 10.3 we represent the financial inclusion

Table 10.1 Variables, their proxies, legend and source

Variable	*Measurement*	*Legend*	*Sources*
Financial inclusion	Natural log of outstanding loans with commercial banks (% of GDP)	LOL	FAS*
Digital finance	Natural log of number of automated teller machines (ATMs) per 100,000 adults	ATM	FAS
	Individuals using the internet (% of population)	INT	WDI**
	Mobile money transactions: value (% of GDP)	MMT	FAS

* Financial Access Survey
** World Development Indicator

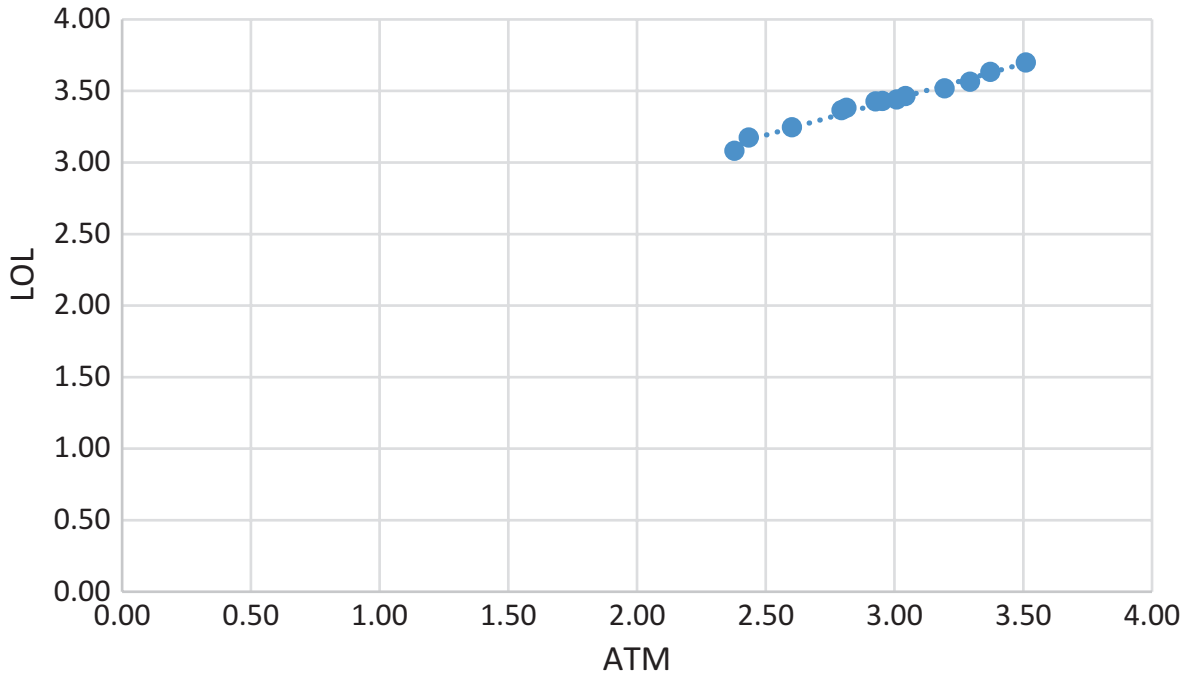


Fig. 10.1 Impact of ATM on LOL

variable, LOL, on the Y axis and digital finance proxies on the X axis. These figures represents the association of digital finance proxies (ATM, INT and MMT) with financial inclusion, LOL. Figure 10.1 depicts an upward, straight trend line between the number of ATMs per 100,000 adults and outstanding loans with commercial banks, which means that there is a positive relation between ATM and LOL.

Figure 10.2 exhibits the relation between internet penetration and financial inclusion. With linear trend analysis, we see that the higher the number of internet users, the higher the degree of financial inclusion. In other words, there is a positive effect of INT on LOL.

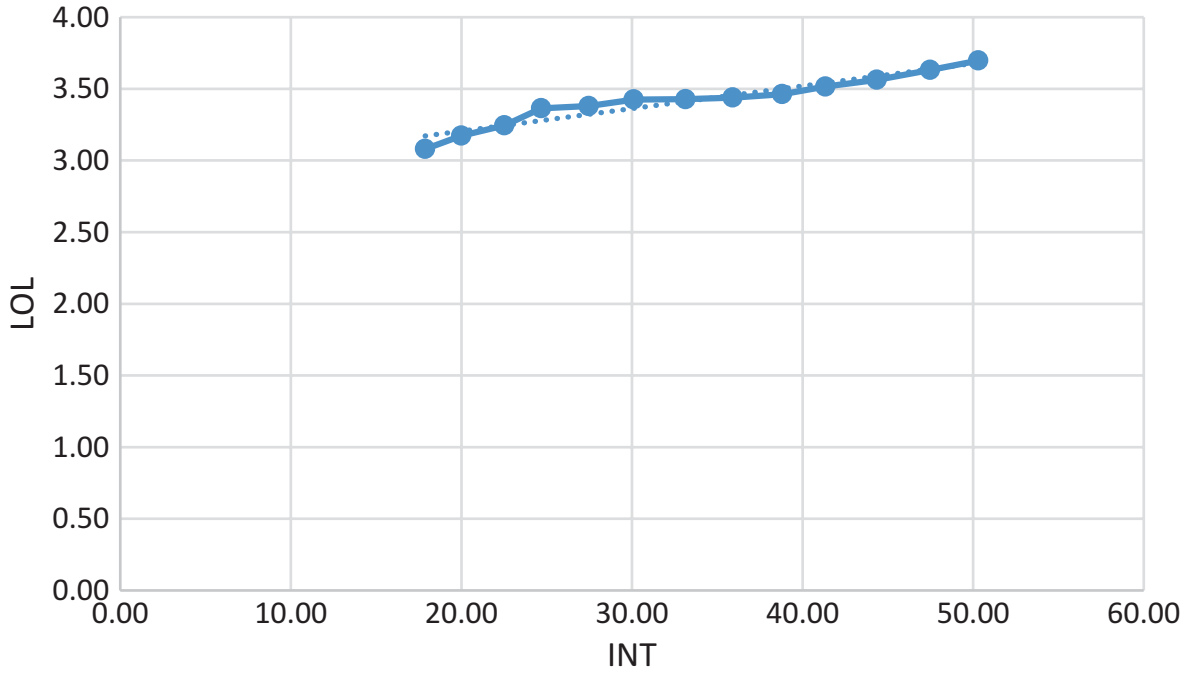


Fig. 10.2 Impact of INT on LOL

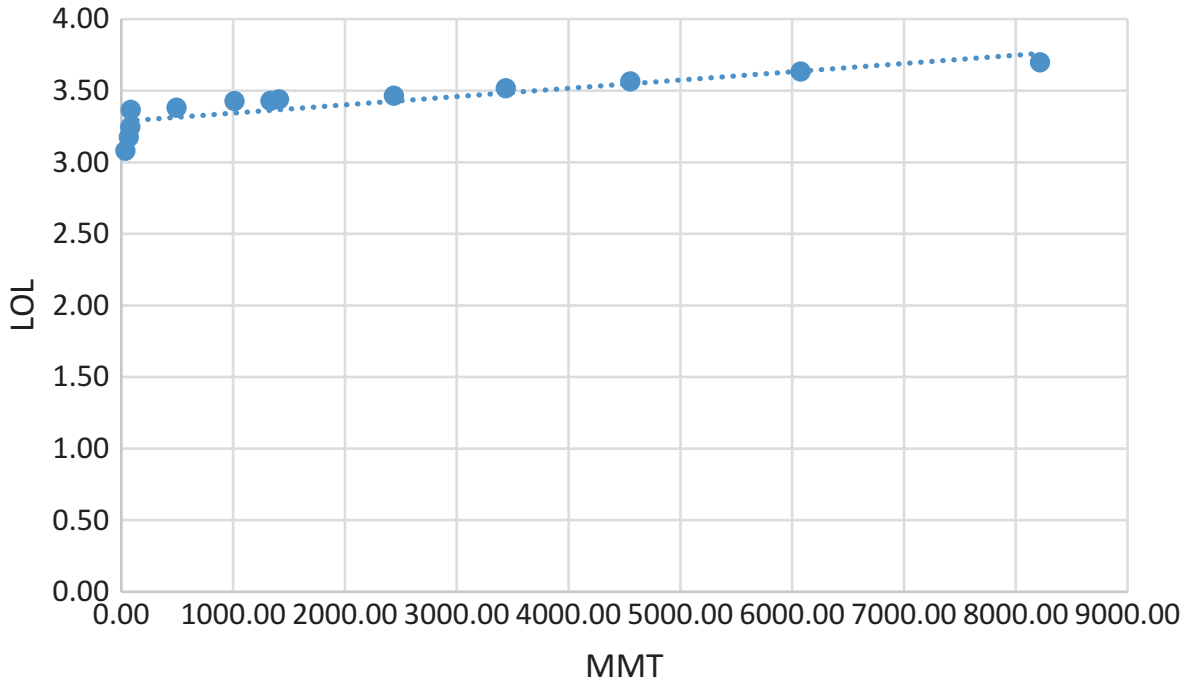


Fig. 10.3 Impact of MMT on LOL

The third proxy we use is the value of mobile money transactions. Since mobile money is one of the most widely used instruments of digital finance which has great global penetration, we have used this proxy to examine whether this contributes to financial inclusion or not. The results of the trend analysis indicate that there is positive movement between MMT and LOL. Thus, we argue that digital finance promotes financial inclusion. Our findings are similar to those of Ozili (2018).

Case Study

In this section we look at India, which is well and truly in the digital era and is a good example of digital financial inclusion. The essence of the case study is the use of technology to support the country's flagship benefit transfer programs and open doors for the poor to the financial mainstream. India is a leading case study of commitment to digital transformation and financial inclusion. With over a billion mobiles, 325 million broadband connections and 306 million new bank accounts, the country has become a case study in digital financial inclusion. The 500-million mobile connections in rural areas offer an opportunity to achieve full financial inclusion.

To make this happen, India has been launching strategic innovations to improve the design and drive a faster rollout of large benefit programs. The innovations pointedly track the delivery of entitlements to beneficiaries and help engage with the excluded millions using mobile-based tools. India's Digital Financial Inclusion program has piloted an integrated cashless digital platform to bring the farming community to center stage, while supporting a technological solution to fund and track rural projects. Another pilot infuses transparency into the delivery of nutrition across anganwadis, or rural childcare centers. A notable new entry is a technology platform to support a nationwide welfare scheme for pregnant and lactating mothers.

The Digital Financial Inclusion program uses insights drawn from work in both the corporate social responsibility (CSR) and government spaces to develop collaborative information technology (IT) platforms that bring corporates and governments together for community development. Moving upstream to benefit design, the program has also deployed analytical tools that use complex information on existing government schemes and databases to help design new programs.

Together with non-governmental organizations (NGOs) and the government, India's Digital Financial Inclusion program collaborates,

conceptualizes and develops innovative solutions for enabling direct benefit transfer (DBT), financial inclusion and financial literacy. The program conducts research studies to understand and analyze factors that could possibly lead to the greater adoption and use of digital services by the poor. Postal accounts offering digital bank accounts could lead the way in this. Recently, the India Post Payments Bank (IPPB) has offered facilities for deposits, money transfers, DBTs and bills and utility payments, among other things.

Case 1: India Post Payments Bank: A Case Study of Digital Financial Inclusion

IPPB offers three types of zero-balance savings accounts: regular savings, digital savings and basic savings bank deposit accounts.

October 9 is celebrated as World Post Day. It also marks the beginning of National Postal Week, which is celebrated until October 15. India Post, which runs postal services in the country, sent greetings to its followers on the occasion, as did the government to its employees. India Post, which has a network of 155,015 post offices, also started the India Post Payments Bank in the year of 2018. IPPB has 650 branches and 3250 access points.

IPPB offers three types of savings account: regular savings account, digital savings account and basic savings bank deposit account.

- *India Post Payments Bank regular savings account*

This regular savings account allows for unlimited cash deposits and withdrawals. The zero balance account offers an interest rate of 4% per annum on end-of-day balances. The interest is paid on a quarterly basis.

This account does not require any monthly average balance and it can be opened with a zero balance. The maximum end-of-day balance that can be maintained is USD 1458.61.

- *India Post Payments Bank digital savings account*

IPPB's mobile app allows customers to open this account on their own, instantly. A maximum end-of-day balance of USD 1458.61 is allowed and a maximum yearly cumulative deposit of USD 2917.73.

This zero balance account is subject to closure if "know your customer" (KYC) verification is not completed within 12 months of opening. KYC

helps banks verify their customers' identities. Aadhaar and PAN cards are mandatory for opening a digital savings account at IPPB.[2]

- *India Post Payments Bank basic savings bank deposit account*

This savings account has all the features and benefits offered by the regular savings account, except that it allows only four cash withdrawals in a month. It also does not require maintenance of any monthly average balance. It can be opened with a zero balance and offers an interest rate of 4% annum, paid quarterly.[3]

Financial inclusion products must be founded on the following principles:

E1: Effectiveness
How effective it is to use, how useful the product is and how it can make a difference in daily life.
E2: Efficiency
How efficient it is to use, whether it is easy to use, easy to handle, child compliant, suitable for senior citizens and differently abled people.
E3: Economical
How affordable it is for people at the bottom of the pyramid. The pricing of the product should be according to the cost of the technology. It should be accessible in small quantities to the majority of consumers.
E4: Environment
The product should be sustainable, eco-friendly and able to be recycled easily.

The earnings from these products can be reflected as a measure of the 3Ps (People, Planet and Profits):

Earnings = E1, E2, E3, E4 $\propto$ P1, P2, P3

[2] Aadhaar number is a unique 12-digit biometric-based random number issued by the UIDAI (Unique Identification Authority of India) to the residents of India after satisfying the verification process and the enrolment process is totally free of cost. A permanent account number (PAN) is a unique 10-digit alphanumeric number, issued by the Indian Income Tax Department, to any "person" who applies for it or to whom the department allots the number without an application and it is valid across India and is unaffected by a change of name, address within or across states in India or other factors.

[3] Sunil Dhawan (2017), 'This savings account doesn't require any minimum balance: Find out if it suits you', article published in Economic Times, India, dated 29 September 2017.

Digital finance products could be subjected to A/B testing. In web analytics, A/B testing (bucket tests or split-run testing) is a randomized experiment with two variants, A and B. It includes the application of statistical hypothesis testing or "two-sample hypothesis testing" as used in the field of statistics. A/B testing is a way to compare two versions of a single variable, typically by testing a subject's response to variant A against variant B, and determining which of the two variants is more effective.

The two versions that are compared are identical except for one variation that might affect a user's behavior. Version A might be the currently used version (control), while version B is modified in some respect (treatment). For instance, on an e-commerce website the purchase funnel is typically a good candidate for A/B testing, as even marginal improvements in drop-off rates can represent a significant gain in sales.

A/B testing has been marketed by some providers as a change in philosophy and business strategy in certain niches, though the approach is identical to a between-subjects design, which is commonly used in a variety of research traditions. A/B testing as a philosophy of web development brings the field into line with a broader movement toward evidence-based practice. The benefits of A/B testing are considered to be that it can be performed continuously on almost anything, especially since most marketing automation software now typically comes with the ability to run A/B tests on an ongoing basis.

Based on the results of A/B testing, products could be redesigned to meet the specific needs of customers within a period of 24 hours. Real-time tracking may not be possible due to the small amount of capital involved. However, A/B testing could produce good results.

Case 2: Direct benefit transfer—a game-changer for financial inclusion in India

History is replete with evidence of the role played by monetary and financial systems in people's lives. Observers describe "golden age" periods in which the quality of life was better, the spirit of innovation encouraged and prosperity widespread. To the extent that this was indeed the case, it was so because of wider participation in formal financial systems.

Today, the extent of such participation is an indicator of economic well-being made possible through opportunities in savings, payments, investments and insurance. Financial services also help small businesses grow, create employment and contribute to economies.

Yet a reported 2.5 billion adults worldwide are excluded from formal financial systems. In India, nearly half of the country's 1.25 billion people have no access to banking services, and close to 300 million people live below the official poverty line. To realize India's potential, its leaders need to ensure that this excluded segment can take part in formal banking and benefit from it.

The majority of people who live below the poverty line—as well as just above it—in India rely on government welfare payments such as those from the National Rural Employment Guarantee Act, old age pensions, scholarships, widows' pensions, discounted liquid petroleum gas (LPG) for cooking, and other subsidies. The government of India makes these payments, whose estimated value is a huge $60 billion per year, through the Direct Benefit Transfer (DBT) scheme, covering beneficiaries under 59 individual schemes.

The sheer extent of its reach and impact makes DBT an excellent platform to bring banking closer to the unbanked masses, but achieving that objective requires long-term vision, not short-term or ad hoc measures.

Over the last decade, technological developments have fast-tracked the way we do financial transactions. Financial inclusion efforts, too, have felt the impact of emerging technologies, with biometrics leading the way in extending banking reach. Technology-aided DBT delivery is the way forward.

However, the issues plaguing the existing transfer system need to be addressed, including broadband and mobile connectivity, lack of power, inaccurate data on beneficiaries, leaks and frauds due to corruption, and so on. This is where technology-driven solutions come into the picture.

JAM (Jan Dhan, Aadhaar, Mobile) is a combination of three platforms intended to drive financial inclusion efforts as envisaged by the government of India. The first is Jan Dhan Yojana, a project for opening bank accounts for all. The second is Aadhaar, a 12-digit unique identity number issued by the Unique Identification Authority of India. And the third is mobile technology, especially the increasing adoption of smartphones. The JAM trinity, as it is called, has the potential to completely digitize DBT delivery and promote financial inclusion.

Under this digital program, the first scheme to be rolled out using DBT is the LPG subsidy payment to nearly 150 million registered beneficiaries. Welfare payments for the National Rural Employment Guarantee Act and kerosene subsidy are expected to roll out next. Intended beneficiaries get their welfare payments directly credited to a bank account and can do financial transactions from multiple access points as is convenient—ATM,

bank branch or business correspondent point—with the unique identity number for authentication (see Fig. 10.4).

Source: National Payments Corporation of India. Infographics by FINO PayTech.

The latest government reports say that the implementation of DBT for LPG subsidy beneficiaries has led to savings of more than $3 billion during the last two years. This was possible because of two primary reasons: weeding out over 30 million fake beneficiaries using the unique identity number, and ensuring that the right beneficiary gets the right amount through the safety of a bank account.

If a single scheme can realize such savings, imagine the potential when $60 billion in welfare payments is seamlessly and directly credited to the bank accounts of all beneficiaries. Gradually, as hitherto unbanked and under-banked people start using bank accounts to make transactions, various other relevant financial products can be made available, allowing them to save, invest and insure their future. This innovative model can thus be a game-changer in promoting financial inclusion and fueling economic growth in India.

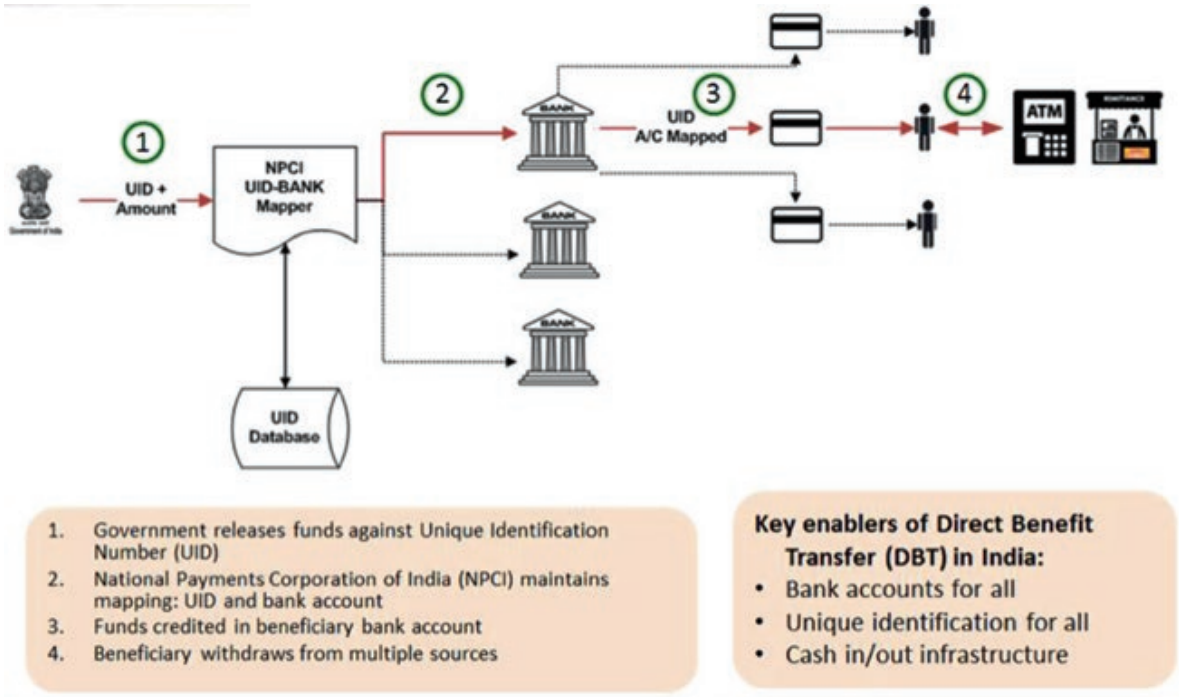


Fig. 10.4 Step direct benefit transfer delivery model. Source: National Payments Corporation of India. Infographics by FINO PayTech

Conclusions

In the age of the digital economy, digital finance with its unique features has been playing an important role in countries' inclusive growth process. It extends the opportunity for all, including the poor and disadvantaged, to participate in economic activities. This study has explored the effects of digital finance on financial inclusion by two methods. Via the first method, using data from 2004 to 2016 on 189 countries, the authors conducted trend analysis and found that all digital finance proxies have positive effects on economic growth. For the second method, considering India as a highly promising country for digital finance, we studied two cases and observed the positive effects of digital finance on financial inclusion. Thus, based on the findings, we argue that policy makers should adopt such policies to promote digital finance, which in turn will promote financial inclusion, resulting in the inclusive growth of the country.

Depending on the availability of data, future research could attempt to examine the relation between digital finance and financial stability with a special focus on periods of financial crisis, so that the economy will have the capacity to meet the challenge of such crises. Digital finance may not serve as an overall panacea for the world's poor, but it can definitely facilitate the better establishment of services to meet working capital requirements, thereby contributing to a corruption-free system leading to better dissemination of services. Predictive analytics combined with biometrics could be one of the futuristic areas of digital finance as the discipline matures and includes more stakeholders across the value chain, making it more inclusive and holistic. The future of digital finance and its implementation is here to stay and would provide ample opportunities for future researchers to expand its aim and scope.

References

Dhawan, S. (2017, September 29). This savings account doesn't require any minimum balance: Find out if it suits you. *Economic Times*, India.

Gabor, D., & Brooks, S. (2017). The digital revolution in financial inclusion: International development in the fintech era. *New Political Economy, 22*(4), 423–436.

Ghosh, S. (2016). Does mobile telephony spur growth? Evidence from Indian states. *Telecommunications Policy, 40*(10–11), 1020–1031.

Gomber, P., Koch, J. A., & Siering, M. (2017). Digital Finance and FinTech: Current research and future research directions. *Journal of Business Economics, 87*(5), 537–580.

GPFI (2010). G20 Principles for Innovative Financial Inclusion—Executive Brief. Retrieved from http://www.gpfi.org/publications/g20-principles-innovative-financial-inclusion-executive-brief

Hannig, A., & Jansen, S. (2010). Financial inclusion and financial stability: Current policy issues. *ADBI Working Paper No. 259*, Asian Development Bank Institute, Tokyo.

Klapper, L., El-Zoghbi, M., & Hess, J. (2016). Achieving the sustainable development goals. The role of financial inclusion. Retrieved May 23, 2016, from http://www.ccgap.org

Manyika, J., Lund, S., Singer, M., White, O., & Berry, C. (2016). *Digital finance for all: Powering inclusive growth in emerging economies*. Washington, DC: McKinsey Global Institute.

Ozili, P. K. (2018). Impact of digital finance on financial inclusion and stability. *Borsa Istanbul Review, 18*(4), 329–340.

Peake, C. (2012). New frontiers: Launching digital financial services in rural areas. In *The 2012 Brookings Blum Roundtable Policy Briefs*. Mercy Corps.

Sarma, M. (2015). Measuring financial inclusion. *Economics Bulletin, 35*(1), 604–611.

Siddik, M. N. A., Sun, G., & Kabiraj, S. (2015). Financial inclusion and its determinants: A study of Bangladesh. *Indian Journal of Finance, 9*(6), 7–29.

CHAPTER 11

Understanding Consumer Behavior in Technology-Mediated Spaces

İlayda İpek

Substantial improvements in technology have been influencing not only strategies and structures of firms but also consumers' social and economic lives. In terms of the former, in order to respond to the alterations in the technological environment, and ultimately, to stay ahead of the competition, firms have been making continuous attempts to fit their internal resources and competencies with the external circumstances. Concerning the latter, having an enormous impact on the manner of communicating, consuming, and progressing, technological advancements have been shaping consumers' daily lives. Specifically, digital technology has affected consumers' routines and habits, which in turn has led to changes in consumer behavior.

Currently, along with the benefits of the Internet, consumers have the ability to easily reach all information regarding firms, which increases the bargaining power of consumers. They have had a more tendency to shop online; however, the critical thing is that they desire to have unique experiences, which stresses the significance of effective customer relationship management (CRM) activities. Being aware of the strength of technology

İ. İpek (✉)
Faculty of Business, Dokuz Eylül University, İzmir, Turkey
e-mail: ilayda.gungor@deu.edu.tr

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_11

on consumer behavior, firms have been more inclined to enhance their offerings, with a particular emphasis on the design of positive consumer experiences. Additionally and remarkably, together with the advances in technology, the roles of firms and consumers have differed considerably. In this sense, firms have been working interdependently with their customers, who are highly involved in co-creating value. Correspondingly, digital marketing has gained huge importance in every aspect of business activities. Building on the aforementioned issues, by drawing on digital marketing, the main purpose of this chapter is to identify how consumer behavior has evolved through the digital transformation in recent decades.

Digital Marketing

Dating back to the early 1990s, the Internet has been commercially utilized since almost three decades (Tiffin & Rajasingham, 2003). Coupled with this digital evolution, boundaries across firms and cultures have been seriously decreased (Wymbs, 2000). In today's business world, rivalry among firms is not based on products or services, but instead on business models, which aim to generate customer value by the combination of digital technologies (Yoffie & Cusumano, 1999). Nevertheless, maintaining competitive advantage in a marketplace is much more burdensome than before by reason of high business risks related to the extremely dynamic nature of the digital environment (Hamel & Sampler, 1998). As it has been well emphasized by Bill Gates (1999), "the Internet changes everything".

Increase in the availability of the Internet; and thus, its extensive commercial usage, have led new players to stand out in today's modern economy. Those players have recorded huge success; such that, according to Forbes (2018a), Google, Facebook, and Amazon have been among the world's five most valuable brands in this year. E-commerce has been dramatically growing; for instance, current retail tendencies indicate that more than half of Americans (51%) have been engaged in online shopping in 2018, with the e-commerce growth rate of 23% year-over-year (BigCommerce, 2018). In addition, it has been expected that US consumers' spending on online shopping will total $632 billion by 2020 (Business Insider Intelligence, 2016).

Moreover, as another distinctive alteration in the technological environment, Web 2.0 has begun to replace Web 1.0. Whereas the contents of Web 1.0 tools are created by providers, Web 2.0 technologies incorporate

interactive and user-generated content (Munar & Jacobsen, 2014). In conjunction with the widespread acceptance of Web 2.0 technologies, a substantial switch from firms to consumers has been reflected (Berthon, Pitt, Plangger, & Shapiro, 2012). In spite of its technological nature, Web 2.0 has also yielded to sociological revolutions, which have exerted an enormous influence on business operations (Berthon et al., 2012). Relying on these facts, in recent years firms have been more inclined to enhance "digital relationships" with their customers (Phillips, 2015).

In order to keep pace with the digital transformation of consumer behavior, firms have been oriented to reshape and restructure their functions, with a special stress on marketing operations (Edelman & Heller, 2015). Accordingly, over the past decades, marketing scholars and practitioners have registered a significant change in the nature of marketing, specifically due to the emergence of new marketing communication channels derived from technological innovations (Lamberton & Stephen, 2016). Currently, digital technology has critically altered the way firms formulate their targeting strategies and interact with their customers. In this respect, "digital marketing" arises as a crucial concept.

The digital marketing term is defined as "an adaptive, technology-enabled process by which firms collaborate with customers and partners to jointly create, communicate, deliver, and sustain value for all stakeholders" (Kannan & Li, 2017, p. 23). Digital marketing represents a novel perspective, not only traditional marketing reinforced by digital components; hence, it includes idiosyncratic characteristics that need to be in-depth evaluated (Taiminen & Karjaluoto, 2015). Instead of concentrating on technology, an effective digital marketing strategy necessitates to comprehend how consumers adopt technology, and correspondingly, to look for channel alternatives in an attempt to strengthen the engagement between consumers and brand (Ryan & Jones, 2009). Without a proper analysis, digital marketing communications may not make sense for target customers, which in turn may result in big losses. Dealing with change and innovation seems to be problematic for all marketing practitioners; however, the situation is notably difficult for digital marketers (Chaffey, 2010).

Being aware of the role of digital marketing in consumers' decision making process, in today's business world, it is imperative for firms to apply digital marketing solutions in an effort to reach their target customers. This is especially valid for both large and small- and medium-sized enterprises to stay competitive in market (Taiminen & Karjaluoto, 2015). Newly, digital advertising expenditures provide evidence on this notion;

digital ad spending worldwide accounted for $283.35 billion in 2018, with a growth rate of 21.4% (eMarketer, 2019). In addition, it has been also anticipated that worldwide spending on digital advertising will amount to $517.51 billion by 2023 (eMarketer, 2019). Critically, the amount of digital touchpoints is rising by more than 20% in every year, because more number of consumers prefers to use digital technologies and "younger, digitally oriented consumers enter the ranks of buyers" (Bughin, 2015). Contacting with consumers by means of digital media is therefore recognized as one of the most promising approaches in the area of strategic marketing for the coming decades (Smith, 2011).

Accompanied by digitization, traditional marketing communication tools started to be supported by new digital channels. These digital tools, providing new and exciting platforms, enable firms to reach target markets in various and creative manners (Ryan & Jones, 2009). Importantly, besides its impact on targeting strategies, digital marketing also allows to leverage existing brand equity (Chaffey, 2010). Augmented reality practices, social media, and mobile apps constitute some late communication channels in digital marketing, which have been broadly preferred by today's success-oriented firms. As an illustration, Uber symbolizes an achievement story, mainly by virtue of the right mixture of targeting and digital media. Operations of the firm are managed with a mobile app and the firm has largely employed online advertising in social media to cultivate its business activities. Presently, Uber announced $48 billion valuation (Donnelly, 2018), which signifies an outstanding accomplishment obtained in just nine years. Another similar example in the sharing economy focusing on digital marketing is Airbnb founded in 2008, with the reported value of $31 billion as of March 2017 (Forbes, 2018b). In a nutshell, it has been strongly suggested that "technology has facilitated novel market behaviors, interactions, and experiences" (Lamberton & Stephen, 2016, p. 146).

It is a foregone conclusion that digital marketing strategies are of vital importance in today's highly dynamic environment; nevertheless, it has been also firmly advocated that firms should be careful about the decisions regarding digital marketing channels. Specifically, firms cannot perform in all digital communication channels, because "being active" is the first and foremost factor of success (Kaplan & Haenlein, 2010). The appropriate digital medium needs to be determined on the basis of customers to be targeted and message to be delivered (Kaplan & Haenlein, 2010). Moreover, the number and engagement degrees of customers in a digital

platform remarkably influence performance (Trusov, Bodapati, & Bucklin, 2010). As a result, pros and cons of each communication medium require to be deeply scrutinized before resources are assigned to digital marketing activities.

E-Customers

It has been widely acknowledged that increase in consumers' involvement in digital devices has given rise to changes in consumer behavior. In particular, information technology innovations have affected to a greater extent how consumers respond to different market settings (Lamberton & Stephen, 2016). The Internet usage rates all over the world have been growing in each year; the statistical report offered by the World Bank (2018) demonstrates that 45.79% of the world's population used the Internet in 2016, while this number was 43.08% in the previous year. Even more stunning is the Internet adoption statistics in some developed countries; for example, in the UK, Japan, and United Arab Emirates more than 90% of the country's whole population utilized the Internet in 2016 (World Bank, 2018).

Consumers' rising inclination toward using the Internet and digital channels can be well explained by the principles of the Technology Acceptance Model (TAM) (Davis, 1989) and Diffusion of Innovations (Rogers, 1962). In line with the TAM, since digital technology encompasses both perceived usefulness and perceived ease of use (Davis, 1989), consumers have been more willing to accept the transformations in the technological environment. As for the Diffusion of Innovations, on the basis of the fit between the digital technology and the characteristics, which affect the acceleration of innovation (i.e., relative advantage, compatibility, complexity, trialability, and observability) (Rogers, 1962), consumers have become more interested in engaging in the Internet and digital tools (Rodriguez, Peterson, & Krishnan, 2018; Takacs & Freiden, 1998). Hence, progresses in the Internet and Web-based platforms have cultivated online consumer behavior, which has shaped the nature of consumers' daily activities (Tiago & Veríssimo, 2014).

The essence of the interaction between firms and customers is not like before, since "The Internet is not just another marketing channel; it's not just another advertising medium; it's not just a way to speed up transactions; the Internet is the foundation for a new industrial order" (Hamel & Sampler, 1998). Web 2.0 technologies have dramatically resulted in a shift

in value creation and power from firms to customers (Berthon et al., 2012). Currently, customers are more empowered than before because of the advantages associated with the employment of digital devices, covering, information access, product variety, competitive pricing, and convenience (Pires, Stanton, & Rita, 2006; Tiago & Veríssimo, 2014). Without much effort, they are able to spread word-of-mouth information about products/services via social networks even with individuals they do not know (Kannan & Li, 2017). Further, as another concern for digital marketers, customers are becoming increasingly sophisticated and their expectations are incessantly expanding (Udo, Bagchi, & Kirs, 2010).

Customers are actively taking part in different digital communication channels such as social media, web forums, blogs, and mobile apps. Whereas previously for product/service purchases, customers were more relying on worth-of-mouth communication with membership reference groups such as family and friends, today they prefer more to gather information about products/services through online reviews, to compare and contrast different offerings' aspects and prices on websites, and to talk over alternatives through social-networking platforms (Edelman, 2010). In conjunction with customers' high authority in digital world, brand messages formed by firms have lost their effect and the possibility of conversion has declined on average (Bughin, 2015). Realizing their bargaining power on firms, customers want firms to build close bonds with them, rather than to command (Edelman, 2010).

Especially, social media acts a critical role for digital customers; in social media customers have an interaction with brands and they exchange information and experiences with others (Lamberton & Stephen, 2016), which have been considered to have a significant influence on buying behavior (Berthon et al., 2012). It has been revealed that in globe 135 minutes were daily spent in social media in 2017, with an increase from 126 minutes in the prior year (Ward, 2018). And interestingly, the statistics referring to the time devoted for mobile devices have indicated that in the US individuals consume more than 90% of their Internet time on smartphone apps (eMarketer, 2017).

In some cases, customers' high involvement in digital mediums may be seen as problematic in terms of firms. As it has been already highlighted, today customers are identified as "unfettered" (Nunes & Cespedes, 2003), and correspondingly, they are free to spread positive and negative words about a firm's offerings and strategies with or without getting permission of the firm (Kietzmann, Hermkens, McCarthy, & Silvestre, 2011). This makes firms defenseless as their offerings' quality, availability, and prices

are completely obvious and any weak points are swiftly disseminated all over the world (Nunes & Cespedes, 2003). Such kinds of online customer-to-customer information exchanges are attempted to be managed by firms with word-of-mouth marketing methods such as social media marketing, viral marketing, and guerilla marketing (Kozinets, de Valck, Wojnicki, & Wilner, 2010), because customers are highly dependent on other customers' opinions and feelings while making buying decisions.

Online Customer Experiences

Developments in technology; logistics, payment systems, and trust in particular—together with the escalation in the Internet usage—have formed a $1.9 trillion global online shopping market, in which "millions of consumers no longer 'go' shopping, but literally 'are' shopping - at every moment and everywhere" (KPMG, 2017, p. 2). Primarily due to the benefits provided by online retailing to consumers such as freedom of shopping in any time and place, online sales have smoothly reached double-digit growth numbers in some countries (Nielsen, 2017). Notably, in the world's biggest online market, China, the growth rate of 32.2% year-on-year in online retailing was recorded in 2017, while this number decreased by 2.1% in the first two quarters of 2018 (National Bureau of Statistics of China, 2018a, 2018b). The nature of online shopping is assumed to alter at an even faster speed coupled with the expansion of voice search and omni-platform/omni-channel alternatives (Popomaronis, 2017).

Apart from these online shopping trends, it is indisputable that consumers are not rational decision makers anymore. During the stages of the decision making process, they are more relying on their feelings and they want to have unique experiences. Accordingly, the present situation is more challenging for firms, since they have to devote more time and effort to improve their offerings through well-designed and positive customer experiences. This fact is not only true for brick-and-mortar businesses; by virtue of highly competitive environment, online operations need to be also strengthened by favorable experiences in order to leverage conversion rates. However, this is much more painful, as providing and managing "a balanced experience" on a digital platform is severely complicated (Chaffey, 2010).

Consumers collect information about products/services either by direct experience, such as trial, or by indirect experience, such as exposure to advertising (Li, Daugherty, & Biocca, 2001). Consumer experiences in digital environment are considerably different; such that, this concept has been viewed as indirect experiences; but it also includes social

and interpersonal cues (Liao & Keng, 2014). Moreover, in Web-based platforms consumers rest solely on sight and sound, whereas in traditional marketplace experiences encompass all senses (Straker, Wrigley, & Rosemann, 2015). In this regard, online consumer experiences are described as psychological and emotional conditions faced by consumers, meanwhile having interaction with products online (Li et al., 2001). Widespread preference of online shopping has resulted in the evolution of the e-customers' buying behavior, allied with the attainment of e-purchasing experiences (Hernandez, Jimenez, & Martín, 2009).

Increase in the strength of online consumer behavior has caused firms to expand their digital existence (Nielsen, 2017); and hence, to look for the ways to create a unified experience across various channels. On the other hand, it has been shown that one of the most important obstacles of online shopping is the lack of delightful experiences and social interplay (Barlow, Siddiqui, & Mannion, 2004). Consumers remark that online transactions are not specific to a certain customer and they do not feel comfortable while purchasing unusual and intricate products; they therefore expect firms to offer customer assistance not only in offline but also in online environments (Holzwarth, Janiszewski, & Neumann, 2006), that places a high emphasis on the significance of unified customer experience. In this respect, customer experiences have been recognized as the pivotal drivers of online shopping behavior (Childers, Carr, Peck, & Carson, 2001).

Further, as a consequence of the diffusion of digital marketing tools, currently, customers have become more engaged with brands; for example, they receive regular e-mails about new season products, they utilize mobile apps to take the advantage of discount prices, or they visit the brand's social media pages to learn about the upcoming events (Edelman, 2010). Nevertheless, considering customers' involvement in different digital channels at once, firms should be more inclined toward generating multi-channel experiences (Straker et al., 2015). This approach has been thoroughly underlined by the term of "perfect customer experience", which is firmly intertwined with the firm's capability "to gather and deploy customer information from all channels and to integrate it with other relevant information" (Payne & Frow, 2005, p. 173). In this sense, digital channels help firms consolidate customer experiences, which in turn produce high online traffic and superior customer loyalty at the end (Edelman, 2010).

Currently, many firms have leaned toward the integration of online and offline customer experiences. L'Oreal is one of those firms which successfully creates unique experiences for its customers through the combination

of online and offline aspects. As an illustration, L'Oreal's Makeup Genius app offers opportunity for customers to virtually try various kinds of makeup products through a webcam (Edelman & Heller, 2015). As another example, last year IKEA introduced IKEA Place app—an augmented reality application, which aims to form memorable customer experiences by enabling people to virtually design their homes, offices, and so on with different furniture (IKEA, 2017).

Furthermore, in recent decades retailing industry has also registered substantial alterations, mainly due to the improvements in online mediums and continuous digitalization (Verhoef, Kannan, & Inman, 2015). In an effort to respond to changes in retailing, firms have begun to simultaneously operate both in traditional and online channels; however, at this point the critical thing is that those retailing activities require to be carried out in an integrated fashion. Assuming that retailing operations are dependent on autonomous systems, fragmented supply chains emerge, which fail to create a coherent and trustworthy customer experience (Saghiri, Wilding, Mena, & Bourlakis, 2017). Concordantly, in the latest years, a shift from multi-channel retailing to omni-channel retailing has been noticed (Verhoef et al., 2015).

Whereas multi-channel retailing fundamentally concentrates on retail channels, in omni-channel retailing the stress is placed on the integration between channels and brands (Picot-Coupey, Huré, & Piveteau, 2016). Symbolizing a challenge for traditional retailers (Rigby, 2011), omni-channel retailing is defined as "the synergetic management of the numerous available channels and customer touchpoints, in such a way that the customer experience across channels and the performance over channels is optimized" (Verhoef et al., 2015, p. 176). It puts a particular emphasis on the interplay of "disparate channels into a single seamless omni-channel experience" (Rigby, 2011, p. 67). The concept has gained a considerable importance in retailing together with increase in customers' demands for unified brand experiences. Concisely, the effective management of the multi-channel customer experiences is particularly vital for outstanding business success and performance in today's highly digital environment.

E-Customer Relationship Management

Relationship marketing, focusing on creating, expanding, and sustaining sufficient relational exchanges, has been seen as a big transformation for both marketing scholars and practitioners (Morgan & Hunt, 1994). It has

been strongly suggested that establishing long-term relationships with internal and external partners generates more positive returns compared to discrete transactions (Morgan & Hunt, 1994); that view has caused a fundamental shift from a transaction-based perspective to a relationship-centered orientation. In this regard, within the context of CRM, in what manner customers should be dealt to establish effective bonds with them has become a crucial concern in marketing theory and practice (Reinartz, Krafft, & Hoyer, 2004).

Drawing on the relationship marketing theory, CRM rests primarily on the notion that in order to be successful and survive in the market, instead of looking for ways to sell products/services, firms should consistently direct themselves toward building high customer value, which in turn creates value for the firm (Boulding, Staelin, Ehret, & Johnston, 2005). The main tenets of CRM defend that as long-term customers are more profitable than short-term ones, firms should maintain their relationships with customers to improve customer retention and loyalty (Zeithaml, Berry, & Parasuraman, 1996). Correspondingly, firms have increasingly started to appreciate that the economic worth of customers to the firm is not the same; and thus, they need to tailor their products/services and marketing communications on the basis of various customers (Reinartz et al., 2004).

Even though CRM is frequently adopted to refer to technology-oriented customer solutions (e.g., sales force automation), the breadth of the concept is wider; such that, CRM covers a strategic vision with a particular stress on the interrelationship of organizational culture, information technology, and customer service (Payne & Frow, 2005). Nevertheless, whereas CRM comprises diverse types of applications, a technology-based view is widely accepted (Keramati, Mehrabi, & Mojir, 2010). The fast progress in information technology has allowed firms to adopt new technology-based tools, entitled CRM technology, which aims to support the process of CRM (Jayachandran, Sharma, Kaufman, & Raman, 2005). Currently, many firms make huge investments in CRM technology in an attempt to effectively manage their CRM activities (Powell, Noble, Noble, & Han, 2018).

In the increasingly digitalized world, CRM technology is of great importance; CRM was identified as the biggest software market in 2017, with the worldwide revenue totaling $39.5 billion (Gartner, 2018). In addition, it has been estimated that CRM will grow by 16% in 2018 and will therefore represent the fastest expanding software market in the same year (Gartner, 2018). Adoption rates also signify the strength of CRM in

current topics of marketing; with approximately 47% in 2017, on average (Lazar, 2017). Firms prefer to employ CRM technology for different purposes such as more accurate customer profitability evaluation, integrated customer data management, and superior customer lifetime value (Powell et al., 2018). However, albeit its popularity, the consequences of CRM technology are questionable in terms of either customers or firms (Jayachandran et al., 2005).

By virtue of the substantial effect of technological advancements on business processes, CRM operations have been also influenced by Internet-based technologies in a considerable way. In recent decades, an increasing integration between CRM and Internet technology has been noticed, that formed a novel approach for CRM strategy. Being called as electronic CRM (e-CRM), traditional CRM applications have been replaced with a more interactive nature, depending fundamentally on regular and effective customer communication (Kımıloğlu & Zaralı, 2009). Together with exploiting the benefits of Internet technology, e-CRM enables to expedite the handling of customer relationships (Harrigan, Ramsey, & Ibbotson, 2008).

A comprehensive definition of e-CRM involves "the marketing activities, tools and techniques, delivered over the Internet (using technologies such as Web sites and e-mail, data-capture, warehousing and mining) with a specific aim to locate, build and improve long-term customer relationships to enhance their individual potential" (Lee-Kelley, Gilbert, & Mannicom, 2003, p. 241). In this sense, e-CRM offers important advantages in respect of scrutinizing customer data, since at the right time it signals what customers really expect (Harrigan et al., 2008). Thanks to this attribute, it prevents resource losses in production operations systems (Harrigan et al., 2008). Importantly, by means of e-CRM, firms are more able to personalize their bonds with customers, notably owing to higher number of customer touchpoints and greater opportunity to customize marketing communications (Day & Hubbard, 2003).

Referring to the reality of digital consumer behavior and the significance of e-CRM with respect to both efficiency and effectiveness, today firms have been more apt to apply online CRM solutions to attain competitive advantage. Besides, in parallel with the late trends in consumer behavior and the growing impact of Web 2.0 technologies, within the context of e-CRM, new concepts also emerged. Among these, social CRM (or CRM 2.0) serves as a popular philosophy that combines social media and dynamic capabilities of traditional CRM (Greenberg, 2010). This new concept has been recognized as a more cooperative and network-driven

perspective to deal with customer relationships (Trainor, Andzulis, Rapp, & Agnihotri, 2014).

Customers have become highly involved in social media channels not only to keep in touch with their friends or colleagues but also to interact with firms (Trainor, 2012). In this regard, the brief purpose of social CRM is to respond to the expectations of social customers (Greenberg, 2010). Via social CRM applications firms endeavor to cultivate their bonds with customers in social networks both with regard to quality and quantity; and thus, to reinforce brand image and overall brand loyalty (Wongsansukcharoen, Trimetsoontorn, & Fongsuwan, 2015). Seen as a new breath in CRM, social CRM is of great benefit to firms. such that it has been found that customer retention rates increase by 26%, thanks to social CRM applications (Lazar, 2017). Considering these facts, traditional CRM service providers (e.g., Salesforce.com and SAP) have begun to merge data from social media channels such as Facebook and Twitter to actively monitor customers and quickly respond to customer requirements (Owyang, 2010).

Customer Co-Creation

Traditional value creation approach is based on the thought that firms establish value and transfer it to customers (Prahalad & Ramaswamy, 2004a). Specifically, with limited or without interaction with customers, firms independently decide on which products to be manufactured, which marketing messages to be delivered, which distribution channels to be used, and so on so forth (Prahalad & Ramaswamy, 2004b). On the other hand, fundamentally because of the upheavals in customers' roles in recent decades, the main tenets of the traditional value creation view started to be doubted (Prahalad & Ramaswamy, 2004a). It has been suggested that this passive perspective regarding customers should be changed with an active approach, since the real value of a product/service can be best reflected in the eyes of customers (Witell, Kristensson, Gustafsson, & Löfgren, 2011).

Coupled with the big transformation in the essence of the relationship between customers and firms, the process of value building has considerably altered; currently, customers have tended toward taking an active part in each step of firms' value creation processes (Prahalad & Ramaswamy, 2004b). Being described as "co-creation" of value, customers have begun to have an impact not only on product development stages but also on service designs (Heidenreich, Wittkowski, Handrich, & Falk, 2015). Based on this, it has been contended that marketing has recorded a significant

change from a goods-dominant logic, focusing mainly on "tangible output and discrete transactions", to a service-dominant logic, which depends strictly on "intangibility, exchange processes, and relationships" (Vargo & Lusch, 2004, p. 2).

Rather than following the traditional definitions of the term, the service-dominant approach identifies services as "the application of specialized competences (knowledge and skills) through deeds, processes, and performances for the benefit of another entity or the entity itself" (Vargo & Lusch, 2004, p. 2) and it places special emphasis on the collaboration between the firm and its stakeholders, which are highly involved in the co-creation of value (Lusch, Vargo, & O'Brien, 2007). This perspective claims that customers represent stakeholders, who consistently make contributions to value generation, whereas firms refer to enterprises with an opportunity to only propose value (Vargo & Lusch, 2004).

Growing engagement in the usage of the Internet and social media channels has facilitated the implementation of co-creation activities. Accordingly, in conjunction with the benefits of digital technology, firms have been increasingly formulating successful co-creation strategies. For instance, Dell's Ideastorm exemplifies one of those impressive applications in value co-creation (Saarijärvi, Kannan, & Kuusela, 2013). In this platform, customers of Dell express their opinions and suggestions about how to improve the offerings of the firm. Owing to this, Dell has an access to customers' actual ideas in relation to Dell's products, and at the same time customers see that their contributions are of great prominence for Dell, that is likely to strengthen customer participation and citizenship behavior (Yi & Gong, 2013).

In this context, Threadless.com also sets a good example (Saarijärvi et al., 2013). The firm calls its customers to design different range of products, covering t-shirts, accessories, and home decor, and so on, and to load their drafts on its website. All customers can give scores for the sketches and total customer scores determine which products will be sold under the name of Threadless.com. By means of this co-creation of value, customers are able to purchase their favorite products and the firm can anticipate future customer demands (Saarijärvi et al., 2013). Furthermore, BMW also cooperates with its customers to put its cars into final form (BMW Group, 2010). In particular, the firm introduced the Co-Creation Lab, which serves as a virtual gathering place where customers are encouraged to comment on automotive world in the future. BMW utilizes the sharing of the customers in order to plan the upcoming projects.

Building on the foregoing, in an effort to exploit the advantages of value co-creation, firms have been trying their best to employ customer resources in an efficacious manner. By means of co-creation activities, firms collect information about future customer requirements and expectations (Gustafsson, Kristensson, & Witell, 2012), which cultivate open innovation (Martini, Massa, & Testa, 2014). By virtue of this close interaction with customers in value creation, firms are able to achieve a competitive positional advantage both in terms of efficiency and effectiveness (Grissemann & Stokburger-Sauer, 2012). The benefits of value co-creation are not confined to firms; but also include customers (Dong, Evans, & Zou, 2008). For example, co-creation of value allows customers to enjoy getting quick response to their demands and having products/services with lower prices (Dong et al., 2008).

Conclusion

In the light of the aforementioned issues, considerable developments in technology have been exerting an enormous impact on both firms and consumers. From the firms' standpoint, necessary adjustments in organizational structures and strategies need to be made in order to keep pace with the latest technological progress; however, more critically, firms should be simultaneously involved in scrutinizing deeply and understanding ever-changing consumer behavior. This is notably because in today's highly digitalized world the way consumers use products/services, communicate with their environment, and interact with firms has been experiencing a huge change. In this sense, tremendous alterations in consumer behavior have been putting a great pressure on firms to be more leaned toward utilizing novel marketing programs and solutions in general, and digital marketing activities in particular.

Importantly, together with the increase in the extension rate of information and communication technologies, the authority of market knowledge has started to be passed from suppliers to consumers, which has yielded to the transformations of consumer empowerment, and ultimately, marketing implications (Pires et al., 2006; Tiago & Veríssimo, 2014). Currently, e-consumers are more able to easily compare and contrast features, prices, and varieties of products/services offered by various firms and even more significant to contact with firms and other consumers without difficulty. Hence, consumers have been considered being more sophisticated and holding more bargaining power than before (Berthon

et al., 2012; Udo et al., 2010), which puts special emphasis on the fact that digital engagement between firms and consumers has become much more prominent.

Specifically, relying on the notion that in conjunction with rational drivers, consumers take emotional factors into consideration during their decision making processes (Schmitt, 1999), a vital requirement emerges to provide positive and unique customer experiences not only through offline but also through online platforms in the contemporary marketplace. However, at this point, a critical concern that warrants attention by marketing practitioners pertains to the design of unified customer experiences which balance and harmonize experiences across different marketing channels. In this respect, omni-channel retailing arises as a valuable concept, of which applications are promising in terms of the management of customer experiences, since "the distinctions between physical and online will vanish, turning the world into a showroom without walls" (Brynjolfsson, Hu, & Rahman, 2013, p. 2).

Additionally, digital consumer behavior has also reflected in the revolution of value creation perspective; such that, customers have increasingly stood out as co-creators of value. In recent decades, the collaboration between firms and customers has been much greater and customers have been taking an active part more in establishing value, echoing successful real life examples. As customers represent the best sources of information regarding a value of a product/service (Witell et al., 2011), firms should be more inclined to work together with customers in an attempt to enhance their offerings and to gain a competitive advantage in the marketplace. Drawing on all of these important evolutions in consumer behavior, e-CRM applications have become even more paramount, which underline the significance of creating long-term and effective bonds with digital customers. All in all, this chapter spotlights how consumer behavior has remarkably changed on the basis of the digital metamorphosis, with the particular stress placed on the main tenets of digital marketing and e-customers.

References

Barlow, A. K., Siddiqui, N. Q., & Mannion, M. (2004). Developments in information and communication technologies for retail marketing channels. *International Journal of Retail and Distribution Management, 32*(3), 157–163.

Berthon, P. R., Pitt, L. F., Plangger, K., & Shapiro, D. (2012). Marketing meets Web 2.0, social media, and creative consumers: Implications for international marketing strategy. *Business Horizons, 55*(3), 261–271.

BigCommerce. (2018). The 19 ecommerce trends + 147 online shopping stats fueling sales growth in 2018. Retrieved September 6, 2018, from https://www.bigcommerce.com/blog/ecommerce-trends/

BMW Group. (2010). BMW Group Co-Creation Lab. Retrieved September 27, 2018, from https://www.press.bmwgroup.com/global/article/detail/T0082655EN/bmw-group-co-creation-lab?language=en

Boulding, W., Staelin, R., Ehret, M., & Johnston, W. J. (2005). A customer relationship management roadmap: What is known, potential pitfalls, and where to go. *Journal of Marketing, 69*(4), 155–166.

Brynjolfsson, E., Hu, Y.J., & Rahman, M.S. (2013). Competing in the age of omnichannel retailing. *MIT Sloan Management Review*, Summer, 1–7.

Bughin, J. (2015). Brand success in an era of Digital Darwinism. *McKinsey Quarterly*. Retrieved September 9, 2018, from https://www.mckinsey.com/industries/high-tech/our-insights/brand-success-in-an-era-of-digital-darwinism

Business Insider Intelligence. (2016). Retail Forecast: U.S. consumers will spend $632 billion online by 2020. Retrieved September 6, 2018, from https://www.businessinsider.com/ecommerce-online-retail-monetization-strategy-planning-2016-11

Chaffey, D. (2010). Applying organisational capability models to assess the maturity of digital-marketing governance. *Journal of Marketing Management, 26*(3–4), 187–196.

Childers, T. L., Carr, C. L., Peck, J., & Carson, S. (2001). Hedonic and utilitarian motivations for online retail shopping behavior. *Journal of Retailing, 77*(4), 511–535.

Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340.

Day, G. S., & Hubbard, K. J. (2003). Customer relationships go digital. *Business Strategy Review, 14*(1), 17–26.

Dong, B., Evans, K. R., & Zou, S. (2008). The effects of customer participation in co-created service recovery. *Journal of the Academy of Marketing Science, 36*(1), 123–137.

Donnelly, G. (2018). Travis Kalanick net worth tops $4 billion with sale of Uber stock. *Fortune*. Retrieved September 8, 2018, from http://fortune.com/2018/01/19/travis-kalanick-net-worth/

Edelman, D. (2010). Four ways to get more value from digital marketing. *McKinsey Quarterly*. Retrieved September 9, 2018, from https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/four-ways-to-get-more-value-from-digital-marketing

Edelman, D., & Heller, J. (2015). How digital marketing operations can transform business. *McKinsey Digital*. Retrieved September 6, 2018, from https://www.mckinsey.com/business-functions/marketing-and-sales/our-insights/how-digital-marketing-operations-can-transform-business

eMarketer. (2017). Mobile time spent in 2018: Will smartphones remain ascendant? Retrieved September 7, 2018, from https://www.emarketer.com/content/mobile-time-spent-2018

eMarketer. (2019). Digital Ad spending 2019. Retrieved May 8, 2019, from https://www.emarketer.com/content/global-digital-ad-spending-2019

Forbes. (2018a). The world's most valuable brands. Retrieved September 6, 2018, from https://www.forbes.com/powerful-brands/list/#tab:rank

Forbes. (2018b). As a rare profitable unicorn, Airbnb appears to be worth at least $38 billion. Retrieved September 8, 2018, from https://www.forbes.com/sites/greatspeculations/2018/05/11/as-a-rare-profitable-unicorn-airbnb-appears-to-be-worth-at-least-38-billion/#2c045ff22741

Gartner. (2018). Gartner says CRM became the largest software market in 2017 and will be the fastest growing software market in 2018. Retrieved September 17, 2018, from https://www.gartner.com/newsroom/id/3871105

Gates, B. (1999). Business @ the speed of thought. *Business Strategy Review, 10*(2), 11–18.

Greenberg, P. (2010). The impact of CRM 2.0 on customer insight. *Journal of Business and Industrial Marketing, 25*(6), 410–419.

Grissemann, U. S., & Stokburger-Sauer, N. E. (2012). Customer co-creation of travel services: The role of company support and customer satisfaction with the co-creation performance. *Tourism Management, 33*(6), 1483–1492.

Gustafsson, A., Kristensson, P., & Witell, L. (2012). Customer co-creation in service innovation: A matter of communication? *Journal of Service Management, 23*(3), 311–327.

Hamel, G., & Sampler, J. (1998). The e-corporation more than just web-based, it's building a new industrial order. *Fortune*. Retrieved September 9, 2018, from https://money.cnn.com/magazines/fortune/fortune_archive/1998/12/07/252120/index.htm

Harrigan, P., Ramsey, E., & Ibbotson, P. (2008). E-CRM in SMEs: An exploratory study in Northern Ireland. *Marketing Intelligence and Planning, 26*(4), 385–404.

Heidenreich, S., Wittkowski, K., Handrich, M., & Falk, T. (2015). The dark side of customer co-creation: Exploring the consequences of failed co-created services. *Journal of the Academy of Marketing Science, 43*(3), 279–296.

Hernandez, B., Jimenez, J., & Martín, M. J. (2009). Adoption vs acceptance of e-commerce: Two different decisions. *European Journal of Marketing, 43*(9/10), 1232–1245.

Holzwarth, M., Janiszewski, C., & Neumann, M. M. (2006). The influence of avatars on online consumer shopping behavior. *Journal of Marketing, 70*(4), 19–36.

IKEA. (2017). IKEA launches IKEA place, a new app that allows people to virtually place furniture in their home. Retrieved September 24, 2018, from https://www.ikea.com/us/en/about_ikea/newsitem/091217_IKEA_Launches_IKEA_Place

Jayachandran, S., Sharma, S., Kaufman, P., & Raman, P. (2005). The role of relational information processes and technology use in customer relationship management. *Journal of Marketing, 69*(4), 177–192.
Kannan, P. K., & Li, H. A. (2017). Digital marketing: A framework, review and research agenda. *International Journal of Research in Marketing, 34*(1), 22–45.
Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons, 53*(1), 59–68.
Keramati, A., Mehrabi, H., & Mojir, N. (2010). A process-oriented perspective on customer relationship management and organizational performance: An empirical investigation. *Industrial Marketing Management, 39*(7), 1170–1185.
Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons, 54*(3), 241–251.
Kımıloğlu, H., & Zaralı, H. (2009). What signifies success in e-CRM? *Marketing Intelligence and Planning, 27*(2), 246–267.
Kozinets, R. V., de Valck, K., Wojnicki, A. C., & Wilner, S. J. (2010). Networked narratives: Understanding word-of-mouth marketing in online communities. *Journal of Marketing, 74*(2), 71–89.
KPMG. (2017). The truth about online consumers: 2017 global online consumer report. Retrieved September 11, 2018, from https://assets.kpmg.com/content/dam/kpmg/xx/pdf/2017/01/the-truth-about-online-consumers.pdf
Lamberton, C., & Stephen, A. T. (2016). A thematic exploration of digital, social media, and mobile marketing: Research evolution from 2000 to 2015 and an agenda for future inquiry. *Journal of Marketing, 80*(6), 146–172.
Lazar, M. (2017). Ecommerce CRM software: 2017 CRM statistics show why it's a powerful marketing weapon. *IBM Community*. Retrieved September 17, 2018, from https://www.ibm.com/developerworks/community/blogs/d27b1c65-986e-4a4f-a491-5e8eb23980be/entry/2017_CRM_Statistics_Show_Why_it_s_a_Powerful_Marketing_Weapon?lang=en
Lee-Kelley, L., Gilbert, D., & Mannicom, R. (2003). How e-CRM can enhance customer loyalty. *Marketing Intelligence and Planning, 21*(4), 239–248.
Li, H., Daugherty, T., & Biocca, F. (2001). Characteristics of virtual experience in electronic commerce: A protocol analysis. *Journal of Interactive Marketing, 15*(3), 13–30.
Liao, T. H., & Keng, C. J. (2014). Online purchase delay: The roles of online consumer experiences. *Journal of Electronic Commerce Research, 15*(2), 133–150.
Lusch, R. F., Vargo, S. L., & O'Brien, M. (2007). Competing through service: Insights from service-dominant logic. *Journal of Retailing, 83*(1), 5–18.
Martini, A., Massa, S., & Testa, S. (2014). Customer co-creation projects and social media: The case of Barilla of Italy. *Business Horizons, 57*(3), 425–434.
Morgan, R. M., & Hunt, S. D. (1994). The commitment-trust theory of relationship marketing. *Journal of Marketing, 58*(3), 20–38.

Munar, A. M., & Jacobsen, J. K. S. (2014). Motivations for sharing tourism experiences through social media. *Tourism Management, 43*, 46–54.

National Bureau of Statistics of China. (2018a). Total retail sales of consumer goods in December 2017. Retrieved September 11, 2018, from http://www.stats.gov.cn/english/pressrelease/201801/t20180126_1577681.html

National Bureau of Statistics of China. (2018b). Total retail sales of consumer goods in June 2018. Retrieved September 11, 2018, from http://www.stats.gov.cn/english/pressrelease/201807/t20180719_1610962.html

Nielsen. (2017). What's in-store for online grocery shopping: Omnichannel strategies to reach crossover shoppers. Retrieved September 11, 2018, from https://www.nielsen.com/content/dam/nielsenglobal/de/docs/Nielsen%20Global%20Connected%20Commerce%20Report%20January%202017.pdf

Nunes, P., & Cespedes, F. V. (2003, November). The customer has escaped. *Harvard Business Review, 81*, 96–105.

Owyang, J. (2010). Social CRM: Not your father's CRM. *Forbes*. Retrieved September 24, 2018, from https://www.forbes.com/2010/04/09/facebook-twitter-social-media-crm-dell-comcast-cmo-network-jeremiah-owyang.html#225e9e303d89

Payne, A., & Frow, P. (2005). A strategic framework for customer relationship management. *Journal of Marketing, 69*(4), 167–176.

Phillips, E. (2015). Retailers scale up online sales distribution networks. *Wall Street Journal*. Retrieved September 6, 2018, from https://www.wsj.com/articles/retailers-scale-up-online-sales-distribution-networks-1447792869

Picot-Coupey, K., Huré, E., & Piveteau, L. (2016). Channel design to enrich customers' shopping experiences: Synchronizing clicks with bricks in an omnichannel perspective—The direct optic case. *International Journal of Retail and Distribution Management, 44*(3), 336–368.

Pires, G. D., Stanton, J., & Rita, P. (2006). The internet, consumer empowerment and marketing strategies. *European Journal of Marketing, 40*(9/10), 936–949.

Popomaronis, T. (2017). E-Commerce in 2018. Here's what the experts are predicting. *Forbes*. Retrieved September 12, 2018, from https://www.forbes.com/sites/tompopomaronis/2017/12/15/e-commerce-in-2018-heres-what-the-experts-are-predicting/#65a9c2596deb

Powell, A., Noble, C. H., Noble, S. M., & Han, S. (2018). Man vs machine: Relational and performance outcomes of technology utilization in small business CRM support capabilities. *European Journal of Marketing, 52*(3/4), 725–757.

Prahalad, C. K., & Ramaswamy, V. (2004a). *The future of competition: Co-creating unique value with customers*. Boston: Harvard Business School Press.

Prahalad, C. K., & Ramaswamy, V. (2004b). Co-creation experiences: The next practice in value creation. *Journal of Interactive Marketing, 18*(3), 5–14.

Reinartz, W., Krafft, M., & Hoyer, W. D. (2004). The customer relationship management process: Its measurement and impact on performance. *Journal of Marketing Research, 41*(3), 293–305.

Rigby, D. (2011). The future of shopping. *Harvard Business Review, 89*(12), 65–76.

Rodriguez, M., Peterson, R. M., & Krishnan, V. (2018). Impact of CRM technology on sales Process behaviors: Empirical results from US, Europe, and Asia. *Journal of Business-to-Business Marketing, 25*(1), 1–10.

Rogers, E. M. (1962). *The diffusion of innovations*. New York: Free Press.

Ryan, D., & Jones, C. (2009). *Understanding digital marketing: Marketing strategies for engaging the digital generation*. London: Kogan Page.

Saarijärvi, H., Kannan, P. K., & Kuusela, H. (2013). Value co-creation: Theoretical approaches and practical implications. *European Business Review, 25*(1), 6–19.

Saghiri, S., Wilding, R., Mena, C., & Bourlakis, M. (2017). Toward a three-dimensional framework for omni-channel. *Journal of Business Research, 77*, 53–67.

Schmitt, B. (1999). Experiential marketing. *Journal of Marketing Management, 15*(1–3), 53–67.

Smith, K. T. (2011). Digital marketing strategies that Millennials find appealing, motivating, or just annoying. *Journal of Strategic Marketing, 19*(6), 489–499.

Straker, K., Wrigley, C., & Rosemann, M. (2015). The role of design in the future of digital channels: Conceptual insights and future research directions. *Journal of Retailing and Consumer Services, 26*, 133–140.

Taiminen, H., & Karjaluoto, H. (2015). The usage of digital marketing channels in SMEs. *Journal of Small Business and Enterprise Development, 22*(4), 633–651.

Takacs, S. J., & Freiden, J. B. (1998). Changes on the electronic frontier: Growth and opportunity of the World-Wide Web. *Journal of Marketing Theory and Practice, 6*(3), 24–37.

Tiago, M. T. P. M. B., & Veríssimo, J. M. C. (2014). Digital marketing and social media: Why bother? *Business Horizons, 57*(6), 703–708.

Tiffin, J., & Rajasingham, L. (2003). *The Global Virtual University*. London: RoutledgeFalmer.

Trainor, K. J. (2012). Relating social media technologies to performance: A capabilities-based perspective. *Journal of Personal Selling and Sales Management, 32*(3), 317–331.

Trainor, K. J., Andzulis, J. M., Rapp, A., & Agnihotri, R. (2014). Social media technology usage and customer relationship performance: A capabilities-based examination of social CRM. *Journal of Business Research, 67*(6), 1201–1208.

Trusov, M., Bodapati, A. V., & Bucklin, R. E. (2010). Determining influential users in internet social networks. *Journal of Marketing Research, 47*(4), 643–658.

Udo, G. J., Bagchi, K. K., & Kirs, P. J. (2010). An assessment of customers' e-service quality perception, satisfaction and intention. *International Journal of Information Management, 30*(6), 481–492.

Vargo, S. L., & Lusch, R. F. (2004). Evolving to a new dominant logic for marketing. *Journal of Marketing, 68*(1), 1–17.

Verhoef, P. C., Kannan, P. K., & Inman, J. J. (2015). From multi-channel retailing to omni-channel retailing: Introduction to the special issue on multi-channel retailing. *Journal of Retailing, 91*(2), 174–181.

Ward, T. (2018). How much social media is too much? *Fortune*. Retrieved September 8, 2018, from https://www.forbes.com/sites/tomward/2018/06/08/how-much-social-media-is-too-much/#299a7b3360e6

Witell, L., Kristensson, P., Gustafsson, A., & Löfgren, M. (2011). Idea generation: Customer co-creation versus traditional market research techniques. *Journal of Service Management, 22*(2), 140–159.

Wongsansukcharoen, J., Trimetsoontorn, J., & Fongsuwan, W. (2015). Social CRM, RMO and business strategies affecting banking performance effectiveness in B2B context. *Journal of Business and Industrial Marketing, 30*(6), 742–760.

World Bank. (2018). Individuals using the Internet (& of population). Retrieved September 7, 2018, from https://data.worldbank.org/indicator/it.net.user.zs

Wymbs, C. (2000). How e-commerce is transforming and internationalizing service industries. *Journal of Services Marketing, 14*(6), 463–477.

Yi, Y., & Gong, T. (2013). Customer value co-creation behavior: Scale development and validation. *Journal of Business Research, 66*(9), 1279–1284.

Yoffie, D. B., & Cusumano, M. A. (1999, January–February). Judo strategy: The competitive dynamics of internet time. *Harvard Business Review, 77*, 70–81.

Zeithaml, V. A., Berry, L. L., & Parasuraman, A. (1996). The behavioral consequences of service quality. *Journal of Marketing, 60*(2), 31–46.

CHAPTER 12

An Evaluation of the National Open Government Data (OGD) Portal of the United Arab Emirates

Stuti Saxena

Introduction

Conceived both as a philosophy and as a set of policies, open government data (OGD) relates to the data published by the government which is freely available for being re-used by different stakeholder groups (citizens, journalists, researchers, entrepreneurs, etc.) to derive value out of the same and improvise upon their services (Alexopoulos, Loukis, & Charalabidis, 2014; Attard, Orlandi, Scerri, & Auer, 2015; Kassen, 2013). Such data may relate to different socio-economic sectors like weather, agriculture, industry, energy, power, education, trade and so on. (Ubaldi, 2013). OGD is an advanced format of e-government which promotes transparency, accountability and citizen trust because the government publishes the hitherto-reserved data for wider usage (Mpinganjira, 2015). Apart from facilitating citizen engagement, OGD initiatives help in improvising upon public policies and contribute toward efficiency,

S. Saxena (✉)
Central University of Haryana, Mohindergarh, India
e-mail: stutisaxena6691@cuh.ac.in

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_12

economy and effectiveness of the government. As such, open data may be defined as "machine-readable data (which is) discoverable, available, and downloadable through dedicated internet portals without cost to potential data users" (Dawes, Vidiasova, & Parkhimovich, 2016, p. 15). For earning benefits from the OGD initiative, it is important that the datasets are published in a timely manner and they should be complete in all aspects (e.g., metadata, data entries, legible formats, provision of visualization and analytical tools, etc.) (Jaeger, Bertot, & Shilton, 2012).

Open data initiatives spur innovation through co-creation activities of the government and the other stakeholders (Verhulst & Young, 2016). For instance, software developers and the government collaborate on the social coding platform GitHub to contribute toward the revamped version of the source code (Mergel, 2015). It has also been underlined that sustainable OGD initiatives help in the economic growth of the country (Charalabidis, Alexopoulos, & Loukis, 2016; Jung & Park, 2015; Saxena, 2017a; Wirtz & Birkmeyer, 2015). Implicitly, the OGD initiative of the United Arab Emirates (UAE) has the potential of facilitating the economic diversification of the country and meeting the goals of the Vision 2021 espoused by the government. However, to assess the potential of the OGD initiative, it is important that the datasets published by the government online are qualitatively and quantitatively adequate. This would facilitate optimum re-use of the datasets for deriving social and economic value out of them.

In line with the aforesaid, this chapter invokes the usability framework (Machova, Hub, & Lnenicka, 2018) and the website quality framework (Sorum, Andersen, & Clemmensen, 2013) to present a case study where we evaluate the national OGD portal of the UAE (https://bayanat.ae/) in terms of the data quality and underline the drivers and barriers in re-using the datasets published via the portal. The key research question guiding the case study is: "What are the drivers and barriers in re-using the datasets published via the national OGD portal of UAE?" Further, the findings from the case study emphasize the need for maintaining the quality of the datasets published therein. As the first detailed study to investigate the national OGD portal of the country, the case study is a significant contribution to the OGD literature.

The chapter is structured as follows: After summarizing the literature on OGD, a brief shall be provided about the research methodology invoked in the chapter wherein the case study on the UAE's national OGD portal shall be elaborated. Thereafter, concluding remarks shall be provided to summarize the major inferences. The penultimate sections of the chapter shall detail the social and practical implications of the case study and leave pointers for further research.

Related Research

OGD is an emerging phenomenon and academic interest in the field has been increasing over a period of time. While most of the studies are conducted in the West, where open data initiatives are being undertaken at an advanced level, studies in the developing countries are few and far between (Huijboom & Van den Broek, 2011; Janssen, Charalabidis, & Zuiderwijk, 2012; Saxena, 2017a; Saxena & Janssen, 2017; Zuiderwijk & Janssen, 2014). Implicitly, there is a need for conducting more research in the developing countries to underline the impediments in the implementation of the OGD initiative (Charalabidis et al., 2016) besides drawing lessons from the countries that have been successful in the implementation of the open data initiatives (Nugroho, Zuiderwijk, Janssen, & de Jong, 2015). Such contextual studies, including the empirical ones, provide cases of the success or failure of OGD initiatives. For instance, Rothenberg (2012) followed a case study approach wherein the OGD frameworks of the US, the UK, Canada and New Zealand were compared. The purpose of selecting these four countries lay in the marked similarities of culture among them. Likewise, a study based in Africa aimed at understanding the extent of OGD implementation regarding the characteristics of an "ideal" OGD portal and the study was conducted in five (Ghana, Kenya, Sierra Leone, South Africa and Tanzania) out of the seven (Ghana, Sierra Leone, Tunisia, Morocco, South Africa, Kenya and Tanzania) OGD centers (Afful-Dadzie & Afful-Dadzie, 2017). In another study, the Unified Theory of Acceptance and Use of Technology (UTAUT) model was adapted in order to conduct an empirical investigation in the Netherlands and identify the factors which impact the usage and adoption of OGD (Zuiderwijk, Janssen, & Dwivedi, 2015). In an empirical investigation conducted on a sample of 210 citizens in Germany, the extended Technology Acceptance Model (TAM) was deployed to underline the relationship between the constructs (ease of use, usefulness, transparency expectancy, participation expectancy, collaboration, intention to use OGD and word-of-mouth intention concerning OGD) and it was found that all the five constructs were significant predictors of intention to use OGD and to collaborate among each other using word-of-mouth (Wirtz, Weyerer, & Rosch, 2017). In another empirical investigation conducted in India with a sample of 244 respondents, OGD use and acceptance among different users was probed and it was concluded that OGD use has increased in

the country and men are more likely to tap OGD than women (Saxena & Janssen, 2017). All these studies underlined the reasons as to why OGD initiatives are yet to realize their cherished aims in terms of promoting accountability in administration and furthering citizen participation in policy-making.

Literature on OGD has provided an understanding that the demand-supply equation of any OGD initiative merits a careful consideration and this entails that the legal provisions regarding publication of datasets and the concern for privacy of the individuals (Martin, 2014). For instance, datasets are "supplied" by the government agencies and datasets may also be "suggested" or "recommended" by the users. Likewise, datasets may be "demanded" by the stakeholders for their purposes. The "supply" side may be bolstered by ensuring that the entire OGD publication value-chain is qualitatively and quantitatively superior and the "demand" side may be strengthened by the user engagement and involvement in the OGD initiative.

Many studies have provided frameworks for investigating OGD initiatives. For instance, Martin (2014) has provided a framework for underlining the social and technological aspects of OGD initiative. The framework includes five dimensions—digital technologies (configurations that include tangible artifacts, the skills of technologists and users and the interfaces of artifacts with the wider technical infrastructure), user practices (manner in which data is being re-used by a large cross-section of stakeholders), public management practices (includes the processes of data and ICT management and established data-related policies), institutions (include the sets of rules that connect data users and government organizations, including data markets and regulatory frameworks for government data) and resources (the resources drawn upon by actors shaping the OGD data agenda as including: social capital (the networks that connect actors); cultural capital (cultural goods and services, such as skills and knowledge); economic capital (money and other assets that can be directly and immediately converted to money); and symbolic capital: the means available on the basis of (perceived) prestige or legitimacy.

Another study provides a model which outlines the stages in the roll-out of any OGD initiative (Kalampokis, Tambouris, & Tarabanis, 2011a; 2011b). Kalampokis et al. (2011a) have classified OGD into "downloadable files" (data is available in simple formats), "linked data" (data is linked

with another one and re-used), "direct data provision" (all data is available via a portal and synchronized with time) and "indirect data provision" (actual data is provided and the user is responsible for further aggregation and processing of the data). Therefore, a four-stage model (aggregation of government data; integration of government data; integration of government data with non-government data; and integration of government data with non-government formal and social data) is identified by Kalampokis et al. (2011a). However, the drawback of the study lies in being more technical in approach wherein the emphasis was laid down on the crude mechanistic dimensions of linked data.

Saxena (2017b) provided a typology of countries on the basis of their OGD-adherence ("laggard," "caged," "forerunner" and "champ") wherein the "laggard" countries are the ones where there are hindrances associated with OGD implementation and OGD-usage; "caged" countries are those with less propensity to implement an OGD initiative but increased potential of usage by different stakeholders; "forerunner" countries as those which hold high potential of rolling out an OGD initiative but low potential of usage by different stakeholders; and "champ" countries as those which ranked high in terms of implementation of an OGD program as well as usage by a diverse set of stakeholders (p. 219).

In another study, another model expounds the degree of interaction among the stakeholders involved in the OGD initiatives directly or indirectly (Sieber & Johnson, 2015). Thus, Sieber and Johnson (2015) provide four models which deal with the "nature of (OGD) delivery (which) shapes the way the data is used" (Sieber & Johnson, 2015, p. 310). Table 12.1 summarizes the four models.

While the aforementioned models are descriptive, the overall comprehensiveness in terms of quality assessment of the datasets is missing. Therefore, the present case study seeks to adopt the evaluation framework proposed by Machova and his colleagues (2018). In their benchmarking framework derived from the reading of the literature on OGD, three dimensions (criteria) and 14 sub-criteria were defined which were a function of data discoverability, data accessibility and reusability. Table 12.2 reproduces the framework proposed by them.

Apart from this, the case study underlines the significance of maintaining the quality of the website dedicated for the publication of datasets. For this purpose, the website quality framework (Sorum et al., 2013) is being referred here. The basic premise of the framework is that the

Table 12.1 Citizen engagement models proposed by Sieber and Johnson (2015)

Model	*Description*
Data over the wall	This is the basic citizen engagement model wherein the government publishes datasets on the online portal directly. Basic features like downloading of datasets, visualization, mapping or sorting are allowed. Datasets are available in formats like PDF, Excel and so on. Programmatic access may be permissible via software-to-software interface (i.e., application programming interface or API). Users are encouraged to report errors in the form of feedback.
Code exchange	Government encourages the re-use of datasets for innovating products and services. Therefore, the government publicizes and promotes the OGD initiative by holding events, conferences, workshops or "app" contests. Software or application "app" developers, civic hackers and social entrepreneurs are encouraged to participate in such promotional activities.
Civic issue tracker	Citizens contribute to the existing datasets in many ways. For instance, citizens report of civic problems (e.g., fire, accidents, drainage problems, floods, potholes, etc.) which necessitate immediate action by the government authorities. By encouraging participation of citizens in the OGD initiative, the government is regularly informed about the civic issues and this facilitates in revising and updating the datasets (Alexopoulos et al., 2014; Dawes et al., 2016).
Participatory open data	As an ideal model wherein citizens and governments enter into a dialogue via the OGD initiative, there is active engagement and participation of citizens in the policy-making and policy-implementation stages. Citizens are encouraged to freely contribute toward the existing datasets via the online portal. Furthermore, all grievances of the citizens pertaining to the quality of the datasets are rectified in the prescribed manner. Datasets are qualitatively superior in this model with metadata and permit statistical analysis, interpretation, visualization and mapping.

website quality is linked with user-friendliness, effective website usage and content-related issues and accessibility. Four types of websites have been identified in the study (Table 12.3).

Furthermore, the website quality dimensions have been identified as presented in Table 12.4.

Table 12.5 summarizes the website attributes in terms of the quality parameters.

Table 12.2 Model proposed for evaluating OGD portals (Machova et al., 2018)

Dimension	*Criteria*	*Description*
1. *Open dataset specifications*	a. *Description of dataset*	Portal provides datasets together with their description and how and for what purpose they were collected
	b. *Publisher of dataset*	Portal provides information about organization that published datasets
	c. *Thematic categories and tags*	Portal provides thematic categories of datasets to address the main topics covered. It distinguishes categories (themes) from tags (keywords)
	d. *Release date and up to date*	Portal provides datasets associated with a specific time or period tag, that is, date published, date updated and its frequency
	e. *Machine readable formats*	Portal provides datasets formats that are machine-readable and allow for easy re-use
	f. *Open data license*	Portal provides license information related to the use of the published datasets
	g. *Visualization and analytics tools*	Portal provides visualization and analytics capabilities to gain information about a dataset, e.g., in charts or visualizations in maps
2. *Open dataset feedback*	a. *Documentation and tutorials*	Portal provides high quality of documentation and tutorials to help users in learning how to use the portal
	b. *Forum and contact form*	Portal provides an opportunity to submit feedback on a dataset from the users to providers and forum to discuss and exchange ideas among the users
	c. *User rating and comments*	Portal provides capabilities allowing the collection of user ratings and comments on a dataset
	d. *Social media and sharing*	Portal provides the integration with social media technologies to create a distribution channel for open data and sharing feedback
3. *Open dataset request*	a. *Request form*	Portal provides a form to request or suggest new type or format type of open data
	b. *List of requests*	Portal provides a list of requests that were received from users, including the current state of request processing
	c. *Involvement in the process*	Portal provides capabilities allowing the involvement in the active requests, that is, express interest in the same dataset

Table 12.3 Four types of websites and their attributes

Website	*Main attributes*
The inviting website (InvW)	Website should always make it possible for the user to make the right choice; Users should be able to find information fast as part of effective use; Website development is a continuous process where there is always room for improvement; Website should be lucid and it should be easy to find things; Personalization and speed; Well-developed content and use of pictures; Users should be satisfied and provided with help and support; More use of interactivity and new ways of presenting and visualization of web content; Presenting search functions in order to make it easier for finding relevant information, personalizing web content and making updates and continuous improvements
The intuitive website (IW)	Website appears to have a "logical structure" and this helps in the easy access and free movement across the website; Users are able to easily find information; Regular updating of the website contents
The easy-to-use website (EUW)	Website should be lucid, and it should be easy to find things on the web site; Supports user-satisfaction and efficiency of use; Information presented on the web site must be currently valid' Personalization of web content; Interactivity; Provision of true, updated and trustworthy information; Use of easy language; Providing faster response for users, fast mail response and feedback to users and deliberate responses; Fast response and feedback to users; Presenting content that users are looking for and expecting to find at websites
The simple website (SW)	Accessibility is highly important, together with user adoption and website content; Users should be able to find the services they need; Content should be organized in a way that clearly separates the different parts from one another; Content should be "easy to find" and well "organized" so that users can discover and retrieve relevant information and services; There is data dependability

Table 12.4 Quality aspects and their dimensions

Quality aspects	*Dimensions*
Information quality (captures the website/application content)	Content Dependability Updated
System quality (measuring the desired characteristics of a system or a website)	Accessibility (standard) Effective use Personalization Search functions Speed Uptime User-friendly
Service quality (covers the overall support delivered by the service provider)	Fast mail response Fast response for the users Service quality (easy help)

Table 12.5 Assessment of website types in terms of the attributes

Major attributes	*Simple website (SW)*	*Easy-to-use website (EUW)*	*Intuitive website (IW)*	*Inviting website (InvW)*
Accessibility	✓	✓	✓	✓
Content	X	✓	✓	✓
Deliberate	X	✓	X	X
Dependability	✓	✓	X	✓
Effective use	✓	✓	✓	✓
Fast mail response	X	✓	X	X
Faster response for the users	X	✓	X	X
Interactivity	✓	✓	X	✓
Personalization	X	✓	X	✓
Search functions	X	X	X	✓
Service quality (easy help)	X	✓	X	✓
Speed	✓	X	X	✓
Updated	X	✓	✓	X
Uptime	✓	X	X	X
User-friendly	✓	✓	✓	✓

Evaluation of the National OGD Portal: A Case Study of the UAE's OGD Initiative

UAE occupies the 17th spot out of 137 countries in the Global Competitiveness Report (2017–2018). The country has been adopting economic diversification drives. Given the prospects of OGD initiatives in furthering economic growth, we posit that the OGD initiative of the country has the potential of aiding the country in its economic diversification moves. Apart from furthering economic growth, the OGD initiative should facilitate citizen participation and collaboration in improvising public services and making them more efficient and effective. Datasets published under the aegis of the OGD initiative may be re-used by a diverse set of stakeholders (academic community, scientists, researchers, public/private/non-profit professionals, journalists, software developers, etc.). By re-using the datasets, the stakeholders may derive social and economic value and spearhead innovation. However, to realize the benefits of re-using the datasets, it is important that the datasets published by the government are of good quality and the government is proactive in instituting a robust OGD initiative by contributing toward publishing datasets that are qualitatively and quantitatively superior. Datasets should be complete in themselves and they should be published in a regular manner. This would ensure a sustainable re-use of the datasets by the stakeholders. Therefore, we present an evaluation assessment of the national OGD portal of United Arab Emirates (https://bayanat.ae/) as a case study in order to analyze the quality of the datasets published via the portal. Specifically, the case study addresses the question: "What are the drivers and barriers in re-using the datasets published via the national OGD portal of United Arab Emirates (UAE)?"

The Federal Competitiveness and Statistics Authority (FCSA) is responsible for managing and updating the portal. The portal has a search option wherein keywords may be provided for accessing the datasets. In all, there are 776 datasets covering themes like agriculture (113), labor force (55), education (45), climate (31), and so on. Furthermore, the datasets are provided by organizations like Ministry of Education (40), Ministry of Health (32), Zakat Fund (18), Zayed University (16), Ministry of Energy (15), and so on. Datasets may be filtered using tags like education (53), crop (33), zone (33), federal (31), gender (27), region (27), and so on. Furthermore, datasets are available in two formats: Excel (XLSX) (774) and Excel (XLS) (2). Software applications that may be tapped via open

data portal of the country include Aamen (an application for handling employees' complaints about their wages), Absher (an application for achieving the goals and indicators of the national agenda of the country), Dubai Culture (an application for connecting the classical with the modern world), and so on. There is a "Youth Data" platform wherein the innovative efforts of the young generation are being showcased (https://data.youth.gov.ae/en). Visualization of datasets is possible only for select datasets. The government has been arranging hackathons for furthering innovation using open data. The "privacy policy" mentions that any personal information such as name and email address shared by the visitors shall be reserved for confidential purposes only. As per the "terms of use" of the portal, the datasets are available for "free use and reuse for commercial and non-commercial purposes." The portal mentions that the following procedures should not be violated:

- Accessing details that are not intended to be provided to this user or logging into a server or an account that the user is not authorized to access.
- Attempting to conduct any test or survey for finding weakness of any system or network of the UAE government or violating applicable procedures or documenting them without an official permit from the UAE government.
- Attempting to interfere in the provided service on the part of any user, host or network including but not limited to placing a virus on the website, increasing load to or immersing it, sending commercial messages to it or avalanching it with electronic messages or even destroying it.
- Sending unwanted electronic messages to the website including commercials and/or advertisements on services or products or falsifying any dispatch control protocol package address/internet protocol or any part of the address details in any electronic message or sending news messages.
- Using this portal by any means for sending an email, anything of it or on its behalf, by referring to it, or assuming the identity of its name involving offence or libel of the UAE government, website or any person whomsoever, announcing any untrue news or information and ascribing it to the UAE government unrightfully. There is a "contact form" for contributing datasets and soliciting datasets.

Conducting evaluation studies by underlining the features of the national OGD portals has been common till now (See, for instance, Kassen, 2013; 2017; Saxena, 2018). This case study adopts the usability evaluation framework proposed by Machova and her colleagues (2018) as well as the website quality evaluation framework by Sorum et al. (2013). We chose to refer this framework for the present study for three reasons—it is a recent contribution which is relevant for the evaluation of any OGD portal; it is comprehensive and all-encompassing; and it has been derived after a review of literature which shows that it is well-researched and may be subject to further reference for academic research. Table 12.6 summarizes the key

Table 12.6 Major facilitators and hindrances in re-using datasets via the national OGD portal of UAE

Criteria	*Facilitators to re-use datasets via the national OGD portal*	*Hindrances to re-use datasets via the national OGD portal*
1. *Open dataset specifications*		
a. *Description of dataset*	Portal is accessed in two languages—Arabic and English. There is a provision to "search" datasets via the main page. Metadata in the form of data publication date, last revision done, identifiers, spatial/geographical coverage, authorship and public access level are limited.	Portal provides datasets together with their description and how and for what purpose they were collected. However, real-time data is not being provided via the portal as of now.
b. *Publisher of dataset*	In all, there are 776 datasets covering themes like agriculture (113), labor force (55), education (45), climate (31), and so on. Furthermore, the datasets are provided by organizations like Ministry of Education (40), Ministry of Health (32), Zakat Fund (18), Zayed University (16), Ministry of Energy (15), and so on.	Not all listed publishers have contributed to the portal.
c. *Thematic categories and tags*	Datasets may be filtered using tags like education (53), crop (33), zone (33), federal (31), gender (27), region (27), and so on.	Portal provides thematic categories of datasets to address the main topics covered. It distinguishes categories (themes) from tags (keywords)

(*continued*)

Table 12.6 (continued)

Criteria	*Facilitators to re-use datasets via the national OGD portal*	*Hindrances to re-use datasets via the national OGD portal*
d. *Release date and up to date*	Metadata reveals some information regarding the date of publication of the datasets.	It is important that the metadata be relevant and complete in all aspects.
e. *Machine readable formats*	Datasets are available in two formats: XLSX (774) and XLS (2).	User-friendly formats should be provided for analysis and statistical interpretation.
f. *Open data license*	Data are regarded as a "natural resource" and should be available for all users.	N/A
g. *Visualization and analytics tools*	Datasets may be arranged in an ascending or descending order.	Visualization and mapping tools are required for optimum re-use of the datasets.
2. *Open dataset feedback*		
a. *Documentation and tutorials*	The manual of "Open Data Policy" provides information regarding the OGD initiative of the country and the utility of the same for the different stakeholders.	N/A
b. *Forum and contact form*	There is a "Contact" form where the users/portal visitors may place requests or contribute datasets.	A blog may be provided where the users may interact with each other and lend constructive views for the improvisation of the portal.
c. *User rating and comments*	Datasets may be ranked (five-star rating).	Users may provide "feedback" regarding the website via the "Contact" form.
d. *Social media and sharing*	Social media networks are not provided.	Social networking plugins may be provided (Twitter, Facebook and YouTube) to facilitate the sharing and discussion of datasets.
3. *Open dataset request*		
a. *Request form*	Only the government agencies may place "requests" for publishing the datasets.	Other stakeholders should be involved in the data publishing process.
b. *List of requests*	The provision of providing the list of requests made by the users is missing in the portal.	The portal should provide information regarding the recommendations and contributions of the users to the datasets via the portal.

(*continued*)

Table 12.6 (continued)

Criteria	*Facilitators to re-use datasets via the national OGD portal*	*Hindrances to re-use datasets via the national OGD portal*
c. *Involvement in the process*	Public participation has been encouraged wherein recommendations from the portal visitors are invited for supporting the OGD initiative of the country.	Events like contests for the stakeholders to improvise upon the datasets should be organized. For instance, software apps may be devised by the software professionals. Businessmen may showcase how they harnessed datasets for improvising their services. Implicitly, incentives should be provided for increased involvement of the stakeholder groups.

observations for the national OGD portal of the country in line with the model (Machova et al., 2018). From the table, it is clear that the OGD portal provides opportunities to re-use datasets and the re-use of the datasets by a large cross-section of stakeholders might result in innovation and improvisation of goods and services thereby contributing to economic development of the country. At the same time, there are challenges in re-using the datasets because the datasets need to be updated on a regular basis and a more proactive stance of the government agencies is required to facilitate publishing and interoperability of the datasets.

Table 12.7 summarizes the key observations from the perspective of Sorum et al. (2013). These observations are derived after a careful perusal of the national open data portal of UAE.

Conclusion

The case study sought to present an evaluation of the national OGD portal of the UAE. After a brief scan of the literature on OGD, the case study provided a brief regarding the drivers and barriers to re-use the datasets published via the OGD portal conceding that data re-use by a diverse group of stakeholders is one of the prime rationale for publishing them and making them available for all. The evaluation of the OGD portal was conducted using the recently-proposed framework (Machova et al., 2018)

Table 12.7 Key insights about the portal using the framework of Sorum et al. (2013)

Major attributes	*Major insights derived from the portal analysis*
Accessibility	The national open data portal is well-accessible for use and reference by the relevant stakeholders. Datasets must be made more accessible in terms of their qualitative and quantitative aspects.
Content	The content of the open data portal should be well-maintained and updated on a regular basis.
Deliberate	The national open data portal of the UAE needs to provide real-time information.
Dependability	It is important that the national open data portal provides datasets that are reliable and dependable.
Effective use	The open data portal must provide more user-friendly features for facilitating the re-use of datasets.
Fast mail response	It is important that the mail responses be provided at the right time.
Faster response for the users	The national open data portal must ensure that responses to any queries made by the visitors are being answered promptly.
Interactivity	The national open data portal must facilitate interactivity by incorporating more user-friendly features.
Personalization	The content of the national open data portal must ensure that there is personalization for ensuring that the visitors perceive that the portal is serving their ends well.
Search functions	The website offers a "search" facility so that the specific dataset may be accessed easily depending upon the keyword/s used.
Service quality (easy help)	The website needs to improve on this feature.
Speed	The national open data portal should provide datasets in a speedier manner.
Updated	It should be the priority of the government to publish updated datasets.
Uptime	It would be a better exercise if efforts are made to publish linked datasets on a real-time basis.
User-friendly	The website should provide user-friendly features in order to facilitate a wider range of stakeholders to re-use the datasets.

and the framework proposed by Solem et al. which are quite comprehensive. It may be deduced from this evaluation of the OGD platform that while the OGD initiative of the country is a forthcoming measure of the government for promoting transparency in administration and forging ties with the citizens, the portal needs to be improvised in terms of the quality of the datasets. Datasets should be published on a timely basis and regular

updates should be made whenever required. For ensuring increased participation and engagement of the users, incentives should be provided for data re-use. Users' contribution to the datasets should be encouraged and interaction via the platform is also encouraged such that the users may interact among themselves and discuss about the improvisation of the goods and services through data re-use. Datasets lack visualization and mapping tools and this is a bottleneck in analyzing the datasets. Likewise, datasets are not amenable to statistical interpretation on account of the absence of compatible features which may facilitate the re-use of datasets in an optimum manner.

Social and Practical Implications

The case study underlines that the success of any OGD initiative lies in its sustainability. For ensuring the sustainability of the OGD initiative, it is important that the datasets be re-used for creating value out of the same. For instance, datasets may be re-used by a cross-section of stakeholders (citizens, journalists, software experts, academia, entrepreneurs, etc.) for appreciating public administration and providing recommendations for policy-making and improvisation in public service delivery. At the same time, they may tap the datasets for improvisation and innovation of their services. On the other hand, the policy-makers may solicit views from the users regarding the portal and the quality of the datasets published via the same and this would help in improving upon the administrative services. In turn, this would lead to better relationship between the policy-makers and the citizens. It is also important for the government to ensure that a robust ICT infrastructure is in place for sustaining the OGD initiative. Thus, on the one hand, the public officials managing the OGD initiative should be trained with regard to the maintenance and monitoring of the OGD portal and data publishing process; and on the other hand, technological advancement should be upheld for providing real-time linked datasets in near future. This would further improvisation in public service delivery and facilitate citizen participation in the co-creation of services.

Further Research Directions

The case study leaves pointers for further research. For instance, further research is merited to assess the manner in which the OGD initiative is being perceived by the different stakeholder groups. Empirical

investigation is required to appreciate the utility of the datasets for different stakeholder groups. Future research may be conducted to analyze the perspectives of the government representatives regarding the factors that impact the data publishing processes. A comparative approach may be undertaken to analyze how lessons may be drawn from the successes of OGD initiatives elsewhere. Research cues may be derived from the significance of data mining and social networks in furthering OGD discussions by the stakeholders. Likewise, it may be interesting to evaluate the significance of Big Data analytics in understanding the efficacy of OGD initiatives (Saxena, 2016). Finally, further research is warranted to analyze the utility of the implementation of OGD initiatives at the regional and local levels of the country.

References

Afful-Dadzie, E., & Afful-Dadzie, A. (2017). Open government data in Africa: A preference elicitation analysis of media practitioners. *Government Information Quarterly, 34*(2), 244–255.

Alexopoulos, C., Loukis, E., & Charalabidis, Y. (2014). A platform for closing the open data feedback loop based on Web 2.0 functionality. *eJournal of eDemocracy & Open Government, 6*(1), 62–68.

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly, 32*(4), 399–418.

Charalabidis, Y., Alexopoulos, C., & Loukis, E. (2016). A taxonomy of open government data research areas and topics. *Journal of Organizational Computing and Electronic Commerce, 26*(1/2), 41–63.

Dawes, S. D., Vidiasova, L., & Parkhimovich, O. (2016). Planning and designing open government data programs: an ecosystem approach. *Government Information Quarterly, 33*(1), 15–27.

Huijboom, N., & Van den Broek, T. A. (2011). Open data: An international comparison of strategies. *European Journal of ePractice, 12*(1), 4–16.

Jaeger, P. T., Bertot, J. C., & Shilton, K. (2012). Information policy and social media: Framing government—Citizen Web 2.0 interactions. In C. G. Reddick & S. K. Aikins (Eds.), *Web 2.0 technologies and democratic governance* (pp. 11–25). New York: Springer.

Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management, 29*(4), 258–268.

Jung, K., & Park, H. W. (2015). A semantic (TRIZ) network analysis of South Korea's "Open Public Data" policy. *Government Information Quarterly, 32*(3), 353–358.

Kalampokis, E., Tambouris, E., & Tarabanis, K. (2011a). *Open government data: A stage model*. Paper presented at International Conference on Electronic Government, 28 August–2 September 2011, Springer, Delft, The Netherlands.

Kalampokis, E., Tambouris, E., & Tarabanis, K. (2011b). A classification scheme for open government data: Towards linking decentralised data. *International Journal of Web Engineering and Technology, 6*(3), 266–285.

Kassen, M. (2013). A promising phenomenon of Open Data: A case study of the Chicago Open Data Project. *Government Information Quarterly, 30*(4), 508–513.

Kassen, M. (2017). Open data in Kazakhstan: Incentives, implementation and challenges. *Information Technology & People, 30*(2), 301–323.

Machova, R., Hub, M., & Lnenicka, M. (2018). Usability evaluation of open data portals: Evaluating data discoverability, accessibility, and reusability from a stakeholders' perspective. *Aslib Journal of Information Management*. https://doi.org/10.1108/AJIM-02-2018-0026

Martin, C. (2014). Barriers to the open government data agenda: Taking a multi-level perspective. *Policy & Internet, 6*(3), 217–240.

Mergel, I. (2015). Open collaboration in the public sector: The case of social coding on GitHub. *Government Information Quarterly, 32*(4), 464–472.

Mpinganjira, M. (2015). Use of e-government services: The role of trust. *International Journal of Emerging Markets, 10*(4), 622–633.

Nugroho, R. P., Zuiderwijk, A., Janssen, M., & de Jong, M. (2015). A comparison of national open data policies: Lessons learned. *Transforming Government: People, Process and Policy, 9*(3), 286–308.

Rothenberg, J. (2012). *Case study international benchmark: Towards a better supply and distribution process for open data*. Retrieved from Logica Business Consulting, The Netherlands.

Saxena, S. (2016). Integrating Open and Big Data via 'e-Oman': Prospects and issues. *Contemporary Arab Affairs, 9*(4), 607–621.

Saxena, S. (2017a). Significance of open government data in the GCC countries. *Digital Policy, Regulation and Governance, 19*(3), 251–263.

Saxena, S. (2017b). Prospects of open government data (OGD) in facilitating the economic diversification of GCC region. *Information and Learning Science, 118*(5/6), 214–234.

Saxena, S. (2018). Open government data (OGD) in six Middle East countries: An evaluation of the national open data portals. *Digital Policy, Regulation and Governance, 20*(4), 310–322.

Saxena, S., & Janssen, M. (2017). Examining open government data (OGD) usage in India through the UTAUT framework. *Foresight, 19*(4), 421–436.

Sieber, R. E., & Johnson, P. A. (2015). Civic open data at a crossroads: Dominant models and current challenges. *Government Information Quarterly, 32*(3), 308–315.

Sorum, H., Andersen, K. N., & Clemmensen, T. (2013). Website quality in government: Exploring the webmaster's perception and explanation of website quality. *Transforming Government: People, Process and Policy, 7*(3), 322–341.

Ubaldi, B. (2013). Open government data: Towards empirical analysis of open government data initiatives. *OECD Working Papers on Public Governance*, 22. https://doi.org/10.1787/5k46bj4f03s7-en

Verhulst, S., & Young, A. (2016). *Open Data impact, when demand and supply meet: Key finding of the open data. Impact case studies.* Retrieved from http://odimpact.org/key-findings.html

Wirtz, B. W., & Birkmeyer, S. (2015). Open government: Origin, development, and conceptual perspectives. *International Journal of Public Administration, 38*(5), 381–396.

Wirtz, B. W., Weyerer, J. C., & Rosch, M. (2017). Open government and citizen participation: An empirical analysis of citizen expectancy towards open government data. *International Review of Administrative Sciences.* https://doi.org/10.1177/0020852317719996

Zuiderwijk, A., & Janssen, M. (2014). Open data policies, their implementation and impact: A framework. *Government Information Quarterly, 31*(1), 17–29.

Zuiderwijk, A., Janssen, M., & Dwivedi, Y. K. (2015). Acceptance and use predictors of open data technologies: Drawing upon the unified theory of acceptance and use of technology. *Government Information Quarterly, 32*(4), 429–440.

CHAPTER 13

Artificial Intelligence and Robotics Technology in the Hospitality Industry: Current Applications and Future Trends

Li Yang, Tony L. Henthorne, and Babu George

Introduction

Artificial intelligence (AI) and robotics science have been penetrating into different kinds of businesses and industries in recent years. AI can be defined as the experimental and theoretical study of perceptual and intellectual processes using computers and to make a computer perceive, understand, and act in ways now possible only for humans (Achacoso, 1990). Robots, as one kind of human-like AI, were first introduced into manufacturing industries for product processing in order to replace monotonous human work decades ago (Murauskiene, 2015). For now, the applications of AI and robotics science are extending into other industries, for example, into transportation as autonomous vehicles (Maurer,

L. Yang (✉) • T. L. Henthorne
University of Nevada, Las Vegas, Las Vegas, NV, USA
e-mail: li.yang@univ.nevada.edu; tony.henthorne@unlv.edu

B. George
Fort Hays State University, Hays, KS, USA
e-mail: bpgeorge@fhsu.edu

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_13

Gerdes, Lenz, & Winner, 2016), into the medical industry for diagnoses and surgeries (Kaur, 2012; Mirheydar & Parsons, 2013), into the education industry for teaching or helping special kids (Fridin & Belokopytov, 2014; Timms, 2016), into the supply chain management for organizing warehouse (Min, 2010), and so on.

In the realm of hospitality and tourism, AI and robots have also arrived and they are catching peoples' attention. The world's first robotics hotel, The Henn-na Hotel, was opened in 2015 in Japan, which is fully staffed by robots (Rajesh, 2015). Meanwhile, in the United States, the Wynn Las Vegas announced in December 2016 that they are going to partner with Amazon to equip all of their hotel rooms with an Amazon Echo, a voice activated system (Raz, 2016). Besides the hotel sector, AI and robotics science are finding their way into other sectors of hospitality. In 2015, Eatsa, a futuristic restaurant where robot cubbies serve quinoa bowls, opened in San Francisco, and later opened another restaurant in NYC (Constine, 2015; Kowitt, 2016). A new bar on the Strip of Las Vegas, Tipsy Robot, featuring two robot bar tenders, started their business on June 30th, 2017 (Stapleton, 2017). AI and robotics science not only offers numerous opportunities to hospitality firms to enhance their daily operations, but also ensures hospitality firms provide consistent quality of products and services to their customers (Perez, 2014). The application of AI in the hospitality industry has caught a few, although not a sufficient amount of, scholars' attention (Borràs, Moreno, & Valls, 2014).

This paper will review current literature in the area of the application of AI and robotics science in the hospitality industry. Our examination will focus on the three areas of the application of AI and robotics science. First, it examines the current status and current application of robotics science in human daily life, highlighting their applications in each segment of the hospitality industry; further, it evaluates the impact of AI and robots on three areas: labor and employment, customers, and corporations and communities; lastly, it explores the future potential application of robots and AI in order to see what benefits or risks can be derived from their application in the hospitality industry.

Business Application of Robotics

The Robot Institute of America, a trade association of robot manufacturers and users, defines robots as follows: "a reprogrammable multifunctional manipulator designed to move material, parts, tools, or specialized

devices, through variable programed motions for the performance of a variety of tasks." (OTA, 1982, p. 8). Currently, robots are labeled as "intelligent physical devices" (Chen & Hu, 2013, p. 161) which have the capability of autonomy, mobility, and sensory to execute intended duties (International Organization for Standardization, 2012; Tan, Mohan, & Watanabe, 2016). Robotics science is definitely changing the way that humans live and work all over the world, and the application area of AI and robotics science is widening (Veruggio & Operto, 2008). Robots are extending their way out of the factories and into normal life: hospitals, schools, shopping malls, airports, hotels, restaurants, and even our homes. Robots are not just automated machines anymore; they are becoming consumer items (Veruggio & Operto, 2008). In certain scenarios, they are even more than an intelligent tool; they could even be considered a partner or a companion (Louie, McCaoll, & Nejat, 2014)

Types of Robots

The International Organization for Standardization (2012) claims most robots can be categorized into two groups: industrial robots and service robots. The categories are based on the robot's field of application as well as the industrial or non-industrial environments they operate in. Some scholars, further, separate service robots into two groups: professional service robots and personal service robots (Murphy, Hofacker, & Gretzel, 2017).

Industrial Robots

These are usually adopted to perform industrial tasks which can be indicated from the name. This is the initial field where robots were applied in manufacturing and production. According to the International Federation of Robotics, industrial robotics can be classified as either linear robots, Selective Compliance Assembly Robot Arm (SCARA) robots, articulated robots, parallel robots, or cylindrical robots based on the mechanical structure (IFR, 2016a).

Service Robots

"A service robot is a robot which operates semi or fully autonomously to perform services useful to the well-being of humans and equipment, excluding manufacturing operations" (IFR, 2016a). Different from industrial robots, service robots are designed to assist, serve, and interact with humans. International Federation of Robotics (IFR) claims that service

robots do not have to have arms, like industrial robots do. According to the definition, an industrial robot could also become a service robot when it is installed in a non-manufacturing environment (IFR, 2016b). Because of the increasing market demand and the different applications, the realm of service robots is expanding year by year, and the variety of service robots is becoming wider. Additionally, their diversity and application is becoming more varied each year. The shape and structure of service robots might change over time as well, but generally service robots can be classified into professional service robots and personal/domestic service robots (Ivanov, Webster, & Berezina, 2017; Murphy et al., 2017).

Professional Robots

From the name, they are deployed by formal organizations and companies. Professional service robots have a much bigger market than domestic/personal robots (IFR statistical department: World Robotics News, 2012). From those statistics, defense, field, logistic, and medical robots are the leading players for professional robots.

Personal/Domestic Robots

Personal/Domestic robots are robots used at home or at other private places by individuals for non-commercial purpose. There are two leading kinds of personal robots. One is cleaning robots, such as floor cleaning, lawn-mowing, and so on. The other is interactive robots, such as butler robots, that are usually equipped with autonomous navigation systems, touchscreens, voice, and face recognition systems. These additional systems can allow them to help or entertain people in public environments or deliver items (Ciupe & Maniu, 2014).

Current Status in Each Hospitality Segment

In the realm of the hotel and tourism industry, AI and robots have also arrived and are attracting great attention and interest from public (Ivanov et al., 2017). The current application of AI and robots in different segments of the hospitality industry, especially service robots will be discussed in this review.

Hotel

Service automation and robotic technologies have made their way into the lodging segment of the hospitality industry, affecting different areas of

hotel operations. At the front desk, hotel operators apply self-service kiosks which allow guests to check in and out by themselves in order to reduce waiting time and to improve service (Kim & Qu, 2014). Just like using a kiosk at the airport to get flight tickets, guests can quickly and efficiently check in to their room without human interaction. This is just the very basic application of service automation. In the concierge department, Connie, a robotic concierge agent, was introduced in Hilton Hotels last year, which is partnered with IBM. Connie is powered by AI and can communicate with hotel guests by answering their questions about hotel amenities and services and providing information for local attractions and activities. Because of the AI principle, Connie can also learn from its interactions with guests, and optimize its future answers (Hilton, 2016). In the housekeeping department, a robot called maidbot is created to provide better hotel room cleaning solutions (Maidbot, n.d.). A few hotels, such as Aloft at Cupertino, California, also implement the automated bellboy, a robot that delivers towels and bottled water at the requests of guests (Markoff, 2014).

Combining all these functions, as mentioned earlier in this paper, the world's first robotics hotel, the Henn-na hotel, was opened in 2015 in Japan (Henn-na Hotel, n.d.), which is fully staffed by robots. When guests check into their hotel room, two front desk agents (one human-like robot speaks Japanese, one dinosaur-like robot speaks English) will welcome them; automated trolleys take guest's luggage to their room; and a block-shaped robot does room service as well as delivers beverages and simple snacks.

Food and Beverage

AI and robots are also applied in the food and beverage industry, mostly in restaurant service and food preparation. In some restaurants, the food order is taken by robots (Curtis, 2016). Pizza Hut in Japan hires a humanoid robot named Pepper to take food orders from customers by voice. Pepper is equipped with a voice recognition and AI system, and talks with customers. In addition, Pizza Hut has partnered with MasterCard to program Pepper with a payment app. In other words, Pepper not only takes food orders from customers and sends them to the restaurant kitchen, but also processes orders for customers which saves time for both the restaurant and the customer. As mentioned earlier, a fully automated restaurant, Eatsa, which serves quinoa bowls, opened their first location in San Francisco. Customers place their order through an iPad in the lobby, and

their name will show up on one of the transparent LCD screen boxes. In less than 2 mins, customers can retrieve their personalized quinoa bowl out of the box. Similarly, Tipsy Robot, a new bar in Las Vegas, is staffed by two robot bar tenders that mix up and serve drinks (Stapleton, 2017). The guests just need to order their drinks on an iPad, and choose from 18 specialty cocktails or create their own drink. The two robotic bar tenders will get the drink ready for the customers in 60 to 90 seconds. In their spare time, the robots will slice fruit or dance to entertain the customers.

All of these robots utilized in the restaurant mitigate the customer's waiting time and eliminates human waiters and cashiers. Robotics technology definitely helps the food and beverage industry save labor cost and keep food quality consistent.

Meetings and Events

AI and robots are also applied in the meeting and convention industry to provide better service and make the demonstration of meetings easier. A lot of corporations and organizations employ interactive robots at their convention booths or at their trade show (S. P. Events, n.d.). Interactive robots at conventions are usually equipped with a mechanical head and arms, a voice and face recognition system, a message screen or iPad, a CD player, and a brochure holder. Some are also fully mobile so as to attract people's attention or sense people when they pass by. S.P. Events claims that this kind of robot is perfect for promotion events, exhibition stands, tradeshows, or products launched by using their AI tool to communicate with people, introduce products on a screen, or even just to greet people to get their attention. Moreover, this application creates a new robot rental business.

Another significant application of robots in the meeting and events industry are drones. For instance, the organizers of a music festival in South Africa deliver beer to over 150 guests by drone (Daily News, 2013). This application not only provides better service, but also saves the labor cost of delivery for the organizers. Drones are also used during weddings as photographers to take beautiful photos for the new couple.

These examples show that AI and robotics science not only offer numerous opportunities to hospitality firms to enhance efficiency of their daily operations, but also ensure they provide a consistent quality of products and services to their customers.

The Impact of Robots in the Hospitality Industry

Impact on Labor and Employment

Positive Impact

The rapid rise of advanced technology has caused deep uncertainty about the future of the job market (Yakowicz, 2016). There are many kinds of AI and robotics technologies in applied hotels, but there is very little academic research explicitly investigating the impact of robots on the workforce of the hospitality industry. Noone and Coulter (2012) argue that modern AI and robotics science can improve restaurant operations because of the two key features of robotics technologies: sensing for demand prediction and planning for production management. They also claim that robotics should be applied to augment workers' cognitive capability rather than direct labor replacement. Robots can help human staff work more efficiently, rather than replacing their jobs, which will positively impact the business (Pacurar, 2017). From a staff perspective, they also think robots do a great job helping out and making jobs better (Yakowicz, 2016).

Other scholars believe the application of AI in the point of sale system of quick service restaurants can avoid employee theft. The National Restaurant Association (1999) reported that the cost of employee theft for its members totaled over $8.5 billion, or 4 percent, of food sales. Stealing cash from an employer is a significant issue. From that number alone, manager and business runners should not underestimate the lure of quick money and the problem of stealing. Therefore, the most effective way to stop this kind of theft is to apply AI detection systems which aim to maintain solid cash controls for the restaurant (Collins, 2013)

Create New Jobs

Robots are becoming possible alternatives, and they have more of an economic advantage over human beings, in some fields. It is clear that robots are becoming a viable alternative to humans as the economic benefit of using a robot can far outweigh the use of a human. The impact of robots and AI could be dramatic on individuals due to different skillsets required in different industries (OTA, 1982). It is not surprising that labor-intensive and dangerous jobs are replaced by robots or AI. As robots are equipped with cognitive systems, the leading job replaced by robots might be transportation, logistics, customer service, and consumer services. Robots are

replacing humans in some of these jobs in order to reduce the human risk and prevent workers from being exposed to dangerous situations or hazardous environments (Chijindu & Inyiama, 2012; Yakowicz, 2016).

History shows the negative impact of technology on employment has been small enough to be insignificant or undetectable. For example, the wrongly created fear that was aroused in the early 1930s about whether vending machines will replace retail store clerks (Slade, 2007, p. 67). And the truth is people still go to the store. Hence, most experts believe robots or AI can create new jobs rather than replace jobs done by human beings (Veruggio & Operto, 2008; Chijindu & Inyiama, 2012). The production of robots and related technologies will create new jobs, but the amount of jobs created will depend on the growth of the robotics industry and the degree that robots can be automated or self-fixed.

Overall, the literature shows that robotics technology will eliminate some jobs, but create more jobs, thus life will get better in the long run.

Impact on Customers

The application of service robots in the hospitality industry is changing the nature of service, and influencing customers' service experiences and customers' relationships with service providers (Ostrom, Parasuraman, Bowen, Patrício, & Christopher, 2015; Rust & Huang, 2014). Because of this new environment, the interaction between customers and organizations might change significantly. To enhance customer experience, companies will need to optimize robotics technology for the customer's frontline experience, thus, to effectively engage customers and develop relationships. A few scholars are researching these topics. Pan, Okada, Uchiyama, and Suzuki (2015) conducted experiments that tested which style of speech attracted guests in a hotel. According to their experiment data "direct speech was more attractive than indirect speech at the beginning period of interaction between guests and the robot"; thus, they recommend the hotel assistive robots should be designed to provide direct verbal interaction with individual guests in order to engage with guests and to quickly attract their attention.

The Design of Human-like Robots or Machines

Other scholars investigate how the look of robots impacts customers. Mori, MacDorman, and Kageki (2012) suggest if artificial service agents look like a human too much, it would be perceived as creepy; therefore,

less warmth for the service. Likewise, Doorn et al. (2017) claim a similar impact will be perceived on the capability dimension of social cognition. In other words, the more the service robots look like a human, it would more likely be conceived as a human and feature lower capabilities. They also note a positive result that because of this verisimilitude, when the human-like robots cause service failures, the customers are more forgivable. Facial gender cues of human-like robots were also examined to investigate their influence on customer behaviors. Eyssel and Hegels' study (2012) shows that a short-haired, male robot is considered as more agentic while a long-haired, female robot is perceived as more communal. This result can inspire the design of service robots and have them change looks depending on what they are programmed to do. Other research is conducted to investigate if human-like features would influence if customers would repeat their service or not (Fan, Wu, & Mattila, 2016). Their results indicate that when customers are alone, they would prefer traditional services, such as talking to human through the phone instead of using touch screens or online services through the phone. This means hospitality firms can avoid adding human-like features to machines in private areas, such as in-room devices. Conversely, adding human-like machines could be practical when used in public, such as a self-service kiosk. Their findings also show that when serving powerful customers, hotels need to consider public image or social presence when using human-like machines. For those hospitality firms that serve more powerful customers in luxury hotels, such as wealthy people, celebrities, and executive corporate staff; these firms need to think of the location and size of the human-like machine to maximize VIP customer satisfaction.

Impact on Corporations and Communities

Increase Production and Reduce Cost

Most literature regarding this subject explores productivity in the context of industrial robotics. Researchers argue that the robotics technology significantly contributes to industrial productivity. The applications of robots to date have extended from large corporations to small businesses; further, AI and robotics technologies will impact small businesses in different areas positively as well, such as predicating demand and planning production (OTA, 1982; Noone & Coulter, 2012; Ivanov et al., 2017). The financial impact of robots in the hospitality firms cannot be ignored, Murphy et al.

(2017) state that industrial robots in the back of the house and service robots in the front of the house both have financial effects.

Cost will be another concern in the context of the application of robots. Many of the newest robots are less expensive than traditional industrial robots. For instance, industrial robots can cost from $100,000 to $500,000 while collaborative robots, which will be adopted in hotels, are priced from $20,000 to $50,000 each (Heires, 2016). A newly built Fairfield Inn & Suites in San Marcos deployed the service robot "Relay" to deliver food, towels, and other items to the guest rooms. Relay costs $2000 per month under a three-year contract with the robot company, Savioke (Hirsh, 2017). The article also mentioned that the application of robots can greatly reduce labor costs for the night shift in hotels and other hospitality segments. The robotics experts believe prices will drop when robots enter into bigger scale manufacturing. In addition, the development of modern batteries and electronics technologies will be lower the cost, advance performance, and broaden the availability of robotics and its related technologies.

The motivation of hotel operators to apply robotics technology is obvious. The trendy application of robots could easily improve their service to some extent, therefore, enhance their reputation and brand image. However, it is doubtful and unclear if hotels can get substantial financial value by implementing robotics technology and reducing labor costs. For example, recall that the Wynn Las Vegas is going to place an Amazon Echo in each their rooms (Raz, 2016). The price of an Amazon Echo is $180, and the Wynn Las Vegas has 2716 rooms for the Wynn location and 2034 rooms for Encore location. After doing a simple math, the total cost will be $855,000 for their initial installment. The Wynn might receive a discount from Amazon, but the future maintenance fee still cannot be underestimated. Yet, Heires (2016) also thinks the application of AI and robots is worth the cost when robots can collect data and map customers' purchasing behavior by directly interacting with them. In this regard, hotels will be able to offer personalization services and achieve higher customer satisfaction in the near future.

In conclusion, some types of the application of robotics science will increase costs while some will reduce costs. Those robots that are applied and on the right track will mostly prove to be valuable (Wells, 2003).

Education and Training

With the growth of robotics technologies, a number of education and training issues are raised by robotics. Some robotics experts also brought up the importance of improving the education system to help people fit into the evolution of the technology; they argued people should focus more on how to extend their knowledge and skills for the age of robots instead of being worried about being replaced by robots (Yakowicz, 2016). Murphy et al. (2017) argued that robotics technologies have impacts on successful tourism and hospitality education. A good start for the robotic world would be to learn about and more understand digital technologies.

Emerging Applications in Each Hospitality Segment

Hotel

Although AI and robots have already entered different hotel departments, the application of robotics is still nascent. The future potential use of robots for the hotel industry is predicted to be broad. It will include, but not be limited to, the laundry department: robots clean and fold sheets and towels; in-room voice-activated systems to help customers control the lighting, AC, and other devices in the room: such as the Amazon Echo mentioned earlier (Raz, 2016). To offer a seamless guest experience, hotel operators will keep chasing better technologies to develop an integrated mobile service ordering system for their guests. This system will eventually enable guest requests to be fulfilled instantly and in no time (Trejos, 2015). Hotels are going to become fully service automated through self-service devices and robotic technologies in the future.

Food and Beverage

For the future application of robotics technology in the food and beverage industry, Ivanov et al. (2017) stated that if AI can offer more advanced technologies, restaurants will also apply those service automation devices or technologies to pursue a fully automated business model similar to the Henn-na hotel, except a restaurant version. They also claim the must-have elements in the future for automated restaurants. These must-haves should contain both front of the house and back of the house automation: at the front of the house, customers can order food through self-service devices or with a robot waiter; at the back of the house, robot chefs will cook the

dishes and deliver the dishes through automation devices, such as conveyors in the sushi bar, or with a robot busser to the customers.

In the quick service restaurant segment, other scholars believe robotic technology or robotic management systems will provide small restaurants opportunities to compete with large chain restaurants, because of their scalability. In other words, although large chain restaurants will have the procurement power and brand equity as competitive advantages, in the scenario of applying affordable robotics management, small businesses will be productive and maintain high efficient daily operations. Robotic management would become the strength of small quick service restaurants (Noone & Coulter, 2012).

Meetings and Events

As mentioned previously, the current application of robots in the meetings industry is mainly to amuse attendees at the meetings. Base on this application, a future potential use of robotics science will combine with other technology, for example, telepresense (MantaroBot, n.d.). Telepresence is an innovative technology for the meeting industry, which allows people who are far away or not available to attend meetings and conventions through the MantaroBot. The key idea of this business model is to enable people to use robots as a physical presence indicator to attend meetings or events (Ivanov et al., 2017). MantaroBots will have screens, cameras, and be mobile like other service robots; in addition, they will have a video phone-call feature.

This technology will be very revolutionary for the meeting industry, and offer the opportunity for more people to attend meetings and conventions virtually without transportation costs, and to still have interaction with others even though they could not physically attend the meetings. A similar idea is to use virtual reality to attend meetings (Revvr Studios, 2017).

The Risks Posed by Developments in Robotics

After forecasting a potential use for robotics technology in the future, it is necessary to explore the possible factors that will influence the development. Kuo, Chen, and Tseng (2017) used expert panel research methods and concluded six factors that will influence the development of robots in the hospitality industry. These six factors are government support, capability for market development, future development of the robotics industry, capability for technology development, capability for raising money, and

capability for talent development. Some research also investigates other factors that could influence service robots. Gorle and Clive (2013) noted that a country with a high lifespan and low birth rate will more likely have an increasing demand for service robots, such as Japan.

Other studies investigate how to make the application of robotics technology successful in the hospitality industry. For example, strategic positions of the hotels in a market could impact the success of using robots (Melian-Gonzalez & Bulchand-Gidumal, 2016). Improving customer satisfaction and developing into new markets (Baker & Magnini, 2018), highlighted with the practical execution of the marketing campaigns, are also important elements to the success of advanced services with robots or AI (Nieves & Diaz-Meneses, 2016).

When hospitality companies enjoy the convenience and consistent service from robots, some of them also begin to be concerned about the risk and safety of adopting AI and robots, or possibly even "fire" robots due to certain defects (Mccrum, 2016). Some experts mention this concern is reasonable because of the self-learning skills and big data capabilities of the robots, especially those of the interactive robots (Heires, 2016). They also argued robots are originally safe, but if they mingle with other technologies or platforms that could modify or add other programs into the robots, the robots might not be safe anymore. So for now, the experts think the biggest risk related to robot industry is cybersecurity.

Conclusion

Artificially intelligence (AI) and robotics technology have already penetrated into different segments in the hospitality industry. As a service industry, the majority of robots used in the hospitality industry are service robots. This review mainly discusses the application of professional service robots in the hospitality industry. In hotels, robots are utilized in the front office, housekeeping, and the concierge department. In the food and beverage industry, robots could be servers, chefs, bartenders, and bussers. During meetings and events, robots aim to amuse people to promote products.

With the growth of the robotics industry, the impacts of AI and robots are more and more significant. The AI system can augment staff capability, thus, improve daily operations. Additionally, a robotics management system can eliminate internal staff theft. Although robots might replace a few human's jobs, they will still create more jobs in the labor market. AI and

robotics has a positive impact on labor and employment. The research in regard to how robots impact customers is mainly about how looks and interactions between robots and human influences customer behavior, especially with human-like robots. This academic area can help design future robots. AI and robots can have a financial impact on companies, increase their productivity, and reduce the labor costs.

The hotel industry wants to reap from these developments and aims to attain full automation. For meetings and events, robots will combine with other technologies in order to create a more innovative application. Telepresence, which can allow human beings to interactive with each other in a meeting without physical attendance, is a particularly interesting application.

Hotels or other hospitality segments invest a lot of money into AI and robots, to make it foolproof. The biggest risk of utilizing robots in the hospitality industry is cybersecurity, because most of the robots need to connect to the internet. To avoid the risk, they have to make sure that all of these new systems are safe and that hackers cannot recode them.

There is still not a lot of research into the application of robotics science or AI in the context of the hospitality industry. Future research may benefit by investigating the following areas: in what locations should robots be placed in a hotel, restaurant, or even at a meeting in order to influence customers' perceptions of robots; what characteristics (e.g. age, nationality, race) should robots have to improve the service experience? Robots have already entered our lives in various guises. Humans have a lot to do to optimize the application of robots and to make our lives even better.

References

Achacoso, T. B. (1990). Artificial ethology and computational neuroethology: A scientific discipline and its subset by sharpening and extending the definition of artificial intelligence. *Perspectives in Biology and Medicine, 33*(3), 379–390.

Baker, M. A., & Magnini, V. P. (2018). The evolution of services marketing, hospitality marketing and building the constituency model for hospitality marketing. *International Journal of Contemporary Hospitality Management, 28(8)*, 1510–1534.

Borràs, J., Moreno, A., & Valls, A. (2014). Intelligent tourism recommender systems: A survey. *Expert Systems with Applications, 41*(16), 7370–7389.

Chen, Y., & Hu, H. (2013). Internet of intelligent things and robot as a service. *Simulation Modelling Practice and Theory, 34*, 159–171.

Chijindu, I., & Inyiama, H. C. (2012). Social implications of robots—An overview. *International Journal of Physical Sciences, 7*(8), 1270–1275.

Ciupe, V., & Maniu, I. (2014). New trends in service robotics. *Mechanisms and Machine Science, 16*, 57–75.

Collins, G. (2013). Safeguarding Restaurants from point-of-sale fraud: An evaluation of a novel theft deterrent application using artificial intelligence. *Journal of Hotel Business Management, 2*(105). https://doi.org/10.4172/2169-0286.1000105

Constine, J. (2015, August 31). Eatsa, A futuristic restaurant where robot cubbies serve quinoa. Retrieved August 8, 2017, from https://techcrunch.com/2015/08/31/eatsa/

Curtis, S. (2016). Pizza Hut hires ROBOT waiters to take orders and process payments at its fast-food restaurants. *Mirror*. Retrieved August 8, 2017, from http://www.mirror.co.uk/tech/pizza-hut-hires-robot-waiters-8045172

Daily News. (2013). Drone delivers beers—not bombs—at South Africa music festival. Retrieved August 8, 2017, from http://www.nydailynews.com/lifestyle/eats/drone-drops-beers-not-bombs-south-africa-article-1.1422617

Doorn, J., Mende, M., Noble, S. M., Hulland, J., Ostrom, A., Grewal, D., & Petersen, J. A. (2017). Domo arigato Mr. Roboto: Emergence of automated social presence in organizational frontlines and customers' service experiences. *Journal of Service Research, 20*(1), 43–58.

Eyssel, F., & Hegel, F. (2012). (S)he's got the look: Gender stereotyping of robots. *Journal of Applied Social Psychology, 42*(9), 2213–2230.

Fan, A., Wu, L., & Mattila, S. (2016). Does anthropomorphism influence customers' switching intentions in the self-service technology failure context? *Journal of Services Marketing, 30*(7), 713–723.

Fridin, M., & Belokopytov, M. (2014). Acceptance of socially assistive humanoid robot by preschool and elementary school teachers. *Computers in Human Behavior, 33*, 23–31.

Gorle, P., & Clive, A. (2013). Positive impact of industrial robots on employment. Retrieved August 8, 2017, from www. ifr.org/uploads/media/Update_Study_Robot_creates_Jobs_2013.pdf

Heires, K. (2016, September 1). Rise of the robots. Retrieved August 8, 2017, from http://www.rmmagazine.com/2016/09/01/rise-of-the-robots/

Henn-na Hotel. (n.d.). Retrieved August 8, 2017, from http://www.h-n-h.jp/en/

Hilton. (2016). Hilton and IBM pilot "Connie," the world's first Watson-enabled hotel concierge. Retrieved August 8, 2017, from http://news.hiltonworldwide.com/index.cfm/news/hilton-and-ibm-pilot-connie-the-worlds-first-watsonenabled-hotel-concierge

Hirsh, L. (2017, May 13). Hotels make room for robots on staff: Bots could reduce labor costs. Retrieved from http://sdbj.com/news/2017/may/18/hotels-make-room-robots-staff/

International Federation of Robotics (IFR). (2016a). World robotics industrial robots. Executive summary. Retrieved August 8, 2017, from http://www.ifr.org/fileadmin/user_upload/downloads/World_Robotics/2016/Executive_Summary_WR_Industrial_Robots_2016.pdf

International Federation of Robotics (IFR). (2016b). World robotics service robots. Executive summary. Retrieved August 8, 2017, from https://ifr.org/img/office/Industrial_Robots_2016_Chapter_1_2.pdf

International Organization for Standardization. (2012). ISO 8373:2012(en) Robots and robotic devices—Vocabulary. Retrieved August 8, 2017, from https://www.iso.org/obp/ui/#iso:std:iso:8373:ed-2:v1:en:term:2.2

Ivanov, S., Webster, C., & Berezina, K. (2017). *Adoption of robots and service automation by tourism and hospitality companies*. Paper presented at the INVTUR Conference, 17–19 May 2017, Aveiro, Portugal.

Kaur, S. (2012). How medical robots are going to affect our lives. *IETE Technical Review, 29*(3), 184–187.

Kim, M., & Qu, H. (2014). Travelers' behavioral intention toward hotel self-service kiosks usage. *International Journal of Contemporary Hospitality Management, 26*(2), 225–245.

Kowitt, B. (2016, December 13). San Francisco's favorite quinoa spot is taking on New York. Retrieved August 8, 2017, from http://fortune.com/2016/12/13/eatsa-new-york-location-quinoa/

Kuo, C., Chen, L., & Tseng, C. (2017). Investigating an innovative service with hospitality robots. *International Journal of Contemporary Hospitality Management, 29*(5), 1305–1321. https://doi.org/10.1108/IJCHM-08-2015-0414

Louie, W.-Y. G., McCaoll, D., & Nejat, G. (2014). Acceptance and attitudes toward a human-like socially assistive robot by older adults. *Assistive Technology, 26*(3), 140–150.

Maidbot. (n.d.). Retrieved August 8, 2017, from http://www.maidbot.co

MantaroBo. (n.d.). Retrieved on August 8, 2017, from http://www.mantarobot.com

Markoff, J. (2014, August 11). "Beep," says the bellhop: Aloft hotel to begin testing 'botlr,' a robotic bellhop. Retrieved August 8, 2017, from https://www.nytimes.com/2014/08/12/technology/hotel-to-begin-testing-botlr-a-robotic-bellhop.html

Maurer, M., Gerdes, J. C., Lenz, B., & Winner, H. (Eds.). (2016). *Autonomous driving: Technical, legal and social aspects*. Berlin, Heidelberg: Springer Open.

Melian-Gonzalez, S., & Bulchand-Gidmumal, J. (2016). A model that connects information technology and hotel performance. *Tourism Management, 53*(C), 30–37.

Mccrum, K. (2016). Robot waiters fired after spilling drinks and food during restaurant service. Retrieved August 8, 2017, from http://www.mirror.co.uk/tech/robot-waiters-fired-after-spilling-7715093

Min, H. (2010). Artificial intelligence in supply chain management: Theory and applications. *International Journal of Logistics, 13*(1), 13–39.

Mirheydar, H. S., & Parsons, J. K. (2013). Diffusion of robotics into clinical practice in the United States: Process, patient safety, learning curves, and the public health. *World Journal of Urology, 31*(3), 455–461.

Mori, M., MacDorman, K., & Kageki, N. (2012). The uncanny valley. *Robotics & Automation Magazine, IEEE, 19*(2), 98–100.

Murauskiene, N. (2015, May 12). Artificial intelligence is here to manage ATMs. Retrieved August 8, 2017, from https://www.linkedin.com/pulse/artificial-intelligence-here-manage-atms-neringa-murauskiene

Murphy, J., Hofacker, C., & Gretzel, U. (2017). Dawning of the age of robots in hospitality and tourism: Challenges for teaching and research. *European Journal of Tourism Research, 15*, 104–111.

National Restaurant Association. (1999). Avoiding an inside job. *Bread & Butter Newsletter*.

Nieves, J. N., & Diaz-Meneses, G. (2016). Antecendents and outcomes of marketing innovations. *International Journal of Contemporary Hospitality Management, 28*(8), 1554–1576.

Noone, B., & Coulter, R. (2012). Applying modern robotics technologies to demand prediction and production management in the quik-service restaurant sector. *Cornell Hospitality Quarterly, 53*(2), 122–133.

Office of Technology Assessment. (1982). *Exploratory workshop on the social impacts of robotics: Summary and issues.* Washington, DC: Congress of the United States.

Ostrom, A. L., Parasuraman, A., Bowen, D. E., Patrício, L., & Christopher, A. V. (2015). Service research priorities in a rapidly changing context. *Journal of Service Research, 18*(2), 127–159.

Pacurar, A. (2017, January 20). Can artificial intelligence help the hotel industry? Retrieved August 8, 2017, from https://www.cpexecutive.com/post/can-artificial-intelligence-help-the-hotel-industry/

Pan, Y., Okada, H., Uchiyama, T., & Suzuki, K. (2015). On the reaction to robot's speech in a hotel public space. *International Journal of Society Robotics, 7*, 911–920.

Perez, S. (2014). NoWait, The app that lets you join a restaurant wait list from your phone, goes nationwide. Retrieved from https://techcrunch.com/2014/01/30/nowait-the-app-that-lets-you-join-a-restaurant-wait-list-from-your-phone-goes-nationwide/

Rajesh, M. (2015). Inside Japan's first robot-staffed hotel. *The Guardian*. Retrieved August 8, 2017, from https://www.theguardian.com/travel/2015/aug/14/japan-henn-na-hotel-staffed-by-Robots

Raz, N. (2016, December 15). Amazon echo a 'game-changer' for Wynn Las Vegas. *Las Vegas Review-Journal*. Retrieved August 08, 2017, from https://

www.reviewjournal.com/business/casinos-gaming/amazon-echo-a-game-changer-for-wynn-las-vegas/
Rust, R. T., & Huang, M.-H. (2014). The service revolution and the transformation of marketing science. *Marketing Science, 33*(2), 206–221.
Slade, G. (2007). *Made to break: Technology and obsolescence in America*. Boston, MA: Harvard University Press.
S. P. Events. (n.d.). Fairs and promotions. Retrieved August 8, 2017, from http://www.spevents.co.uk/our-events/event-selector/exhibition_stand_attractions/interactive_robot_hire.html
Stapleton, B, (2017, June 30). Go on a Bender at the new Tipsy Robot. Retrieved August 8, 2017, from https://vegas.eater.com/2017/6/30/15898204/tipsy-robot-now-open
Revvr Studios. (2017, February 17). Where are the virtual conventions? Retrieved August 8, 2017, from http://www.revvrstudios.com/where-are-the-virtual-conventions
Tan, N., Mohan, R. E., & Watanabe, A. (2016). Toward a framework for robot-inclusive environments. *Automation in Construction, 69*, 68–78.
Timms, M. J. (2016). Letting artificial intelligence in education out of the box: Educational cobots and smart classrooms. *International Journal of Artificial Intelligence in Education, 26*(2), 701–712.
Trejos, N. (2015). Marriott to hotel guests: We're app your service. *USA Today*. Retrieved August 8, 2017, from http://www.usatoday.com/story/travel/2015/05/13/marriott-hotels-mobile-requests-two-way-chat/27255025/
Veruggio, G., & Operto, F. (2008). Roboethics: Social and ethical implications of robotics. In *Springer Handbook of Robotics*, Siciliano, B., & Khatib, O. (Eds.). Heidelberg: Springer. 1499-1524.
Wells, P. N. T. (2003). Can medical robots reduce health care costs? *Industrial Robot: An International Journal, 30*(1). Retrieved from https://doi-org.ezproxy.library.unlv.edu/10.1108/ir.2003.04930aaa.002
World Robotics News. (2012, August 28). Professional service robots—Continued increase. Retrieved from https://ifr.org/ifr-pressreleases/news/professional-service-robots-continued-increase
Yakowicz, W., & Inc. (2016, October 17). Robots might take your job—but you'll get a better one. *Chicago Tribune*. Retrieved from http://ezproxy.library.unlv.edu/login?url=https://search.proquest.com/docview/1829440782?accountid=3611

Check for updates

CHAPTER 14

Risk Management in the Digital Era: The Case of Nigerian Banks

Tankiso Moloi and Oluwamayowa Olalekan Iredele

Introduction

Moloi (2018) argued that in the context of the public sector, more emphasis tends to be biased to processes and systems, discounting the role of people, which renders the enterprise risk management program unsuccessful. Patterson (2015) concurs with this point as he points out that the enterprise risk management program would only be effective if it is supported by people, processes and technologies. People, processes and systems are all recipes for a successful risk management program. This however requires that the organization has reliable data to base their analysis and decision on.

One of the main pitfalls of the risk management programs in organizations has been a lack of reliable information and data to base decisions on. Patterson (2015) points to two factors that result in unreliable data from the risk management perspective and that is risk data is often scattered across the organization and not shared across business unit, in essence, the culture of silos is an impediment. Further, Patterson points to the fact that

T. Moloi (✉) • O. O. Iredele
University of Johannesburg, Johannesburg, South Africa
e-mail: smoloi@uj.ac.za; oiredele@uj.ac.za

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_14

many risk management functions lack the tools they need to capture and use risk information more effectively. What this effectively means is that it could be expected that there could be opportunities that the organization fails to capture.

By utilizing technology, Ernst and Young (2014) assert that we could begin to see the reduction in missed opportunities and realized risks. This is due to the fact that technology facilitates "improved insight, efficiency and effectiveness in risk management activities, thus freeing up time for risk management resources to spend on risk identification, assessment and control of future risks". Knopjes (2017) supports this view by stating that organizations could take advantage of emerging technologies to integrate different systems, collect and analyze massive volumes of data from an unlimited number of sources across multiple locations. Through this process, organizations will achieve enhanced operational processes, improve timeliness of reporting and utilize data to drive preventative actions resulting in them being miles head of potential risks (Knopjes, 2017).

Due to the nature of their business, which includes being the repository of liquidity, the core payment mechanism and the principal source of finance to at least a large part of the economy, banks assume a variety of risks (Moloi, 2014; Liu, 2004). The traditional finance theory has limited these risks into major balance sheet risks, namely, liquidity risk, interest rate risk, capital risk and credit risk (Hempel, Coleman, & Simonson, 1990). Traditional finance theory had not extended itself to think about the manner in which these risks and many other facing banks should be governed. As can be noted, there had not been a realization that technology would rise and lead to the circumstance where banks are dominantly relying on it for almost all their activities.

Moloi (2014) argues that banks are important to the economy as they lubricate the financial system, serve as the point of connection between various sectors of the economy and reassure a significant level of specialization, competence, economies of scale, and settings conducive to execution of different economic policies of the government (Sanusi, 2011) and as such, there is a need to identify, understand and monitor the circumstances, which could result in the banking systems' vulnerability (Moloi, 2014). Akande (2016) and Tennant and Tracey (2014) concur with this and they point out that banks are involved in financial intermediation by sourcing for savings at a rate that will entice depositors and lend funds to creditors at an affordable rate to maximize profits in a competitive environment.

The manner in which these institutions function is even more critical in developing nations such as Nigeria as the country battles to attract foreign direct investments for infrastructure and other needs. The Nigerian banking sector has faced a tumultuous period in recent times which has led to the restructuring and consolidation exercise (Aransiola, 2013). The reform led to the reduction in the number of banks from 89 to 25.

It had been thought that the restructuring and consolidation exercise would lead to stability, however; this situation was short lived as the banks got affected by another financial crisis that resulted in the survival of only 24 banks (Akande, 2016). The challenges that lead to the collapse of these institutions were attributable to risk management. As of 2018, the Central Bank of Nigeria has indicated that there are only 22 commercial banks that are healthy and are standing in Nigeria (CBN, 2019).

As the institution battled the credibility question and lack of confidence, the CBN responded by instituting the CBN code of good governance in order to give guidance to the Nigerian banks on the minimum corporate governance requirements. The central bank indicated that this was developed with the primary purpose of promoting a transparent and efficient banking system that will engender the rule of law and encourage division of responsibilities in a professional and objective manner (CBN, 2014). Adeoye and Amupitan (2014) advise that the concept of good governance in banking industry would need to be stretched in order to empirically include total quality management, which has six performance areas, namely, capital adequacy, assets quality, management, earnings, liquidity and sensitivity risk. The degree of adherence to these parameters determines the quality rating of the banks.

Given the importance of banks in the economy, the recent challenges faced by the banking sector in Nigeria, we are of the view that the evolutions in computing and risk technology which has led to the developments in technologies that could be used to exploit big data, conduct complex analytics, these technical advancements offer the Nigerian banks and authorities such as the CBN with better abilities for enhancing risk management effectiveness, thereby reducing vulnerabilities and exploiting opportunities that could arise.

This study therefore investigated whether or not the Nigerian banking sector was exploiting technology in order to reduce missed opportunities and realized risks, whether technology has been exploited in order to integrate different systems, collect and analyze massive volumes of data from an unlimited number of sources across multiple locations. Essentially, our

research question was answered through the use of content analysis of integrated reports of selected Nigerian banks where the focus is on the manner and the way in which risks have been captured. We content-analyzed this section of the integrated report in order to determine whether Nigerian banks indicated in these reports have deployed technology or not and whether or not advanced technology would have been deployed in order to reduce missed opportunities and realized risks.

The remainder of the chapter is organized as follows: Section "A Brief Overview on Corporate Governance" provides a brief overview on corporate governance, Section "A Brief Overview on Recent Corporate Governance Challenges in Nigerian Banks" briefly provides an overview on recent corporate governance challenges in the Nigerian banking sector, Section "Code of Corporate Governance for Banks in Nigeria" discusses the manner in which the codes of good practices including the CBN code are guiding banking firms on the governance of risk and whether or not they do highlight the role of technology, Section "Risk Management in the Digital Age" discusses risk management in the era of digital infrastructure and Section "Empirical Analysis of the Current State of Risk Management of Banks in Nigeria" presents the findings relating to the use of technology and technological infrastructure in risk management.

A Brief Overview on Corporate Governance

The financial crisis around the world and the consequent collapse of major corporate institutions in both developed and developing economies has brought to the fore the issue of corporate governance. Today, corporate governance has attracted considerable attention of policy makers and academic researchers across the globe. The emphasis is on the need for the practice of good governance both at the public and private enterprises and this is due to the economic primacy of publicly quoted firms in most national economies. Corporate governance is increasingly understood among policy makers as a value enhancing strategy in a competitive environment and there is a growing consensus globally that corporate governance has a positive link to national growth and development (Akinkoye & Olasanmi, 2014).

According to Alushna (2017), codes of best practice represent the self-regulation of listed companies and provide a range of recommendations (Larcker & Tayan, 2011; Mallin, 2004; OECD, 2004, 2015; Tricker, 2012) addressing:

- Rights of shareholders, stressing the equal treatment of shareholders holding shares of the same class
- The procedures and rules shaping the functioning of the annual shareholder meeting (ASM) and measures empowering shareholders motivating them to active participation in ASM
- Responsibilities of executives who are accountable to shareholders and stakeholders
- The procedures and rules shaping the functioning of the board
- Transparency standards which describe the scope and content of information policy (company operation and strategy, financial situation, ownership structure, composition, structure and procedures of the board, company by-laws and regulations, executive compensation)

The Cadbury report was the first step where corporate governance code had required that companies follow the "comply or explain" principles, which means that listed firms had to comply with the code guidelines. Those companies that could not comply were expected to report the non-compliance and provide explanation for the non-compliance with the possible measures to be taken for improvement (Cadbury Report, 1992).

Today, pushing for higher governance standards has become a regular campaign with the participation of an increasing number of parties such as academics, media, regulatory authorities, corporations, institutional investors and international organizations shareholders right watchdog amongst others. Corporate failures have necessitated considerable interests on empirical research on the effectiveness of various corporate governance institutions and mechanisms. We note that particular attention has understandably been drawn to addressing and researching the underlying issues and factors that led to the financial and corporate crisis that characterized both the developed and developing economies. Many nations today have taken numerous initiatives such as the introduction of code of best practice, new listing/disclosure rule, mandatory training for board directors and enforcement of code of governance amongst other things; as measures to address and enhance the issue of corporate governance practice. International organizations such as the Basel Committee on Banking Supervision, International Monetary Fund (IMF) and Organisation of Economic Co-operation and Development (OECD) have become major proponents of governance (Moloi, 2009, 2014). In this regard, Akinkoye and Olasanmi (2014) point out that IMF, for instance, demands that governance improvement should be included in its debt relief program. These efforts to ensure good corporate governance practice had also been extended to the banking industry (Moloi, 2016a, b).

A Brief Overview on Recent Corporate Governance Challenges in Nigerian Banks

In the Nigerian banking context, corporate governance has recently received increased attention because of high-profile scandals involving abuse of corporate power, and in some cases, alleged criminal activities by those entrusted with governance. The abuse of corporate power and criminal activities has directly been attributable to weak corporate governance in banks. Recently, the Governor of the Central Bank of Nigeria, Godwin Emefiele went as far as indicating that that corporate governance practices in the Nigerian banking sector leaves much to be desired. His main concern in this regard appears to be the fact that these fraudulent incidents were capable of undermining financial stability by heightening vulnerability of financial institutions to external shocks (The Punch, 2017).

In the banking sector good corporate governance practices are regarded as important in reducing risk for investors, attracting investment capital and improving the performance of companies. The series of widely publicized cases of accounting improprieties recorded in the Nigerian banking industry in 2009, for instance cases involving Oceanic Bank, Intercontinental Bank, Union Bank, Afri Bank, Fin Bank and Spring Bank were related to the lack of vigilant oversight functions by the boards of directors. Accordingly, boards in these banks appear to relinquish governance to corporate managers who then pursued their own self-interests leading to the reported improprieties (Uwuigbe, 2011).

Due to the fact that banks have an important role in the intermediation process, any disturbance in this sector will likely affect the economy. Monnin and Jokipii (2010) posit that modern financial system contributes to economic development and the improvement in living standards by providing various services to the rest of the economy. These include clearing and settlement systems to facilitate trade, channeling financial resources between savers and borrowers and various products to deal with risk and uncertainty. The ability of the bank therefore to manage risk being a major threat to its going concern is crucial, not only for itself but for various stakeholders associated with its business.

Code of Corporate Governance for Banks in Nigeria

Following the conclusion of the consolidation program in 2005, a Code of Corporate Governance for Banks in Nigeria was issued to the banking industry. The code which became effective in April 2006 was designed to

enhance corporate governance practices within the banking industry in view of the fact that governance mechanisms in banks was notably weak and board members of financial institutions were unaware of their statutory and fiduciary responsibilities.

The update on the code of governance became necessary in view of the ambiguity in the code and its conflict with the provisions of Companies and Allied Matters Act (CAMA) 1990 as well as the need to align the 2005 code with contemporary developments and international best practices, hence the updated 2014 edition. With its promulgation, there was an expectation that the revised code would provide clear guidelines on all aspects of governance and that it would enhance Corporate Governance practices for banks in Nigeria.

The 2014 code covers seven key aspects of corporate governance as it affects the banking sector in Nigeria, namely, board and management, shareholders, right of other stakeholders, Disclosure in annual reports, risk management, ethics and professionalism and conflict of interest and finally, sanctions. The board has a special role in carrying out an oversight of functions and the effectiveness with which the boards of financial institutions discharge this responsibility is critical to the bank survival. The code requires that boards be free to drive their institutions forward, but exercise that freedom within a framework of transparency and effective accountability.

For the purpose of this study, we indicate that while other aspects of the code are relevant to the going concern of banks, risk management is central in view of the nature of the risks that banks face and the fact that they are in custody of other people's money (Moloi, 2014, 2016a). The CBN code makes it the responsibility of the board of a bank to ensure that policies, procedures and controls are put in place to manage the various types of risks with which it is faced. Given the multifaceted nature of risk, the complexities surrounding today's organizations, particularly banks, governing risks in an institution like a bank would be complex. The speed in which technology is emerging provides an opportunity to the boards to insist to managements of these complex organizations that technology that integrates different systems and collects and analyzes massive volumes of data from an unlimited number of sources across multiple locations is a step forward toward the enhanced operational processes, an improvement in timeliness of reporting and utilization of data for the purpose of driving preventative actions which would result in organizations being miles ahead of potential risks (Knopjes, 2017).

It is the purpose of this study to investigate whether or not the Nigerian banking sector is exploiting technology in order to reduce missed opportunities and realized risks, and further, how technology has been exploited in order to integrate different systems and collect and analyze massive volumes of data from an unlimited number of sources across multiple locations. The objectives of this study are achieved through content-analyzing integrated reports of selected Nigerian banks where the focus is on the manner and the way in which risks have been captured. We have content-analyzed this section of the integrated report in order to determine whether or not Nigerian banks indicate that they utilize technology in risk management and whether these reports mention advance risk management technology and have deployed the same in order to reduce missed opportunities and realized risks.

Risk Management in the Digital Age

The CBN Code of Corporate Governance for Banks and Discount houses provides that

> *[e]very bank shall have a risk management framework specifying the governance architecture, policies, procedures and processes for the identification, measurement, monitoring and control of the risks inherent in its operations. The Board is responsible for the bank's policies on risk oversight and management and shall satisfy itself that management has developed and implemented a sound system of risk management and internal control.*

We live in a rapidly changing world. There is no time in history when virtually every aspect of human life has been affected by the rapid change brought about through information technology (Harari, 2018). According to Marwala and Hurwitz (2017), the world has witnessed four phases of revolution in human history. These revolutions are described below:

- The first industrial revolution brought mechanical innovations with the development of steam engine which was key to the then industrial revolution;
- The second industrial revolution which started in the second half of the nineteenth century brought the oil-powered internal combustion engine and electrical communication. Major technological advances during this period included the telephone, light bulbs and phonograph (Marwala & Hurwitz, 2017);

- The third industrial revolution or digital revolution which came in the 1980s brought computerization, that is, mainframe computers, personal computers and the internet, and the information and communication technology (ICT) available today. This has been a period of advancement of technology from analog electronic and mechanical devices to the digital technology (Agrawal, Gans, & Goldfarb, 2018; Marwala & Hurwitz, 2017); and
- The fourth industrial revolution which has arrived at the beginning of the twenty-first century is the advent of cyber-physical systems representing new ways in which technology becomes embedded within societies, that is, business, government, civil society, and so on, and the human body; it is driven by the rapid convergence of advanced technologies across the biological, physical and digital worlds. It is marked by emerging technology breakthroughs in a number of fields, including robotics, artificial intelligence, biotechnology, and so on. All of these revolutions came with unique impacts on every aspect of human lives including business (Harari, 2018; Agrawal et al., 2018; Marwala & Hurwitz, 2017).

The banking sector is undergoing its own revolution as well with significant implications for risk management (Moloi, 2014; Moloi, 2016a). Risk management is very significant to the operations of any business entity due to serious consequences that the occurrence of risk portends (Moloi, 2016b). It implies that for a business organization to be rest assured of the achievement of its objectives besides survival and growth, risk management becomes imperative (Ayodele & Alabi, 2014). Risk is the exposure to loss arising from the variation between the expected and actual outcomes of investment activities (Nzotta, 2002; Owualla, 2000). Therefore, in a broad term, risk management can be related to a mechanism which embraces planning, organizing and controlling resources and operational activities of business for effective reduction or elimination of risk or the adverse effects of risks (Ayodele & Alabi, 2014). Management of risks is thus imperative for banks given the nature of their business.

According to Ayodele and Alabi (2014), Onyekwelu and Onyeka (2014) as well as Moloi (2014), some of the common risks that are managed by banks include:

- Credit (default) risk which is also known as default risk. It is related to a loss that may occur from the failure of repayment of a credit advance made by a bank to its customers. Credit risk is inherent to the business of lending funds. It may lead to losses when banks' customers experience deterioration in financial condition, making it impossible to recover principal and interest on loans, securities and other monetary claims outstanding. Management of this type of risk is the most fundamental task in banking operations. The objective of credit risk management therefore is to minimize the risk and maximize bank's risk adjusted rates of return by assuming and maintaining credit exposure within the acceptable parameters.
- Liquidity risk which refers to the ability of a bank to fund increases in assets and meet obligations as they come due, without incurring unacceptable losses. Liquidity risk therefore occurs when a bank is not able to meet the payment of commitment it has made. To manage liquidity risk, the banks must periodically examine the structure of fund sources and uses and implement measures needed to improve this structure.
- Market risk arises when financial institutions actively trade assets and liabilities (and derivatives) rather than holding them for longer-term investment, funding or hedging purposes. Market risk implies not merely the risk of loss but also the potential for gain from interest rate, equity return and foreign exchange risk in that as these risks increase or decrease, the overall risk of the financial institution is affected. Market risk could therefore be defined as the incremental risk incurred by a financial institution when interest rate, foreign exchange and equity return risks are combined with an active trading strategy, especially one that involves short trading horizons such as a day. Market risk can be in the form of interest rate risk, refinancing risk and reinvestment risk. It is therefore required that bank should conduct strict management and control of market risk based on the awareness that the possibility of substantial losses is inherent in the nature of market transactions.
- The Bank of International Settlement (BIS) (1997), views operational risk (inclusive of technology risk) as the risk of loss resulting from inadequate or failed internal processes, people and systems or from external events. The breakdown of internal controls and corporate governance leading to "error, fraud, performance failure, compromise on the interest of the bank resulting in financial loss". The

objective of operational risk management is to find out the extent of financial institutions operational risk exposure; understand what drives it; to allocate capital against it; and identify trends internally and externally that would help in predicting it.

- Strategic risk is viewed by Onyekwelu and Onyeka (2014) as the risk arising from the overall strategy of the bank that includes the quality of the strategic planning process and the implications of such strategy especially for risk appetite and track record of implementation. It explains risk associated with business targets as risk from bank's product, services and customers.

It is clear from the risks that are said to be managed by banks above that technology related risks have not received prominence. This is probably because of the manner in which we have traditionally viewed banks. Of importance here is that whatever package or transaction that banks offer today, technology has an important role. In developing countries such as Nigeria where risks have materialized and caused vulnerabilities in the banking system, we are of the view that exploiting technology in order to reduce missed opportunities and realized risks, and using this technology in order to integrate different systems and collect and analyze massive volumes of data from an unlimited number of sources across multiple locations for the benefit of stakeholders is important.

In highlighting the importance of adequate management of technological risk, the Centre of Excellence in Financial Services (COEFS, 2018) stated how in this era of the fourth industrial revolution, technological advancements like artificial intelligence means that human and digital systems can interact more profoundly than ever before. Applying this technology in financial services has the potential to reduce costs and improve efficiency, allow customers to transact seamlessly and in real time, and improve providers' understanding of customer behavior and needs, allowing for the personalization of financial services. On the other hand, the fast pace of technological innovation and adoption within financial services means that banks will increasingly face a common set of operational risks around new technology. If these risks are not properly managed by banks, they have the potential to create instability and undermine consumer trust in the banking system as instances of cybercrime, systems failure and compromised data increase in frequency.

Particularly, as the guardians of sensitive private and financial data, fraud and cybercrime is the greatest reputational risk facing banks in the

digital age. One obvious reason contributing to the intensity of this risk is that banks have a number of fronts they need to protect. The digitization of bank channels and underlying processes create multiple avenues through which virtual attacks can take place, notwithstanding the need to protect banks' physical premises and systems from intrusion. In addition to banks' internal systems and processes, the increasing use of outsourced service providers and the integration of banks' systems with external vendors adds another layer of vulnerability. It is difficult for banks to cover all of these fronts comprehensively (COEFS, 2018). All of these require banks to develop comprehensive cyber and data resilience strategies. Failure in this regards in the information revolution era could launch the banking sector into another round of crisis that may be difficult to recover from.

Previously, the digitization efforts of banks have been customer-centered (such as online marketing) and the operations that support those efforts (customer onboarding, customer servicing). Only recently have banks expanded their transformations into other parts of the organization, including the risk function. In this regard, it is clear that banks concede that there is value in deploying technology as part of managing risks.

A survey conducted by the Institute of International Finance/McKinsey shows that 70 percent of respondents reported that senior managers are paying moderate attention to risk-digitization efforts; 10 percent say that senior managers have made these efforts a top priority. Risk digitization is clearly an established topic in the executive suite (Institute of International Finance/McKinsey, 2018). The survey lays out seven key elements a bank must develop for a digital risk transformation, namely:

- Accordingly, enhanced data governance and operating models are likely to improve the quality of the data, make risk and business decisions more consistent, and ensure responsiveness to risk's data needs. The report is of the view that one important enhancement is the need to consider data risk as a key element of the risk taxonomy, linked to a specific risk-appetite statement and data-control framework and to accommodate far more varieties of data. The survey found that approximately 30 percent of the respondents say that new data sources will probably have a high impact on their work. This is in line with Agrawal et al.'s (2018) assertion that the power of artificial intelligence will allow resources to be focused somewhere else whereas technology will focus on predictions.

- On process and workflow automation, it is expected that as risk automates tasks such as collateral data entry, often through robotic process automation (RPA), it can combine several of them into smart workflows: an integrated sequence performed by groups of humans and machines across an entire journey (e.g., credit extension fulfillment). In addition to greater efficiency, smart workflows create a more seamless and timely experience for customers. The survey found that about a quarter of respondents believe that more than 15 percent of costs can be cut across different risk disciplines, except in credit, where the number is a bit above 60 percent. Around 30–45 percent of respondents see 5–15 percent cost-reduction potential from automation, depending on risk type. Ninety percent see benefits from increased precision and 55 percent believe automation will improve compliance with regulation. As a knock-on effect, risk resources will focus more on the value-adding activities they have been trained for. And 84 percent of respondents expect an increase in customer and employee satisfaction.
- On advanced analytics and decision automation, the survey is of the view that sophisticated risk models (e.g., those built on machine-learning algorithms) can find complex patterns (such as sets of transactions indicative of invoice fraud) and make more accurate predictions of default and other risk events. The survey found that nearly three-quarters of risk managers surveyed expect advanced analytics to have a significant impact on their work, while 50 percent say credit decision times will fall by 25–50 percent and a few respondents even believe that times could fall by 75–100 percent. This is in line with Agrawal et al.'s (2018) assertion that the power of artificial intelligence is on prediction, that is, it will make prediction accessible and faster.
- With regard to a cohesive, timely and flexible infrastructure, if risk infrastructure is digitized, it will evolve to support several other building blocks: innovative data-storage solutions, new interfaces, easier access to the vendor ecosystem, and so on. It will use techniques like application as a service, obtained from application service providers (even on open banking platforms). Approximately 45 percent of the respondents see innovative technologies as a high-impact building block. "No code" and "low code" solutions will put control further in the hands of risk executives and reduce the number of end-user computing tools. Nearly 60 percent of the respondents expect innovative data-storage structures to have a significant impact on risk management.

- On smart visualization and interfaces, risk digitization will deliver its insights in more intuitive, interactive and personalized ways through risk dashboards, augmented-reality platforms for customers, and other interfaces. The survey found that nearly 20 percent of risk managers expect nascent technologies, such as augmented reality, to have a high impact.
- On external ecosystem, risk digitization will facilitate easy partnerships with external providers to vastly improve customer on-boarding, credit underwriting, fraud detection, regulatory reporting and many other activities. The survey found that two-thirds of respondents see fintechs more as enablers than disruptors, while 63 percent of North American respondents plan to use industry utilities to deal with regulatory burdens.
- On talent and culture, risk digitization is expected to have a far greater share of digital-savvy personnel with fluency in the language of both risk and the business, operating within an agile culture that values innovation and experimentation. The survey reports that new profiles seen as most critical in a digitized risk function include data scientists and modeling experts. Accordingly, many risk leaders think that their teams will need to develop these skills rather than hire non-risk professionals and expect them to learn risk.

Empirical Analysis of the Current State of Risk Management of Banks in Nigeria

Given the recent challenges faced by the banking sector in Nigeria, we are of the view that the evolutions in computing and risk technology which has led to the developments in technologies could be used to exploit big data, conduct complex analytics and these technical advancements could offer the Nigerian banks and authorities such as the CBN better abilities for enhancing risk management effectiveness thereby reducing vulnerabilities and exploiting opportunities that could arise.

We set to investigate whether or not the Nigerian banking sector was exploiting technology in order to reduce missed opportunities and realized risks, whether technology has been exploited in order to integrate different systems and collect and analyze massive volumes of data from an unlimited number of sources across multiple locations. The results of content-analyzing integrated reports of selected Nigerian banks where the focus had been on the manner and the way in which risks have been cap-

tured and the determination whether Nigerian banks indicate in these reports whether they have deployed technology or not and whether or not the advanced technology would have been deployed in order to reduce missed opportunities and realized risks.

Table 14.1 presents the empirical investigation of 13 banks that rank within the top 40 firms in Nigeria in terms of market capitalization. The latest annual reports of the sampled banks were content-analyzed to extract information on the current risk management approach. Basically, four aspects of risk management approach were examined, namely, the specific risk management framework guiding the risk activities of the bank; the presence of risk committee whose responsibility is to provide oversight functions on behalf of the Board; basic technology employed to manage all forms of risks in the banks; and the application of advanced ICT for risk management in the information revolution era.

The first three aspects are fully manifest in the sampled banks. These are regarded as the traditional approach to managing risks in the banking industry and have existed for some time. Only two banks appear to be proactive in their risk management approach. Although the depth of risk digitization in these two banks are not verified, the latest annual reports issued by them have some elements of these advanced ICT technologies for managing risks.

Table 14.1 Current risk management approach in Nigerian banks

SN	*Banks*	*Risk management framework*	*Risk committee*	*Basic ICT for risk management*	*Advanced ICT for risk management*
1	GTBank	√	√	√	
2	Zenith	√	√	√	
3	Stanbic	√	√	√	
4	FBN	√	√	√	√
5	UBA	√	√	√	
6	ECOBANK	√	√	√	
7	Access	√	√	√	√
8	Union	√	√	√	
9	Fidelity	√	√	√	
10	Sterling	√	√	√	
11	FCMB	√	√	√	
12	Diamond	√	√	√	
13	Wema	√	√	√	
	Total	*13*	*13*	*13*	*2*

Conclusion, Recommendations and Policy Implications

It is clear in the integrated reports that banks have not embraced advanced technology for the purpose of managing risk. To reduce the likelihood of the reoccurrences of recent challenges faced by the banking sector in Nigeria, we are of the view that the evolutions in computing and risk technology which has led to the developments in technologies could be used to exploit big data, conduct complex analytics and these technical advancements could offer the Nigerian banks and authorities such as the CBN better abilities for enhancing risk management effectiveness thereby reducing vulnerabilities and exploiting opportunities that could arise.

With the rise of technology, majority of banking activities are already executed through technological instruments. As a policy recommendation, in addition to the traditional risk and audit committees that are clearly outlined in the code, the CBN would have to incorporate ICT committee in the codes of corporate governance for banks in Nigeria. Alternatively, the code should make it mandatory that some of the external members of the audit and risk committee should possess technological competencies.

References

Adeoye, A., & Amupitan, M. D. (2014). Corporate governance in the Nigerian banking sector: Issues and challenges. *European Journal of Accounting Auditing and Finance Research, 3*(5), 64–89.

Agrawal, A., Gans, J., & Goldfarb, A. (2018). *Prediction machines: The simple economics of artificial intelligence*. Boston, MA: Harvard Business Review Press.

Akande, O.B. (2016). *Corporate governance issues in the Nigerian banking industry*. PhD dissertation. Walden University.

Akinkoye, E. Y., & Olasanmi, O. O. (2014). Corporate governance practice and level of compliance among firms in Nigeria: Industry analysis. *Journal of Business and Retail Management Research (JBRMR), 9*(1), 13–25.

Alushna, M. (2017). *Corporate governance best practice: Tasks and shortcomings. Responsible corporate governance towards sustainable and effective governance structure*. Springer.

Aransiola, S. Y. (2013). The impact of consolidation on profitability of commercial banks in Nigeria. *Financial and Quantitative Analysis, 1*, 15–25. https://doi.org/10.12966/fqa.05.01.2013

Ayodele, T. O., & Alabi, R. O. (2014). Risk management in Nigeria banking industry. *Research Journal of Finance and Accounting*, *5*(2), 131–136.

Cadbury Report. (1992). *Report of the committee on the financial aspects of corporate governance*. Retrieved September 12, 2018, from http://www.ecgi.org/codes/documents/cadbury.pdf

Central Bank of Nigeria (CBN). (2014). *Code of corporate governance for banks and discount houses*. Abuja, Nigeria.

CBN. (2019). List of financial institutions. Retrieved from https://www.cbn.gov.ng/supervision/Inst-DM.asp.

COEFS. (2018). *The impact of the 4th industrial revolution on the South African financial services market*. Retrieved October 12, 2018, from https://www.genesis-analytics.com/uploads/downloads/COEFStheimpactofthefourthindustrialrevolutiononfinancialservicesinSouthAfrica-final-1-FR.pdf

Ernst & Young. (2014). Retrieved from https://www.ey.com/Publication/vwLUAssets/EY-leveragingtechnology-and-data-for-cost-effective-risk-management/$File/EYLeveraging-technology-and-data-for-cost-effective-risk-management.pdf.

Harari, Y. N. (2018). *21 lessons for the 21st century*. London: Jonathan Cape.

Hempel, G. H., Coleman, A. B., & Simonson, D. G. (1990). *Bank management: Text and cases* (3rd ed.). New York: John Wiley & Sons, Inc.

IIF/McKinsey. (2018). The future of risk management in the digital era. Digital risk survey of 35 banks, companies, and other institutions. Retrieved March 30, 2018, from https://www.mckinsey.com/~/media/McKinsey/Business%20Functions/Risk/Our%20Insights/The%20future%20of%20risk%20management%20in%20the%20digital%20era/Future-of-risk-management-in-the-digital-era-IIF-and-McKinsey.ashx

Knopjes, B. (2017). Why IoT is the future of integrated risk management. Retrieved October 12, 2018, from https://www.isometrix.com/iot-in-risk-management/

Larcker, D., & Tayan, B. (2011). *Corporate governance matters: A closer look at organizational choices and their consequences*. Upper Saddle River: Pearson Education.

Liu, V. (2004). Environmental considerations and business operations of commercial banks in China—A case study of the project loan appraisal policy of ICBC. Lund, Sweden: IIIEE.

Mallin, C. (2004). *Corporate governance*. Oxford: Oxford University Press.

Marwala, T., & Hurwitz, E. (2017). *Artificial intelligence and economic theory: Skynet in the market*. Springer.

Moloi, T. (2009). *Assessment of corporate governance reporting in the annual report of South African listed companies*. Masters dissertation, University of South Africa.

Moloi, T. (2014). Leading external and internal indicators of credit risk in the top South African banks. *Journal of Risk Governance and Control: Financial Markets and Institutions*, *4*(3), 51–65.

Moloi, T. (2016a). The nature of credit risk information disclosed in the risk and capital reports of the top-5 South African Banks. *Banks and Bank Systems, 11*(3), 84–90.

Moloi, T. (2016b). A cross sectoral comparison of risk management practices in South African organizations. *Problems and Perspectives in Management, 14*(3), 99–106.

Moloi, T. (2018). Analysing the human capital capabilities in the enterprise risk management function of South Africa's public institutions. *Business and Economic Horizon, 14*(2), 375–388.

Monnin, P. & Jokipii, T. (2010). *The impact of banking sector stability on the real economy*. Swiss National Bank. Working group paper.

Nzotta, M. S. (2002). *Corporate financial decisions*. Owerri: Oliverson Industrial Publishers.

OECD. (2004). *Principles of corporate governance*. Retrieved October 9, 2018, from http://www.Oecd.Org/Corporate/Ca/Corporategovernanceprinciples/31557724.Pdf

OECD. (2015). G20/OECD *principles of corporate governance*. Retrieved October 9, 2018, from http://www.oecd.org/daf/ca/Corporate-Governance-Principles-ENG.pdf

Onyekwelu, U. L., & Onyeka, V. N. (2014). Financial risk management: A review of the role of the central bank of Nigeria. *Research Journal of Finance and Accounting, 5*(5), 40–52.

Owualla, S. I. (2000). *Principles of financial management*. Lagos: G. Mag Investment Ltd.

Patterson, T. (2015). *The use of information technology in risk management*. IBM.

Sanusi, L.S. (2011). The impact of the global financial crisis on the Nigerian capital market and the reforms. *7th Annual Pearl Awards and Public Lectures*. Muson Centre, Onikan. Retrieved from http://www.cbn.gov.ng

Tennant, D. F., & Tracey, M. R. (2014). Financial intermediation and stock market volatility in a small bank-dominated economy. *Journal of Developing Areas, 48*, 73–95. https://doi.org/10.1353/jda.2014.0073

The Bank of International Settlement (BIS). (1997). Principles for sound management of operational risk. Retrieved from https://www.bis.org/publ/bcbs195.pdf.

The Punch (2017, September 17). *Tackling weak corporate governance in banks*. Retrieved October 9, 2018, from https://punchng.com/tackling-weak-corporate-governance-in-banks/

Tricker, B. (2012). *Corporate governance: Principles, policies and practices*. Oxford: Oxford University Press.

Uwuigbe, O. R. (2011). *Corporate governance and financial performance of banks: A study of listed banks in Nigeria*. Ota: Covenant University.

CHAPTER 15

Digital Technology to Enhance Project Leadership Practice: The Case of Civil Construction

John Ekechukwu and Thorsten Lammers

Introduction

In business, the last few decades have belonged to a certain type of character or worker. A manager who initiates, plans, executes, monitors and controls. These people may be seen as reactive "firefighters", who put out flames, focus on completing tasks and respond to unfolding events (Barber & Warn, 2005). However, the future belongs to a very different kind of person, with a different kind of mind. Leaders, who are "creators, empathizers, pattern recognizers, and meaning makers" (Pink, 2006). These people are proactive "firelighters", who create change, envision the future and evoke passion (Barber & Warn, 2005). And as the complexities of organizational projects become more volatile, the demands for these characters heighten. As such management now also implies leadership.

J. Ekechukwu (✉)
Turnbull Engineering, Sydney, NSW, Australia
e-mail: John.Ekechukwu@alumni.uts.edu.au

T. Lammers
University of Technology Sydney, Sydney, NSW, Australia
e-mail: thorsten.lammers@uts.edu.au

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_15

Leadership, however, is conventionally associated with organizations, not projects (Cavaleri & Reed, 2008). Its theories and respective characteristics are rooted from the context of permanent organizations (Tyssen, Wald, & Spieth, 2014). These lasting forms of organizing facilitate a sense of autonomy and unbounded freedom that may enable leaders to imagine alternative visions for the future (Cavaleri & Reed, 2008). Yet, projects do not have the same magnitude of autonomy. A project environment of limited and predefined duration creates a sense of short-term orientation and hampers the development of deeper social relations such as trust and camaraderie (Tyssen et al., 2014). Additionally, a project evokes focus toward the management of daily objectives and immediate achievements, thus inhibiting actions of leadership and long-term vision (Toor & Ofori, 2008). Each project is unique and conditions for team selection and motivation are often far from ideal (Anantatmula, 2010). At times, project members are engaged in more than one project, causing additional burdens on focus and commitment (Anantatmula, 2010). Because of this, role ambiguity becomes an issue as well as the uncertainty of allocated control (Zhang & Fan, 2013).

Therefore, the role of a "project leader" is significantly different to that of an "organizational leader". Leading an organization allows for freedom, whereas leading a project, involves fulfilling a "promise" to a customer, ultimately constricting a leader to operate within the confines of such promise (Cavaleri & Reed, 2008). In this sense, a project leader is one who practices leadership within predetermined parameters. This makes the role of a project leader a rare commodity that is valued now more than ever before. So far, literature has focused on endorsing particular leadership behaviors and styles, while ignoring the difficulties faced by professionals in practicing these behaviors.

This study will explore digital transformation as a mechanism to facilitate leadership development using the case of the Australian civil construction industry. The focus is to shift the emphasis of digital advancements toward the social phenomenon of leadership. Digital technology alters the way people communicate, interact and relate to one another (Nagy & Koles, 2014). To neglect its potential to enhance project leadership development may be a substantial oversight in the evolution of leadership practice. To date, there have been few attempts to explore digital technology as project leadership technology, particularly within an industry such as construction, in dire need of support to elevate project managers into project leaders.

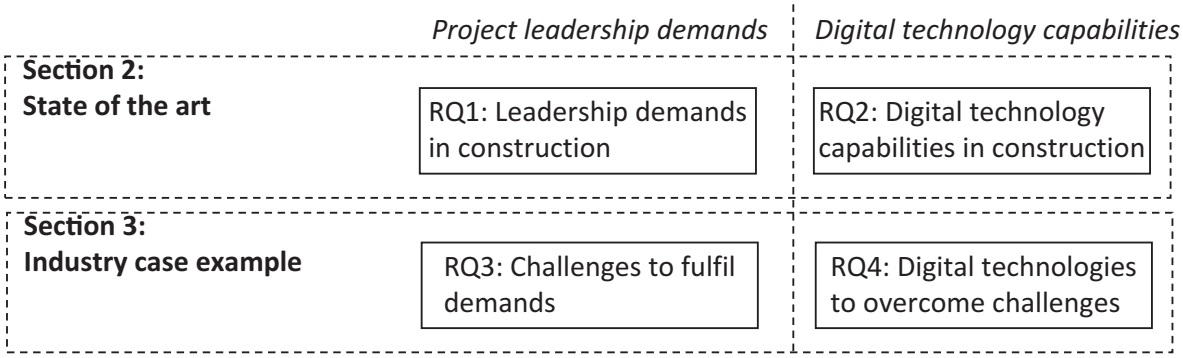


Fig. 15.1 Research design overview

This chapter aims to explore the potential for digital technology to enhance project leadership practice in construction projects. To investigate this potential, we will shed light on the following questions:

1. What are current leadership demands in construction?
2. What are current digital technological capabilities relevant to construction?
3. What challenges are construction professionals facing in meeting the leadership demands?
4. What is the feasibility of digital technology assisting professionals in overcoming these challenges?

The methodology employed to answer these questions involves a literature review coupled with a case example from the Australian civil construction industry (see Fig. 15.1).

Leadership Demands and Technological Capabilities in Construction

In the following section, we provide an academic assessment of the existing body of knowledge regarding leadership demands in the construction industry as well as capabilities of new digital technologies relevant for the construction sector. The focus is to discover important insights that can be used to guide and facilitate the investigation into the industry case example. In reviewing existing literature on those topics, the criteria listed in Table 15.1 have been established to ensure that relevancy and validity is maintained throughout the study.

Table 15.1 Overview of literature review

Source	The source is attainable in digital form.
Content	Content principally focuses on leadership and/or technology.
Keywords	Search is conducted using a combination of relevant words that depict leadership/technology (including plural or connotative equivalents).
Context	Sources are published within an international journal in the domain of construction, engineering, project management and the built environment.
Reliability	The source is academic/scholarly.
Relevancy	The source is published in English language within the date range specified (RQ = Research Question): • RQ1: sources published within the last ten years (from 2007 onward) • RQ2: sources published within the last three years (from 2014 onward)

Leadership Demands in Construction

Leadership challenges are amplified within the construction sector, an industry that will spend US$97.7 trillion globally in the next decade and account for 13.2% of world GDP by 2020 (Richardson, 2017). Construction projects are often large, multi-disciplinary and time-limited undertakings that exacerbate coordination and communication, as well as impose a great degree of uniqueness in setup, activities and content (Tyssen et al., 2014). Yet without tangible support, education and/or training, project managers are expected to "select, equip, train and influence" followers with diverse gifts, abilities and skills while operating within predefined parameters of time, budget and scope (Ameh & Odusami, 2014). The industry is suffering from a leadership crisis. Not from the professionals themselves, but from the disparity between the demands in project leadership and the investment and research in leadership development and training (Toor & Ofori, 2008). There is too much focus on the outcomes of leadership in neglect of the means in delivering these outcomes. Therefore, project managers need assisting tools and mechanisms that can balance the playing field and enable them to maximize their project leadership potential. As Khan (2015) claims, 80% of project failures are a result of poor project leadership or lack of project leadership skills. Management and technical competency alone are no longer sufficient for project success (Schuhmann, 2010). The following results highlight the variety of leadership demands relevant to the construction industry.

A US case study conducted by Cavaleri and Reed (2008) utilizes Seivert's five elements model to present a leadership strategy for improving project performance. The model constitutes of "Essence"

(defining purpose and values), "Air" (planning, analyzing and envisioning new possibilities), "Fire" (motivation and passion), "Water" (team building, loyalty, collegiality and camaraderie) and "Earth" (structure and policies). The expectation is that the project leader must be sensitive to the needs facing the team at the time and find a balance between the five elements. In teams in which the leader places emphasis on "Air" and "Fire" in neglect of "Water" and "Earth", projects are seen to fall behind, because important details that are necessary for execution are ignored. In a similar vein, when a project leader focuses on "Fire", "Water" and "Earth", while overlooking "Essence" and "Air", teams are likely to veer off in a direction that is misaligned with the project mission and objective.

Similarly, a study conducted by Patterson (2010) uses a case study of the Oresund Bridge construction (linking the cities of Copenhagen, Denmark and Sweden) to present a five-E model of leadership. "Envision", "Engage", "Energize", "Enable" and "Execute". By envisioning, "leaders see what the future could be rather than what it is". By engaging, leaders build relationships and seek collaboration. By energizing, leaders inspire and motivate. By enabling, leaders provide the necessary resources to facilitate followers in completing objectives. And by executing, leaders combine all elements together successfully.

These elements are further emulated in a study conducted by Slattery and Sumner (2011), who utilize the Leadership Practice Inventory (LPI) developed by Kouzes and Posner (2003) to examine the most favorable leadership characteristics among a group of 151 high potential construction project managers. The LPI purports to measure the ability of an individual to model the way (clarify values and set the example), inspire a shared vision (envision possibilities), challenge the process (take initiative, innovate and improve), enable others to act (foster collaboration, trust and facilitate relationships) and encourage the heart (provide emotional support).

Furthermore, surveys and structured personal interviews of 69 engineering professionals conducted by Anantatmula (2010) were used to develop a project management leadership model that consists of significant people-related leadership factors that influence project performance. These factors include: Create clarity in communication, define roles and responsibilities, communicate expectations, employ consistent processes, establish trust, facilitate support and manage outcomes.

The demand for authenticity and trust is accentuated throughout literature. Toor and Ofori (2008) conduct a literature review based in Singapore to rationalize the need for Authentic Project Leadership Development (APLD) for construction project leaders. The study explores four components of authenticity; "Awareness" (trust in one's motives and cognitions), "Unbiased Processing" (refraining from denying or distorting private knowledge), "Behavior" (acting in accord with one's values) and "Relational Orientation" (openness and truthfulness in relationships). Findings from this study argue that authentic leaders must know themselves, define their values, understand their motivations, build their support team, and "stay grounded by integrating all aspects of their life". Authenticity demands positive energy, integrity, moral character, self-discipline, optimism, resilience, clear purpose, concern for others and personal values.

As articulated by Kappagomtula (2017), without trust, leadership cannot function. Using investigative surveys of 36 experts in leadership across China and India, Kappagomtula models an analogy of a leadership "wheel" which must be maintained within cross-cultural and multicultural project teams. The model stipulates that the hub of the steering wheel is "Trust". The spokes of the wheel are "Loyalty", "Communication", "Integrity", "Competency" and "Consistency". The wheel itself represents "Culture", and the lubricant for the wheel is "Conflict Management". The emphasis of this analogy is the need for the leader to understand its components (the hub and spokes), maintain its integrity (lubricant) and steer the project team toward the destination of the drive. However, the leader must also learn when to let go of the "wheel" and surrender control.

This notion is also explored by Clarke (2012) who develops a theoretical understanding of shared leadership within projects. He argues that no individual performs all leadership functions, but rather members of the team collectively contribute to the leadership system. Therefore, a leader must focus on enriching established connections between team members and developing new connections to create a cohesive system in which leadership is distributed across all members.

Multiple studies have identified the need for project leaders to embody different leadership styles based on the varying stages of construction projects. Arain (2012) uses pilot questionnaires and interviews of 78 project management professionals working on the Suvarnabhumi

Airport (the largest construction project in Thai construction history) to develop a Quadrilateral Model of project leadership. The model focuses on a leader's ability to balance between four key leadership styles: "Coaching" (aligning goals and facilitating interaction among a team), "Connecting" (listening, empathizing, engaging and revering), "Controlling" (implementing rules) and "Charismatic" (supporting, inspiring and selling a vision). The research posits that leaders need variations in their leadership styles, at different stages of the project, until a balance between personal values, concerns for tasks and concern for people is maintained.

This concept is also explored in interviews across three Sri-Lankan construction projects conducted by Senaratne and Samaraweera (2015). This empirical study highlights the four stages of team development within projects: "Forming", "Storming", "Norming" and "Performing". The authors posit that a construction project manager should employ different leadership roles based on the needs and behaviors of the team members of each stage. For example, a leader should break down interpersonal barriers and establish team goals in the forming stages, but should only intervene where necessary in the performing stage.

This demand of leadership style adaptability is further reinforced by Oke's (2010) Multifactor Leadership Questionnaire (MLQ) of 57 Nigerian construction professionals. Findings from this study demonstrate that construction project leaders should possess the ability to adapt different leadership styles depending on the situation at hand. Leaders should embody a relations-oriented style at the beginning of a project, while adopting a much more task-oriented style during the closing of the project.

Leadership demands have also focused on the need for positive psychology and emotional intelligence among construction professionals. Emil Berg and Terje Karlsen (2014) conduct in-depth interviews with three experienced project managers to accentuate the need to encourage and develop positive emotions among followers. The study reveals how "Positive Meaning" (creating something), "Positive Emotions" (humor and optimism), "Positive Relations" (collaboration) and "Signature Strengths" (emotional intelligence) can create positive results within a project. These emotional elements play a major role in the development and trajectory of relationships within project settings (Pryke, 2015).

In a study, conducted by Pryke (2015), which uses surveys and live observations of 68 construction professionals, leaders are described as "managers of group emotions". Findings in this study demonstrate a clear correlation between emotional intelligence and project performance and further suggest that "Emotional Sensitivity" (ability to sense and interpret the non-verbal messages of others) and "Emotional Expressiveness" (ability to express emotions) are fundamental ingredients for "leader-follower chemistry". The expectation is that a leader should create a harmonious, smooth and in-tune connection between members of the team through emotional exchange and interaction.

Zhang and Fan (2013) further lament on emotional intelligence as a leadership requirement in their questionnaire-based survey covering 112 Chinese project managers. They find that characteristics of "Self-awareness", "Self-confidence", "Self-control", "Empathy", "Positivity" and "Cultural understanding" are key ingredients for project performance.

This is again emphasized by Galvin, Gibbs, Sullivan, and Williams (2014), who survey 38 project managers to define leadership competencies of "Emotional Resilience", "Intuitiveness", "Interpersonal Sensitivity", "Motivation" and "Conscientiousness". The study further postulates that an effective project leader must possess varying levels of sensitivity to political, technical and emotional elements of a project.

This is supported by Müller and Turner (2010), who use a web-based "Leadership Dimensions Questionnaire" on 400 project managers around the world to emphasize the need for project leaders to not only develop "Emotional Intelligence", but to balance between "Emotional", "Managerial" and "Intellectual Competencies".

Researchers have even attempted to appropriate religious teachings into necessary spiritual demands of leadership. Yngvason (2013) reviews the four gospels of the Bible to propose how project leaders can enhance their own practices by embracing the teachings of Jesus. The study proposes that Jesus utilized leadership tactics that enabled him to create a sense of purpose for a small number of followers that eventuated into a way of life for billions. The authors assume that like a project, the initial goal of spreading "the good news" was a planned endeavor, in which Jesus demonstrated transformational characteristics that included; "Individualized Consideration" (attentiveness to followers needs), "Intellectual Stimulation", "Inspirational Motivation" and "Idealized Influence" (a role model for ethical behavior).

Similarly, in a case study of eight Hungarian construction companies conducted by Lazányi and Dóka (2015), it is suggested that teachings from the Old Testament provide clues on leadership behavior. Leaders should be worthy, righteous men, who refrain from covetousness.

This is further explored from an Islamic perspective by Senam, Rashid, Sarkawi, and Zaini (2014), who review the Quran to highlight the need for authentic, servant and ethical leadership practices among construction professionals. The study unearths spiritual and moral values of leadership and emphasizes the need for leaders to carry out leadership activities with minimal waste or destruction and the upmost of ethical integrity.

Table 15.2 summarizes the demands identified in the above empirical studies of leadership relevant to the construction industry and groups them into three focus groups.

Digital Technology Capabilities Relevant to Construction

Digital transformation is the "application of digital technologies to fundamentally impact all aspects of business and society" (Gruman, 2016). The digital component involves technology, while the transformation component involves people (Del Rowe, 2017). It can enable individuals and organizations to transcend from one way of working to another.

This section focuses on reviewing the technological opportunities relevant to the construction industry. It provides a non-exhaustive list of technology examples, current and future trends of digital technology and their capabilities and application potentials for the construction industry.

Smartphones

Smartphones serve as the hubs of interconnectivity (Ahmed & Ahmad, 2014). They allow for knowledge sharing and exchange, as well as collaboration through a plethora of functionalities and applications. Smartphones can be used as a single mode of analysis for digital technology such as augmented reality (AR), to aid learning, understanding and remote collaboration (Joan, 2015); computer-aided design (CAD), to clarify important design details (Moreno, 2014); near-field communication (NFC), to track and monitor construction worker health and safety

Table 15.2 Overview of key leadership demands in construction

Leadership demand	*Source*						
	Cavaleri and Reed (2008)	*Patterson (2010)*	*Slattery and Sumner (2011)*	*Anantatmula (2010)*	*Toor and Ofori (2008)*	*Kappagomtula (2017)*	*Clarke (2012)*
Team focus							
Empower relationships	•	•	•		•		•
Enable effect. communicat.				•		•	
Establish trust & loyality			•	•	•	•	
Provide emotional support			•				
Align team's goals							•
Leader focus							
Envision new possibilities	•	•	•				
Give control to team							•
Lead by example	•	•	•			•	
Manage conflicts						•	
Balance all demands	•	•				•	
Project focus							
Define work struct. & policy				•			
Define purpose and values	•		•		•		
Communicate expectations				•			
Provide necessary resources		•		•			
Adapt style to project phase							

Arain (2012)	*Senaratne and Samaraweera (2015)*	*Oke (2010)*	*Emil Berg and Terje Karlsen (2014)*	*Pryke (2015)*	*Zhang and Fan (2013)*	*Galvin et al. (2014)*	*Müller and Turner (2010)*	*Yngvason (2013)*	*Lazányi and Dóka (2015)*	*Senam et al. (2014)*
	•	•	•	•						
•										
					•					
•			•	•	•		•	•		
•	•									
•			•					•		
	•									
•			•	•	•			•	•	•
•							•			
•		•								
•										
•		•								

(Akhavian & Behzadan, 2016); and credential verification services (CVS), to monitor employee training records and renewals. Smartphones are an extension of their users, and with the many applications that can be integrated within these devices, they act as a remote for digital transformation (Azhar, Jackson, & Sattineni, 2015).

Wearable Technology

Wearable technology (or "wearables") describes the integration of smart electronics within items that may be worn or implanted (Wright & Keith, 2014). Smart glasses, for example, incorporate an optical display that can advance information exchange at construction sites (Moon & Seo, 2015). In a split second, electronic documents, video and images can be shared, saved and retrieved (Heembrock, 2015), allowing construction professionals to be freed from the burden of holding paper documents, such as drawings and specifications (Moon & Seo, 2015). Wearable wristbands that measure sleep patterns and circadian rhythm can aid in managing unsafe working conditions by identifying fatigue, illness and other physical risk factors (Cullen, 2016). These functions can also be extended to mental health and well-being, by reminding construction professionals to stretch, hydrate and replenish (Heembrock, 2015). With the integration of woven sensors, wearable clothing can inform workers if they are exceeding lifting thresholds, have been overexposed to UV rays or are at risk of a slip and fall incident (Heembrock, 2015). Self-charging shoes that detect temperature and track location and motion can also ensure worker well-being is maintained when working alone (Linder, 2017). With the variety of sensors that can be integrated within clothing or accessories, leaders can understand the well-being of their team and advance their role in improving health and safety within construction projects (Linder, 2017).

Drones

Drones function as aerial imaging vehicles that can transcend data collection capabilities within construction projects. They allow for site progression and operations to be accurately tracked and reported (Dillow, 2016) and can enhance efficiency with site surveys, photography and mapping. With sensory features integrated, drones can turn data into 3D structural models, topographical maps and volumetric measurements (Dillow, 2016),

which can translate to enhanced productivity, safety, quality and security within projects. Additionally, "smart drones" possess the capability of recognizing objects and can incorporate a "sense-and-avoid-system" to self-monitor project operations (Hambling, 2015). These aerial capabilities add a greater dimension to project tracking and can allow project leaders to effectively communicate, mitigate potential issues, reduce costs and limit delays (Dillow, 2016).

Virtual Reality, Augmented Reality and Mixed Reality

Advanced visualization technologies can impact upon the way groups and individuals collaborate, communicate and engage (Phelps, 2014). Augmented reality (AR) overlays the visible natural world with a layer of digital content, while mixed reality (MR) integrates virtual objects into the natural world (Heiskanen, 2016). These capabilities for digital content to interact with the real world enable new possibilities for safety, planning, quality and management (Heiskanen, 2016). AR used through smart glasses can enable for a digitally enhanced view of the world that can facilitate the learning and awareness for student and professional development in construction (Kivrak, Arslan, & Tuncan, 2015; Ahmed & Ahmad, 2014). For example, the Smart Helmet by American augmented reality company DAQRI creates an intuitive experience as it connects users to the work environment and provides relevant information instantaneously (Heiskanen, 2016). Additionally, holographic technology such as Microsoft Hololens can allow remote teams to experience the spatial presence of the design model, collaborate and share information instantaneously over three-dimensional holograms (Heiskanen, 2016).

Virtual reality (VR) however creates a completely simulated environment (Sampaio & Viana, 2014) and can allow for four-dimensional viewing and interaction with all stages of the construction process. Project team members can therefore optimize planning and design and create an interactive environment for construction resourcing and scheduling VR can also be used to support learning and awareness for safety and risk management and develop a platform for users with limited knowledge to easily simulate the activities needed (Kassem, Benomran, & Teizer, 2017). These visual capabilities can aid in prediction and decision making by allowing users to view, interact and simulate their projects throughout its life cycle (Sampaio & Viana, 2014).

Building Information Modeling (BIM)

BIM is a three-dimensional (3D) model-based technology process for creating and managing project information (Succar, 2009). BIM will reduce the "blind spots" of two-dimensional (2D) construction drawings (Lee et al., 2015) and facilitate for extensive stakeholder collaboration (Kim et al., 2016). It can improve resource coordination, quality control, information exchange, safety, security, decision making and reduce manual tasks (Park, 2016). However, its collaborative platform (Lee et al., 2014) is what can facilitate project leaders to effectively communicate in real-time among project stakeholders (Park, 2016). Furthermore, new advancements of four-dimensional (4D) BIM and five-dimensional (5D) BIM enable further improvements in areas such as design, information management and sustainability. 4D BIM incorporates time to establish links between project activities and scheduling, which enables construction professionals to conduct constructability checks before construction (Kim, 2016). Whereas 5D BIM incorporates costing within the model to ensure the effective management of costs throughout a project life cycle, including cost estimation and budgeting (Kim, 2016).

Artificial Intelligence and Software

Artificial intelligence (AI) can redefine how processes are undertaken within the construction industry (Blanco, Fuchs, Parsons, & Ribeirinho, 2018). Collisions and inefficiencies within models such as BIM can be pre-empted, tracked and controlled (Garcia De Soto & Adey, 2016). Knowledge contained in daily reports, schedules, weather forecasts, and many more can be transformed into a living knowledge base and means of deductive and plausible conclusions (Klashanov, 2016). Garcia De Soto and Adey (2016) also believe that by integrating the case-based reasoning and neural networks of AI, resource estimates and decision support can be drastically enhanced.

Emotion aware technologies possess mood tracing capabilities with the potential to track and analyze an individual's emotional state (Morsy, 2016). This can enable construction professionals to keep mood journals and emotional data to understand what influences their emotion (Morsy, 2016). Additionally, authentication technology such as radio-frequency identification (RFID); quick-response codes; and fingerprint, vein, iris and facial recognition can allow for manpower management, credential verification and access control (Chin, Kim, & Choi, 2017).

Auto-Analytics

Auto-analytics and cognitive computing can be used to radically improve emotional intelligence and social relationships within the future construction workforce (Hansen, 2017). It is the process of intelligent logical analysis through technology, in which data is collected autonomously and used for analysis (RGA, 2013). Data can be collected through wearables such as watches, headbands and rings that can detect physiological data, such as heart rate, skin temperature and brain waves to accurately read a person's mood and provide real-time data to the user (ABI Research, 2017)

The Case of Australian Civil Construction

Construction has become one of the most important industries in Australia (Leviäkangas, 2017), and by 2020, Australia will become one of seven key countries that will account for 65% of the growth in global construction (Richardson, 2017). Using the case of a Sydney-based contract civil engineering firm specialized in design and asset management of transport infrastructure, land development and renewable energy projects, exploratory interviews were conducted to first unearth challenges faced by construction professionals in meeting leadership demands and expectations and then to explore the feasibility of existing digital technologies to assist in overcoming these challenges. The interviewees were a senior project manager with 16 years' experience, a project manager with 5 years' experience and a junior project manager with 2 years' experience.

Challenges in Addressing Leadership Demands in Civil Construction Projects

Interviewees expressed the significance of project leadership and its importance for the well-being of construction projects. They considered leadership as a critical component of their role and when asked whether they perceive themselves as a leader, one interviewee proclaimed "only because I have to be". Leading was considered as a way of being "consistent" and not allowing life experiences to interfere with project team experiences. The more junior respondents emphasized their thoughts on how leadership "comes with experience". Yet, the senior respondent did not entirely feel he had gained project leadership mastery or even gained a superior grasp of project leadership behavior. Rather, this respondent felt it was an ongoing challenge of his role.

The following sections outline the key challenges extracted from the interviews. They were clustered around the previously identified focus areas of leadership demands "team focus", "leader focus" and "project focus" (see Table 15.2).

Challenges Relating to the Team Focus

Respondents emphasized the necessity of cohesion and its importance to project success, highlighting that a difficult project with a "good team" can be successful, yet an easy project with a "poor team" may perform terribly. Interviewees felt that the difficulty in maintaining cohesion stemmed from the constant "chopping and changing" of project team personnel, which made it difficult to align team goals. Respondents also felt that leadership elements of team chemistry and cohesion were dependent upon "the people". As described, there are "different people" and "no one's good, no one's bad". It's all about "who they are, where they came from, and what their strengths and weaknesses are". This creates difficulties in dealing with the different types of personas, and aligning the way they operate with each other, with the best interests of the project. However, difficulties associated with cohesion were often related to difficulties associated with lack of control. With one respondent suggesting that "you don't necessarily build the team, you get given the team".

Respondents felt that the key challenges associated with communication arose from the need to communicate differently, with different types of people, and in different types of situations. As one respondent proclaimed; "it's not always rainbows and unicorns", at times you can "hurt people's feelings". This was further reinforced as interviewees lamented challenges with communicating with team members in the ways they expect to (or would like to) receive information. Getting the desired outcome from communicating was also seen as a pressing difficulty, with one respondent pointing out the difference in communicating, and "communicating effectively". "It's easy to tell someone to dig a hole, but it's more difficult to convince someone why they should be digging a hole". The challenge in communication is convincing people the reasoning behind what they are doing and "how it impacts the bigger picture".

Challenges Relating to Leader Focus

Respondents felt a key challenge in leadership awareness was in understanding people and how they operate. They believed that understanding which leadership style to adopt "takes up so much of your time" and if you get it wrong, "it can come back and bite you". Another interviewee raised the issue of proactivity, and the difficulty in being proactive in an environment that often imposes reactivity. Getting appropriate feedback was also deemed challenging as respondents felt they could never really judge how others perceived them or their practices. Furthermore, respondents felt that maintaining a consistent persona was a challenging feat, as personal issues could change their behaviors and often result in poor leadership practice that could have "spill-over" effects throughout the project. Respondents articulated that; "when you are at home, you are not always expected to be consistent with your leadership". But at work, leadership is a full-time commitment, and regardless of your situation, you still must be "calm", give "direction" and be able to "think" clearly.

Interviewees revealed the notion of fluctuating control. They felt that at times there were inconsistencies with regards to their control over and control within the project, with one junior respondent describing how it can often be confusing "who is in charge". The more senior interviewee emphasized how the lack of control can be "frustrating at times", and how "you just need to manage" working within the limits of your authority, whereas the junior project manager articulated the notion of leading "upward". Describing how leading also incorporates leading "your bosses" and the difficulties that arise when "convincing them to do things that you need". This respondent felt that you could only lead as much as you could and ultimately the final "decision can be out of your hands".

Challenges Relating to Project Focus

Respondents believed that operating as a leader within the civil construction industry was much more difficult than it would be in any other industry. One respondent highlighted the construction industry's culture of "resistance". He felt that there was a constant "resistance to change and resistance to do things differently" within the sector. Respondents also felt that there was a culture of "responsibility avoidance", in which industry professionals would constantly disassociate themselves from any levels of ownership or autonomy. Thus, making it much more difficult to lead, as individuals would often refuse to take ownership of their work. One

respondent also touched on the topic of health and safety appreciation and described how the civil construction industry is not "mature enough". This respondent felt that leading a safer environment was difficult due to the resistance from workers in which they perceive health and well-being initiatives "as making their job harder".

The overall consensus was that a project can significantly dictate leadership behavior. Respondents highlighted that the project setting meant that they were forced to "assess people quickly" and make "on the spot" decisions and judgments. They felt short projects imposed "minimal time upfront" to plan and implement leadership structure. Therefore, they lamented on "time" as an influential factor in leadership practice. There were many leadership activities that they wanted to adopt but felt they were unable to do so due to time and budget constraints. The emphasis from respondents was that they were not afforded the freedom to properly indulge in leadership practice, but given the time, they would have the opportunity to lead.

Table 15.3 summarizes the identified key challenges to meet leadership demands.

Table 15.3 Overview of challenges to meet leadership demands

Challenges relating to	
Team focus	• Effectively communicating in ways people expect to, or would like to receive information • Convincing members the reasoning behind what they are doing and how it impacts the bigger picture • Communicating differently, with different types of people, and in different types of situations • Constant changing of project team members • Understanding differences in styles, strengths and weaknesses • Difficulties with influencing the team you are given
Leader focus	• Understanding people and how they operate • Getting appropriate feedback • Pressures of maintaining consistency with leadership, even in times of difficult personal circumstances • Limits of authority • Final decisions are at times out of your control • Fluctuating levels of control can make authority ambiguous
Project focus	• Resistance of construction professionals to change • Responsibility avoidance among construction professionals • Lack of worker appreciation for health and well-being • Planning and thinking ahead with project time constraints • Budget doesn't allow for major leadership implementation • Project environment means you need to assess people very quickly

Potential for Digital Technology to Address the Leadership Challenges

In the second part of the exploratory interviews, the previous results were used to discuss how the capabilities of digital technologies could help overcome leadership challenges in the construction industry.

Respondents were able to articulate general and often vague methods they used in dealing with leadership challenges, ranging from "being consistent" to separating "work life from personal life". In short, the consensus was that "you just have to deal with it" as it comes. However, the junior respondent did suggest the use of the smartphone application "FaceTime" as a remedy to some communication difficulties faced. According to this interviewee, the lack of face-to-face communications was prominent in his position and made it difficult for him to effectively communicate. Therefore, the use of FaceTime allowed him to overcome these challenges and "bridge the gap" between less meaningful and meaningful communications. Respondents also felt that effectively communicating goals to all significant stakeholders could allow leaders to gain more authority within their positions. Furthermore, respondents believed that a great project leader could expand the objectives of the project itself beyond time, budget and quality, to involve "building" a "team" to "higher levels" of achievement in which everyone has "learnt", "grown" and "gained something".

Although the senior respondent was somewhat skeptical, the two more junior respondents could see immediate potential for digital technology to enhance leadership practice in some areas. One respondent made immediate connections with visual technologies, such as augmented reality, artificial reality, holographic display, as well as 3D, 4D and 5D BIM. He felt that these technologies would allow leaders to "sell the story" and help explain the reasoning behind objectives and goals. Moreover, this respondent felt that these technologies would be beneficial for leadership to "be able to show someone rather than tell them". "Having the capability of showing someone, having them see exactly what you are talking about, would help to sell that story". Respondents also felt that if the capabilities of auto-analytics were available, they would "use it tomorrow". The respondents were particularly excited about the capabilities to gauge factors like mood, sleep, physical and mental performance, as well as to provide real-time feedback on positive and negative words used in conversion, thus enhancing their leadership behavior. Furthermore, respondents believed that wearable technologies could assist with health and safety and would aid in a leader's responsibility to maintain workplace well-being and

Table 15.4 Digital technologies to assist with leadership challenges

Technology	*Capability*	*TF*	*LF*	*PF*
Smartphones	Can allow leaders to have a single mode of analysis for the project team, the project and themselves.	•	•	•
Wearables	Can allow leaders to maintain health and well-being of workers and positively influence organizational culture.	•		•
Drones	Can allow leaders to better track and report progress to team members, keeping them constantly informed and aware of their influence.	•	•	•
AR, VR & BIM	Can allow leaders to "show" rather than "tell", making it easier to "sell" the project story.		•	
Auto-analytics	Can allow leaders to understand themselves and their interactions with others.		•	

TF = team focus; LF = leader focus; PF = project focus

a positive environment. Respondents felt that utilizing wearables to build a cohesive team could influence the organizational culture, in which other organizations would "try to replicate" and "learn" from them, thus changing cultural perspectives within the industry.

However, all respondents argued that it would take more than technology to have an influence on the challenges outlined, and although they may assist with providing information, it would be up to the leader to use this information accordingly.

Table 15.4 outlines the key capabilities of digital technologies to assist with overcoming leadership challenges.

Discussion and Outlook

Digital transformation is fundamentally influencing all aspects of business and society. It can enable individuals and organizations to transcend from one way of working to another. In construction, emphasis is shifting from project management to project leadership, as professionals need assisting tools to not only aid in managing tasks and activities, but to aid in leading people.

Using the example of the Australian civil construction industry, we reviewed the existing body of knowledge on both industry-specific leadership demands and relevant digital technology capabilities to identify areas of improvement and to guide interviews with construction project managers. Findings indicate that digital technologies will have a positive influ-

ence on project leadership challenges. We find that project managers of civil contractors within Sydney understand the significance of leadership, but are often overwhelmed by its complexity. They demonstrate a willingness to adopt digital technologies as assisting mechanisms for leadership. Participants also lament on leadership challenges relating to project scope, industry attitudes and culture, level of authority, leadership habits/behavior, communication and team development. Construction professionals demonstrated specific examples in which they could utilize certain technologies to address challenges in their leadership practice.

Digital technologies can act as the catalyst that disrupts the behavior of a project team, which ultimately influences the environment in which the team operates. The input of digital technology within the leadership system can therefore transform leadership challenges into opportunities. As such, the adoption of digital technologies may hold the key in taking project leadership practice into the future. This chapter aims at opening up a discussion about exploring areas in which disruptive technologies can be used to create benefits beyond their primary technical applications.

References

ABI Research. (2017, December). *Wearable data analytics bring humans into the IoT*. Retrieved from ABI Research: https://www.abiresearch.com/press/wearable-data-analytics-bring-humans-iot/

Ahmed, M., & Ahmad, S. (2014). Transformation of smart phone to super phone: A future oriented gadget. *Pranjana, 17*(2), 1.

Akhavian, R., & Behzadan, A. (2016). Smartphone-based construction workers' activity recognition and classification. *Automation in Construction, 71*(1), 198–209.

Ameh, O., & Odusami, K. (2014). The leadership profile of Nigerian construction project managers. *Scientia Iranica: Transactions A Civil Engineering, 21*(4), 1241–1248.

Anantatmula, V. S. (2010). Project manager leadership role in improving project performance. *Engineering Management Journal, 22*(1), 13–22.

Arain, F. M. (2012). The quadrilateral model of leadership: Findings from a study on a mega project. *International Journal of Construction Project Management, 4*(2), 125.

Azhar, S., Jackson, A., & Sattineni, A. (2015). Construction apps: A critical review and analysis. In *ISARC. Proceedings of the 32nd International Symposium on Automation and Robotics in Construction* (p. 1).

Barber, E., & Warn, J. (2005). Leadership in project management: From firefighter to firelighter. *Management Decision, 7*(8), 1032–1039.

Blanco, J., Fuchs, S., Parsons, M., & Ribeirinho, M. (2018, April). *Artificial intelligence: Construction technology's next frontier*. Retrieved from McKinsey & Company: https://www.mckinsey.com/industries/capital-projects-and-infrastructure/our-insights/artificial-intelligence-construction-technologys-next-frontier

Cavaleri, S., & Reed, F. (2008). Leading dynamically complex projects. *International Journal of Managing Projects in Business, 1*(1), 71–87.

Chin, S., Kim, I., & Choi, C. (2017). What authentication technology should be chosen for construction manpower management? *Procedia Engineering, 196*, 309–314.

Clarke, N. (2012). Shared leadership in projects: A matter of substance over style. *Team Performance Management: An International Journal, 18*(3/4), 196–209.

Cullen, G. (2016). *The role of technology in health & safety*. Municipal World.

de Soto, B. G., & Adey, B. T. (2016). Preliminary resource-based estimates combining artificial intelligence approaches and traditional techniques. *Procedia Engineering, 164*, 261–268.

Del Rowe, S. (2017, September 29). *Digital transformation needs to happen now*. Retrieved from Destination CRM: https://www.destinationcrm.com/Articles/ReadArticle.aspx? ArticleID=120789

Dillow, C. (2016, September 13). *The construction industry is in love with drones*. Retrieved from Fortune: http://fortune.com/2016/09/13/commercial-drone-construction-industry/

Emil Berg, M., & Terje Karlsen, J. (2014). How project managers can encourage and develop positive emotions in project teams. *International Journal of Managing Projects in Business, 7*(3), 449–472.

Galvin, T., Gibbs, M., Sullivan, J., & Williams, C. (2014). Leadership competencies of project managers: An empirical study of emotional, intellectual, and managerial dimensions. *Journal of Economic Development, Management, IT, Finance, and Marketing, 6*(1), 35.

Gruman, G. (2016). What digital transformation really means. *InfoWorld, 18*(1), 1–3.

Hambling, D. (2015). A drone that learns. *NowScientist, 226*(3017), 20.

Hansen, S. (2017, May 25). *Analytics in the construction sector*. Retrieved from Sourceable: https://sourceable.net/analytics-construction-sector/

Heembrock, M. (2015). The risks of wearable tech in the workplace. *Risk Management, 62*(1), 10.

Heiskanen, A. (2016, July 27). *Augmented and mixed reality in construction*. Retrieved from AEC Business: https://aec-business.com/augmented-mixed-reality-construction/

Joan, D. (2015). Enhancing education through mobile augmented reality. *Journal of Educational Technology, 11*(4), 8–14.

Kappagomtula, C. L. (2017). Overcoming challenges in leadership roles–managing large projects with multi or cross culture teams. *European Business Review, 29*(5), 572–583.

Kassem, M., Benomran, L., & Teizer, J. (2017). Virtual environments for safety learning in construction and engineering: Seeking evidence and identifying gaps for future research. *Visualisation in Engineering, 5*(1), 16.

Khan, S. S. (2015). Importance of transformational leadership in project success: A theoretical framework. *Актуальні проблеми економіки, 1*, 67–76.

Kim, K. P. (2016). Investigation of readiness for 4D and 5D BIM adoption in the Australian construction industry. *Management Review: An International Journal, 11*(2), 43.

Kivrak, S. A., Arslan, G., & Tuncan, M. (2015). Implementing augmented reality in construction projects. *Applied Mechanics & Materials, 719–720*, 197–201.

Klashanov, F. (2016). Artificial intelligence and organizing decision in construction. *Procedia Engineering, 165*, 1016–1020.

Kouzes, J. M., & Posner, B. Z. (2003). *The leadership practices inventory (LPI): Participant's workbook* (Vol. 47). John Wiley & Sons.

Lazányi, K., & Dóka, L. (2015). Leadership practices in the Hungarian construction industry. *Managerial Challenges of the Contemporary Society. Proceedings, 8*(2), 55.

Lee, W., Kang, S., Moh, R., Wu, R., Hsieh, H., & Shu, Z. (2014). Application of BIM coordination technology to HSR Changhua station. *Visualization in Engineering, 3*(5). Retrieved from https://viejournal.springeropen.com/articles/10.1186/s40327-014-0008-9

Leviäkangas, P. P. (2017). Keeping up with the pace of digitization: The case of the Australian construction industry. *Technology in Society, 50*, 33–43.

Linder, C. (2017, June). *Wearable technology could save lives and dollars in construction industry*. Retrieved from Pittsburgh Post-Gazette http://www.post-gazette.com/business/tech-news/2017/06/12/Wearable-technology-construction-industry/stories/201706040035

Moon, S., & Seo, J. (2015). Integration of smart glass technology for information exchange at construction sites. *ISARC. Proceedings of the 32nd International Symposium on Automation and Robotics in Construction*, p. 1. Vilnius Gediminas Technical University, Department of Construction Economics & Property.

Moreno, C. (2014, June 21). *Free CAD software for mobile devices*. Retrieved from Cadalyst: http://www.cadalyst.com/cad/free-cad-software-mobile-devices-20019

Morsy, A. (2016). Emotional matters: Innovative software brings emotional intelligence to our digital devices. *IEEE Pulse, 7*(6), 38–41.

Müller, R., & Turner, R. (2010). Attitudes and leadership competences for project success. *Baltic Journal of Management, 5*(3), 307–329.

Nagy, P., & Koles, B. (2014). The digital transformation of human identity: Towards a conceptual model of virtual identity in virtual worlds. *Convergence, 20*(3), 276–292.

Oke, A. (2010). An examination of project management leadership styles of Nigerian quantity surveyors. *Journal of Building Performance, 1*(1), 57–63.

Park, J. K. (2016). Framework of automated construction-safety monitoring using cloud-enabled BIM and BLE mobile tracking sensors. *Journal of Construction Engineering and Management, 143*(2). Retrieved from https://www.academia.edu/35328870/Framework_of_Automated_Construction-Safety_Monitoring_Using_Cloud-Enabled_BIM_and_BLE_Mobile_Tracking_Sensors

Patterson, J. (2010). Leadership: The project management essential. *Production and Inventory Management Journal, 46*(2), 73.

Phelps, K. (2014). "So much technology, so little talent"? Skills for harnessing technology for leadership outcomes. *Journal of Leadership Studies, 8*(2), 51–56.

Pink, D. (2006). *A whole new mind: Why right-brainers will rule the future*. Penguin.

Pryke, S. L. (2015). The effect of leader emotional intelligence on leader–follower chemistry: A study of construction project managers. *Construction Management and Economics, 33*(8), 603–624.

RGA—Reinsurance Group of America. (2013, October 31). *Auto-analytics—Quantifying the self*. Retrieved from RGA—Reinsurance Group of America https://www.rgare.com/knowledge-center/media/articles/auto-analytics%2D%2D-quantifying-the-self#

Richardson, C. (2017). Technology offers the construction industry a brighter future. *Construction Engineering Australia, 3*(1), 16.

Sampaio, A., & Viana, L. (2014). Virtual Reality technology used as a learning tool in civil engineering training. In *Human System Interactions (HSI), 2014 7th International Conference* (pp. 156–161). IEEE.

Schuhmann, R. (2010). Engineering leadership education—The search for definition and a curricular approach. *Journal of STEM Education: Innovations & Research, 11*, 61–69.

Senam, M., Rashid, K., Sarkawi, A., & Zaini, R. (2014). Construction project leadership from the perspective of Islam. *International Journal of Islamic Thoughts, 6*, 46.

Senaratne, S., & Samaraweera, A. (2015). Construction project leadership across the team development process. *Built Environment Project and Asset Management, 5*(1), 69–88.

Slattery, D., & Sumner, M. (2011). Leadership characteristics of rising stars in construction project management. *International Journal of Construction Education and Research, 7*(3), 159–174.

Succar, B. (2009). Building information modelling framework: A research and delivery foundation for industry stakeholders. *Automation in Construction, 18*(3), 357–375.

Toor, S., & Ofori, G. (2008). Taking leadership research into future: A review of empirical studies and new directions for research. *Engineering, Construction and Architectural Management, 15*(4), 352–371.

Tyssen, A., Wald, A., & Spieth, P. (2014). The challenge of transactional and transformational leadership in projects. *International Journal of Project Management, 32*(3), 365–375.

Wright, R., & Keith, L. (2014). Wearable technology: If the tech fits, wear it. *Journal of Electronic Resources in Medical Libraries, 11*(4), 204–216.

Yngvason, Y. R. (2013). Jesus Christ as a project leader. *Procedia-Social and Behavioral Sciences, 74*, 398–407.

Zhang, L., & Fan, W. (2013). Improving performance of construction projects: A project manager's emotional intelligence approach. *Engineering, Construction and Architectural Management, 20*(2), 195–207.

CHAPTER 16

The Making of Data-Driven Sustainable Smart City Communities in Holiday Destinations

Sudipta Kiran Sarkar, Michalis Toanoglou, and Babu George

Introduction

Sustainable development is a term that emerged from the Brundtland Commission's report of 'Our Commons Future' in 1987 (Saarinen, 2006). The concept of sustainable development entails a responsible practice of resource consumption by current generations so that future ones are not deprived of the same resources for their basic survival (WCED, 1987; Saarinen, 2006). The argument in favour of advocating sustainable development got further impetus in institutional terms in the 'Earth Summit' in 1992 held in Rio de Janeiro, Brazil since the Rio Conference

S. K. Sarkar (✉) • M. Toanoglou
Woosong University, Daejeon, South Korea
e-mail: mtoanoglou@sis.ac.kr

B. George
Fort Hays State University, Hays, KS, USA
e-mail: bpgeorge@fhsu.edu

© The Author(s) 2020

B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2_16

in 1992 on account of the observation of manifold negative ecological and social impacts caused by decades of industrialization (Hanharan, 2010; Saarinen, 2006).

The perspectives of sustainability in tourism emanated from the Agenda 21, an action plan on sustainable development that emerged from the discourses of the Rio Conference. Sustainable development issues have become crucial in current tourism discourses (Roberts & Tribe, 2008)—given the negative impacts of tourism operations in terms of social, political, economic and environmental aspects. According to United Nations World Tourism Organization (UNWTO) 'Sustainability principles refer to the environmental, economic and socio-cultural aspects of tourism development, and a suitable balance must be established between these three dimensions to guarantee its long-term sustainability' (UNWTO, 2004).

Since the last three decades, several initiatives primarily by non-profit organizations as well as academic, governmental and inter-governmental organizations have pursued sustainability, both in practice and dialogue through regional, inter-regional and global projects, grassroots-based activism as well as workshops and conventions. In addition, private sector/industry initiatives towards sustainability have also been evident in spite of the conflicting interests emanating from a dominant neo-liberal economic environment that has shaped the industry perspectives in tourism. However, climate change issues and inequality in tourism destinations have been ever increasing despite various pursuits of sustainable development by different sectors in the tourism industry. Hence, the quest for sustainability in tourism is far from over due to the increasing environmental and social challenges faced globally.

Technological developments and innovations have brought us to a new phase in our modern society. Cutting edge technologies in digital, environmental and transportation domains have paved the way towards a new epoch of industrialization that provides accessibility, communication and socialization to the masses via affordable means. Developments in more efficient, faster and safe mass public transit systems, as well as developments in green energy that have made it relatively more affordable and cost-effective to a much larger section of consumers and businesses, have brought sustainability and technology closer and compatible with each other than ever before. Clean energy and cleaner modes of transportation driven by new innovative technologies and coupled with smart digital technologies have brought new opportunities to the process of transformation towards a sustainable future. Clean energy, cleaner modes of trans-

portation systems as well as smart digital systems have fostered the development of sustainable urban tourism by virtue of hospitality outlets and urban attractions increasingly adopting green energy technologies—clean modes of transportation providing affordable, environmentally clean and efficient mobility services to tourists and digital technologies (powered by widely available Wi-Fi services) enabling communication and availability of information critical for destination knowledge and travel decision-making. Public/state funded technological projects around the world are leading the way for more innovations in cleaner automobiles and cleaner energy production—ARPA-E in the United States and Germany's KfW, a state investment bank have funded a number of clean technology initiatives with relation to energy efficient and renewable projects in the recent past (Tate, 2017). Tesla, a leading producer of electric cars in the United States has been funded by the US Department of Energy for producing electric automobiles with cutting edge technology. Returns from green technology investments initiated both by state and the private sector has resulted in creation of wealth for the public providing more coherence and impetus to the role of technology in fostering sustainable development (Harris, 2017).

The bond between sustainability and technology will mostly be necessitated by the imperative needs in cities and urban areas of the world. Growing complexities in ever expanding cities of the world resulting from the demands for quality of life, environmental protection and competitiveness will drive the need for sustainable development powered by clean energy technology, soft mobility technology and more intelligent digital technologies (Siemens AG, 2013). Moreover, cities and other urban areas around the world, which currently produce approximately 70% of the greenhouse gas emissions, will increasingly become the epicentres of high volumes of consumption leading them to be major contributors to climate change (Siemens AG, 2013). Hence, the demand for innovative technologies that can facilitate sustainable and inclusive urban growth will necessitate effective solutions and strategies that can ensure resilient and environmentally compatible ways of urbanization of our cities in the future.

One of the prime areas of economic activity and consumption in cities around the world is tourism. Cities and urban areas vibrant with arts, culture, music, architecture and design and heritage along with centres generating robust volumes of business and commercial activity have attracted all kinds of visitors. Moreover, ease in mobility and communication due to rapid digitalization and internet connectivity via Wi-Fi as well as expansion

of different modes accessible and efficient public transportation networks have made tourism in urban areas grow manifold. Urban tourism as referred to by UNWTO (2012) is 'trips taken by travelers to cities or places of high population density'. Cities which are administrative capitals of nations are gateways to many destinations and the first point of reference for tourists in acquiring an impression of the nation as a destination. This makes large number of tourists travel to and via cities and hence crucial for destination management organizations to manage cities as destinations in a way that they can facilitate quality experiences and enhance satisfaction of such large mass of tourists.

This chapter discusses the magnitude of technology being strongly implanted in the process of sustainable development of cities as manifested through urban tourism—one of the main pillars of urban economy. The chapter identifies four critical dimensions of sustainability in urban tourism—urban ecotourism, food tourism, urban culturescapes and smart transportation. Correspondingly, critical dimensions of technology in urban tourism—social media and user-generated content, gamification, M-learning (mobile-based learning), green energy technologies and green/soft mobility, are also discussed comprehensively. The chapter then discourses the integration between technology and sustainability in urban tourism contexts based on existing cases and models around the world.

Dimensions of Sustainability in Urban Tourism

Culture, nature and communities in urban areas have been instrumental in shaping the tourism landscape of cities around the world. City-based parks, gardens and waterbodies as well as museums, arts galleries, performing arts, avant-garde arts and countercultural communities have attracted large number of culturally inclined independent travellers. On the rise are emerging forms of urban travel activities like food tourism combined with walking tourism that involves trips to areas of cities noted for local food—wet markets, restaurants, confectionaries and all other forms of eateries and food outlets. Also associated with such tours are the communities engaged with food tourism businesses, their enterprises, the aesthetic aspects of their culinary offerings, their skills as well as the socio-cultural environment they create and belong to as a whole. Another aspect that culture-oriented city walking tours involve are visitations to prime natural and cultural attractions of urban areas harnessing mass transit public transport systems. All these elements reflect socio-cultural inclusivity,

environmental compatibility and offer deeply engaging lived as well as socializing experiences to tourists. Such increasingly popular forms of urban tourism have been receiving adequate patronage from tourists and destination planners in the recent times as the modern tourism discourses progressively reflect the virtues and the essentiality of sustainability.

Urban Ecotourism

Towns and cities have large tracts of nature either in the form of national parks, sanctuaries, wildlife reserves and other organic forms or in modified or artificially created forms. Urban centres around the world are increasingly trying to harness the value of such nature tracts in providing sustainable living to both its residents and visitors. Urban tourism now significantly involves travel to nature-based areas in and around cities especially to such natural areas, which are of high significance in terms of flora or fauna or both. Such trips help both residents and visitors understand the value of nature as well as enable them to have a sustainable approach to experience cities. Such a phenomenon can be referred to as 'urban ecotourism'. Though the term 'urban ecotourism' was first coined in 1996, the phenomenon has been in practice since long. Planeta.com, an award-winning website on ecotourism, in a conference in 2004, described it as 'simply nature travel and conservation in a city environment' (Kastelein, 2004). The definition laid down by the Toronto Green Tourism Association (2006) was 'travel and exploration in and around a city that provides visitors and residents with a greater appreciation of the cities' natural resources and cultural resources'.

Cities like New York, Toronto, Singapore, Hong Kong and Kuala Lumpur have attracted local and international visitors towards both natural and modified nature-based attractions. Nature-based attractions in these cities offer opportunities for various leisure and adventure activities like bird-watching, nature-viewing, photography, caving, hiking and rock-climbing. Such opportunities have made nature-based attractions in these cities an integral part of their urban tourism offerings to all kinds of visitors. One of the prime features of urban nature-based attractions in most cases is that they are often well-connected by mass transit public transportation systems due to their location within or on the periphery of cities and major towns. Hence, travel to urban nature-based attractions or urban ecotourism contributes towards sustainable development as it creates opportunities for alternative recreational activities as against the traditional

highly consumptive urban tourism activities. Urban ecotourism apart from primarily focusing on travel to nature-based areas in urban areas also involves trips to cultural sites like museums, art galleries, handicraft streets, wet markets and food streets. This provides the social and anthropological element to the urban ecotourism phenomenon making it inclusive and environmentally inclined at the same time. Urban ecotourism offers the opportunity to all sections of the society to be educated on the value of ecology to human societies as nature-based attractions offers easy access by the virtue of being located within or close to cities and by being well-connected efficient and affordable mass transit transport systems.

Food Tourism

Food tourism is the phenomenon of travelling and experiencing other cultures through food (Herrera, Herranz, & Arilla, 2012). Food tourism is also sometimes referred to as culinary tourism or gastronomique tourism. Food tourism entails a sense of place in geography (or the link between land and people). Travel for the purpose of experiencing exquisite culinary offerings of a destination has become one of the main motivators of travellers in terms of their travel behaviour and activities at a destination (Herrera et al., 2012; OECD, 2012). Such culinary experiences are offered in an array of settings—the most exciting and engaging ones being local street food at popular street food enclaves of towns and cities, wet markets or night markets as well as local/traditional food outlets which by virtue of their unique offerings (taste, aroma and aesthetic style of presentation) for a considerable period of time have become transformed into a heritage due to the sustained popularity that they have enjoyed (Gillan, 2014; Nualkhair, 2015). Experiencing food related events like food festivals or experiencing hands-on cooking (participating in cooking classes) as well as food tasting and/or experiencing the attributes of specialist food production are also some of unique experiences that urban destinations offer to foodie travellers (Peltier, 2015).

Apart from the hedonic and consumptive aspects associated with the experiences of food tourism, the communal and cultural elements that are strongly embedded in it often construct the basis of such experiences (Peltier, 2017). A critical aspect associated with these communal and cultural elements is sustainability. Food tours are often characterized by walking tours for short durations of 4–5 hours as well as harnessing the services of mass transit public transportation systems. The use of such soft modes

of mobility enable food tours to deeply involve the communal element as it provides a social interface between foodie travellers and the communities engaged with food tourism businesses, the small scale nature of the enterprises they run as well as the aesthetic aspects of their culinary offerings along with their culinary skills (Gillan, 2014; Nualkhair, 2015). This social interface constructs a socio-cultural environment that immerses the visitors and their gazes with deeper social meanings of their experiences of local culinary cultures. Hence, the utilization of soft mobility and the socially immersive experience reflect the sustainable nature of the communal and cultural elements of food tourism.

Urban Culturescapes

Cultural activities ranging from visitations to art galleries and museums depicting different aspects of aesthetic and social scenes of modern society as well as radical/avant-garde and contemporary forms of arts and events mainly advocated by the urban youth shape the culturescapes of our cities (Merrick, 2011; Buffenstein, 2016). Urban areas around the world encompass both mainstream and alternative forms of visual and performing arts along with edifices of heritage like museums and monuments that attract a large number of travellers to urban areas.

Cities like Berlin, London and New York have attracted culturally inclined travellers to art galleries and art collectives that focus both on popular and 'avant-garde' forms of aesthetic expressions. Attractions like the Museum island in Berlin which consist of museums and galleries as well as alternative art centres like Kreuzberg, Friedrichshan and Prenzauler Berg attract all kinds of politically and art-inclined visitors. Many of such culture streets represent underground and alternative urban youth cultural expressions in the form of street musicians and street art collectives. Cultural centres like these in many instances have been the backdrop of radical urban socio-political movements that gave rise to countercultural thoughts challenging traditional Capitalist consumption models and advocating collective action and Do-It-Yourself ethos. Such actions and ethos have become underlying principles of many urban sustainable projects. Urban cultural centres like these often undergo continuous evolution and in the process have attracted a wide array of visitors—art-inclined people and artists themselves, youth belonging to avant-garde movements as well as with visitors based on socio-political factors like immigrants and anarchists. Urban culturescapes that involve such vibrant youth cultural vibes

offer visitors the experience and learning on ecological and social dimensions of urban sustainability through different artistic expressions and other culturally expressive activities.

As in the case of urban ecotourism and food tourism, visitations to urban landscapes also involve the harnessing of the urban mass transit public transport system that connect major centres of cultural sites with various important mass transit hubs.

Soft Mobility

Soft mobility refers to low emission transportation systems that provide mobility in urban spaces to the masses. Such low emission transportation systems operate in a variety of modes like inter-modal, waterborne, rail, public/collective transport and non-motorized individual mobility (Rocca, 2009). Soft mobility should involve—the consideration of the urban destination geographical features, legal and planning structures anticipated ecological, social and financial impacts of development as well as the needs and wants of the residents of the urban destination (Kramer, 2009; Transdanube, 2014).

Cycling/biking in urban streets has become a popular phenomenon around the world. Tourists and residents alike commute considerable distances in their daily schedules in cities with the help of cycling/biking. Several urban destinations have become bike-friendly cities encouraging visitors to utilize cycling paths as a way to reduce congestion and automobile-based pollution. Cheap and affordable mass transit systems like subway trains, monorails, trams and electric buses have provided easy accessibility to different cultural, nature-based and gastronomique streets to all types of visitors (Rocca, 2009). Therefore, one of the emphasizing factors emerging from soft mobility is to convert more tourist movement into sustainable modes of transport which can also be referred to as modal shift (Gebhardt et al., 2016). An inter-modal chain model of mobility for tourists involving the combination of waterborne, rail and public/collective transport as well as walking and cycling ensures a sustainable consumption from tourism movement within an urban destination (Gebhardt et al., 2016).

The rise of collective transport or transport options that have emerged under the sharing economy has also reduced the rate of consumption and offered a more convenient, reliable and safe mobility options for tourists in cities and towns. Sharing economy in transport refers to carpooling

services where car journeys are shared by more than one person. Several carpooling initiatives like Lyft, MyCotra and Uber have become commercially successful and are contributing towards less congestion and reducing individuals' travel costs.

In the Asian context, major cities like Singapore, Hong Kong, Seoul and Tokyo offer a combination of transport options ranging from subways or referred to as MTR and MRT (in Hong Kong and Singapore respectively) supported by extensive citywide bus networks along with options for biking-cycling at different parts wherever possible.

The Case of Hong Kong as a Soft Mobility Destination

In Hong Kong, the MTR (Mass Transit Railway), a combination of heavy rail, light rail and feeder bus service, is one of most extensive and profitable mass transit systems in the world and offers itself as a successful model for many other mass transit systems coming up around in the world and particularly within Asia. In addition, Hong Kong offers the option of using tram services in certain parts of the highly congested Hong Kong Island as well as extensive cycling opportunities in the new territories. Besides, waterways—a traditional mode of transport that have existed in Hong Kong since long still operates significantly offering services from Hong Kong Island (Central) and Kowloon ferry terminals to different parts of Kowloon, Hong Kong Island, New Territories and several outlying islands. The tram and ferry services are also a major part of the tourist experiences of Hong Kong. Tourist—centric transport like the furnicular rail or the Peak tram which covers a distance of 1.4 kilometres approx. starting from downtown Hong Kong Island to the Victoria Peak, a point of commercial and tourist importance at the higher elevations of Hong Kong Island enables both tourists and local residents a convenient commuting option. Another tourist-centric transport is the Ngong Ping cable car which starts from the Tung Chung MTR station to the tourist attractions in the higher elevations of Lantau island (Lantau peak) covering a distance of 5.7 kilometres and providing breathtaking views of the Hong Kong International Airport, the dense flora and topography of Ngong Ping plateau as well as the urban areas around. The Hong Kong International Airport and the

(*continued*)

(continued)
Hong Kong Disneyland are connected with the main commercial and residential areas of Hong Kong by means of MTR's special services—the Airport Express and the Disneyland Resort Line respectively. The urban bus network operated by different companies as well as the Green Liquified Petroleum Gas (LPG) light buses (minibuses) offers a comprehensive and highly well-connected system of road transportation to residents and tourists alike. Road transport in Hong Kong is also featured by three kinds of LPG run taxicabs—the red, green and blue taxicabs. The red taxicabs operate in the downtown areas of Hong Kong Island and Kowloon, the green taxicabs operate in the New Territories of Hong Kong and the blue taxi cabs operating only in the Lantau Island.

This comprehensive mass public transport system of Hong Kong offers a perfect example of an inter-modal model of soft mobility in an urban tourism destination. Firstly, the MTR services, the ferry services and bus network (especially the green LPG minibus services) as well as the extensive cycling-biking options which are highly interconnected enable tourists to commute to different famous and offbeat folk cultural, local culinary and heritage attractions of Hong Kong. Secondly, the 'Peak Tram' and the Ngong Ping Cable car offer exciting and safe tourist-specific experiential commuting to a large of urban tourists and resident visitors. Thirdly, the engaging experiences tourists obtain while availing ferry services and the tram services to commute around Hong Kong transform such transport services into tourist attractions. The extensive MTR network, cycling-biking options as well as the extensively operating LPG run cabs and minibuses along with tourist-centric transport options of the 'Peak Tram' and the Ngong Ping Cable car aptly manifest an eco-friendly and inclusive mass-oriented soft urban mobility system. Due to this dynamic nature of the urban inter-modal in the city, Hong Kong has been ranked first in terms of having the most sustainable transport as per the 2017 Sustainable Cities Mobility Index by Arcadis (Arcadis, 2017).

Dimensions of Technology and Their Integration with Dimensions of Sustainability in Urban Tourism

Cities and urban regions around the world are the most technologically vibrant in terms of use and applications. Urban infrastructure provides the basis and urban consumption patterns necessitate the presence and use of technology at considerable levels. Cities are connected with Wi-Fi technologies and its commercial activities are highly dependent on the use of innovative tech applications for various interconnected systems to work efficiently. Urban regions around the world are generally equipped with technological connectivity in terms of Wi-Fi availability, internet zones and social media sharing that benefits visitors and tourists to derive satisfaction and intention to re-visit urban attractions.

Social Media and User-Generated Content

Social media founded on the technological applications of user-generated content (Web 2.0) plays a major role in communication and socialization among tourists (Buhalis & Law, 2008; Kaplan & Haenlein, 2010). Urban destinations facilitated by easy availability of Wi-Fi services drive online activity of tourists who prefer to be connected while they are in the process of consumption of urban leisure-based experiential products (Bock, 2015). Facebook, which enjoys the highest number of users in the world, followed by Twitter and travel-oriented platforms like Trip Advisor—have been the widely used social media platforms by tourists for online social interactions, sharing of vital travel information influencing travel decision-making by tourists (Santos, 2011). Online social exchanges and sharing of information by tourists in online platforms have led to co-creation of collective knowledge (Seraj & Ayesugul, 2012).

Eco-conscious tourists in urban settings have the ability of setting ecological citizenship through continuous co-creation of knowledge that is collectively generated through their online social interactions and sharing of information on ecological and social elements they experience in nature-areas in urban settings (Rokka & Moisander, 2009; Sarkar, Au, & Law, 2013). The advent of Flickr, a photo-sharing social networking site and recently, Instagram—a photo-sharing social networking app has initiated online co-production of destination knowledge through pictures and photographs. This has enabled tourists create aesthetically credible pictorial connotations of destinations that have deeper influences on fellow-tourists.

Urban tourism destination-based authorities and promotional bodies have harnessed the co-creative and collective knowledge building potential of social media platforms like Facebook and Instagram (Sarkar et al., 2013). Promotional efforts on social media platforms and initiating co-creative practices among consumers have become a key innovative marketing strategy that transforms consumers into pro-sumers. Co-created content by tourists on social media can become invaluable marketing content for destination management organizations (Özdemir & Çelebi, 2015).

The advent of smartphones in the last 5–7 years have enhanced the role of social media use among urban tourists. With the increasing availability of Wi-Fi facilities and high-speed internet connectivity in public places, cultural sites as well as natural sites, smartphones have accelerated the volume of online activity of tourists. Smartphones have drastically reduced the time between tourists experiencing an attraction and their sharing of pictorial, audio-visual as well as textual information in various social media platforms based on such experiences. Tourists can very quickly engage in co-creative practices on social media platforms in the context of urban destinations due to the widespread availability of Wi-Fi and increasing use of smartphones and smartphone-based apps. Tourist-generated (and co-created) content in social media platforms primarily via smartphones have become more reliable and easily available than supplier-generated content resulting in effective travel decision-making and collective knowledge building by existing and prospective tourists.

In the context of urban tourism, both tour operators and tourists make use of social media platforms like Facebook and Instagram to co-create experiences and products of nature, culture and gastronomical tourism. Moreover, blogs maintained by tourists on urban nature, food and cultural attractions play a major role in initiating collective knowledge building through sharing and exchange of key information. Such collective knowledge or tourist/consumer generated knowledge become more intense and reliable than supplier-generated knowledge as it involves a dialectic process of critical and comprehensive introspection of content. Moreover, through such dialectical process of collective knowledge building, tourists establish and advocate ecological and cultural citizenship online.

Gamification in Tourism

Gamification in tourism is ‘the use of game mechanics and experience design to digitally engage and motivate people to achieve their goals’ (Burke, 2014). Gamification can be a great method to digitally engage

visitors with a use of their smartphones, tablets or other digital devices (Sever, Sever, & Kuhzady, 2015).

Gamified tourism experiences often involve challenges in the form of quizzes, puzzles, making and submitting photos according to a task. Scores based on these challenges determine how well the participant has accomplished the challenge (Kazakova, 2015). To monitor the participants there are leaderboards who keeps them updated how they are progressing in the game in terms of the rounds of challenges to be completed and the scores scored. Those completing all stages of the game successfully with high scores are rewarded in both virtual and real (tangible) forms (Kazakova, 2015).

Gamified tourism experiences are offered in different forms especially in the urban context. Some of the prominent forms are location-based augmented reality games, gamified tours specific to urban environments, augmented experiences of theme parks and transmedia storytelling, gamified immersive experiences in cultural and natural heritage and virtual cultural heritage experience, gamified restaurant and hotel experiences.

As mentioned earlier, one of the main aims to offer gamified experiences to tourists is to immerse them in a situation that is highly subjective and enduring in nature and therefore can provide high levels of satisfaction to them. Gamified experiences have the potential to bring loyalty among tourists for the destination and enable repeat visits. Gamified experiences are an attempt to add value to the generic trip experience of tourists while they are visiting an urban destination (Sever et al., 2015).

Applying gamified experiences in the area of sustainable tourism can facilitate social interactions, cultural understanding, inculcate sustainable consumption habits and bring about an enhanced quality of experience (Negruşa, Toader, Sofică, Tutunea, & Rus, 2015). Gamification in sustainable tourism is aimed at engaging as well as encouraging tourists to advocate and internalize new sustainable habits, sustainable consumption or even engage in new sustainable activities and visit new destinations that are sustainable. Gamified experience in the context of sustainable tourism has the potential to provide intrinsic incentives to tourists while they are engaging in activities that enable them to derive self-esteem and social recognition (Negruşa et al., 2015). The merits of applying gamification in sustainable tourism involve usher engagement before during and after any nature-based experience as well as marketing and promotional benefits. Gamification by means of improving the experience offers an opportunity to attract new visitors and new markets.

M-Learning or Mobile Learning

M-learning or mobile-based learning is a process of learning and education through the means of movable electronic computing devices like smartphones and tablets. The use of M-learning in the hospitality industry has been evident in the recent times in areas of hospitality workers' training and education. Hospitality chains like the Marriot and Hilton are applying M-learning applications for recruitment of new workforce in their respective hospitality properties.

The potential application of mobile learning in the context of the touristic use can be realized in the form of environmental interpretation and visitor education within the context of urban natural settings like botanical gardens, city parks, urban groves and other similar places. Utilizing the current developments and innovations in mobile broadband networks, smartphone technology and mobile software applications, the M-learning process can provide engaging interpretive programmes involving free-choice learning and mindful visitor experiences to tourists in urban natural settings.

Soft Mobility and Green Energy Technologies in Urban Tourism

As discussed earlier, soft mobility involves a set of different forms of low emission transport in urban destinations—an inter-model mode of transport. Soft mobility forms have facilitated accessibility to the masses and brought about social transformations and quality of life in many parts of the world. Inter-modal soft mobility has provided inclusivity and an egalitarian mode of transport in which lower income urban communities have found easy and convenient access to different parts of urban areas where they have to commute for earning a means of livelihood and education. Efficient and affordable inter-modal soft mobility available for local residents from all sections of the society also provide accessibility and connectivity to tourists and visitors in urban destinations. Such inter-modal soft mobility forms are driven by technologies that provide access to difficult terrains and dense settlements evident in urban destinations. Transport technologies like cable car and low emission automobiles using new forms of renewable fuel as well as electricity have provided greater social sustainability. These are the emerging smart mobility technologies that have enabled visitors from all economic sections or urban societies to affordably have access to attractions of ecological and socio-cultural interest.

The Case of Gondolas or Cable Cars in Latin American Cities and Hong Kong

Gondolas or cable cars are considered as a cheap option which requires lesser investment for infrastructure and lesser need of displacement of people and communities compared to other mass transit forms like subways or rail tracks. Gondolas may not have the ability alone to transport a large number of people from one to another but support significantly other forms of mass transit system by playing a major role in inter-modal connectivity (Barber, 2017). Gondolas on one hand can facilitate tourist movement within and to attractions in urban destinations and on the other hand also enable to move local residents from their neighbourhoods to major centres of cities at an affordable cost especially where the urban geographical landscape is characterized by challenging topographies. In particular, gondolas have been able to provide a faster, time-saving and affordable mode of transport to urban residents living in the poorer parts of cities and linking them with commercially important parts where there are opportunities for earning a livelihood. Latin American cities like La Paz (Bolivia), Medellin and Cali (Columbia) and Caracas (Venezuela) are major examples of such socially inclusive and effective gondola systems in commuting terms for the masses (Barber, 2017).

The Metrocable in Medellin city, Columbia is regarded as the world's first integrated multi-line gondola system. The Metrocable connects with major business and commercial centres of Medellin city with many outlying and less wealthy areas of the city in upper elevations of hillsides. The Metrocable was opened in 2004 and since then has been credited on one hand, in reducing commuting times, providing cheap and efficient connectivity to many low income residents of outlying areas and on the other hand, has been significantly instrumental in reducing crime rates and urban poverty levels in the city(Barber, 2017).

Mi Teleférico, the gondola system in La Paz, Bolivia is considered as one of the most developed and longest in the world. The Mi Teleférico was opened in 2014 and since then has enabled 50 million people to commute from poorer parts of the city to the upscale business and commercial areas. It has been reported to have saved 652 million minutes for all commuters it has served since 2014. The Mi Teleférico has been claimed to be a very effective project and the city authorities are considering to expand it to seven more lines which are expected to be covering a distance of 20 kilometres.

(*continued*)

(continued)

The Caracas Metrocable in Caracas, Venezuela operates with the similar philosophy of providing connectivity for the city's poorer parts with the affluent parts. It was built to avoid displacing people and providing poorer parts of Caracas like the San Agustin a safe and efficient mode of transport to commute especially within such areas. The Caracas Metrocable primarily connects with other major forms of public transport systems of Caracas and provides transportation within the San Agustin area of Caracas.

All the above examples manifest an environmentally significant and socially inclusive efficient form of urban soft mobility driven by the Cable car or gondola technology. Such a gondola system has proven to have provided safe and crime free transportation which becomes an important factor in tourist's use of such soft mobility systems.

The Ngong Ping 360 cable car in Hong Kong and the Singapore Gondola are on the other hand, urban projects that have been purpose-built for tourist mobility. These cable car projects have been able to divert tourist movement away from the vehicular traffic-prone areas of the city. Moreover, these projects have reduced the dependence of road transportation that involves automobiles run by fossil fuels especially in the case of Lantau Island, Hong Kong.

The Case of Electric Automobiles and Bio-Fuel-Driven Urban Buses

Another variety of smart mobility technologies are electric/hybrid vehicles exemplified in the form of electric cars and bio-fuel driven vehicles. These vehicles run on fuel from renewable sources like bio-fuels and electricity.

One of the recent and best examples of bio-fuel driven is the Bio-bus experimentations in the UK. Bio-bean, a London-based enterprise, has taken the initiative to run a project of Bio-buses driven by the bio-fuel made out by combining oil extracted from old coffee grounds and diesel (Hirtenstein, 2017). Utilizing old coffee grounds from major coffee shops across the UK, it is claimed that 6000 litres (1583 gallons) of bio-fuel or B20 bio-fuel can be produced to run a sizable fleet of buses for public transport in London city (Hirtenstein, 2017). Though bio-fuel is already used to run 9500 buses, the 'coffee-powered' bio-bus is an innovative soft mobility technology to be used for the first time in the city (BBC, 2017).

(*continued*)

(continued)

Another bio-fuel-driven bus project also from the UK is the Bio-Bus project also known as GENeco Bio-Bus. The Bio-Bus is a project powered by bio-fuel in the form of gas formed out of human wastes of food, sewage and commercial liquid wastes (Geneco, 2016). Though two Bio-Buses were operating in 2015, the Office for Low Emission Vehicles turned down the bid for funding and operating this project in 2016 (BBC, 2017). However, officials at GENeco—the enterprise which ran this project—believe the concept that emerged from this project can inspire other similar projects that can become successful since bio-methane gas powered buses have been found to reduce significant amount of air pollution in urban areas around the world (BBC, 2017).

Electric cars are now becoming popular in many countries around the world and in the United States, sales of electric cars rose to 37% in 2016 compared to the previous years and a 70% year-to-year increase in monthly sales (Rapier, 2017). This has made Hotel chains like the Marriot, Starwood and the Hilton to open Electric Vehicle (EV) charging stations in collaboration with Tesla (the largest electric automobile maker) and GE in many of their properties. This move by hotels to welcome electricity powered low emission vehicles also serves their strategic purpose of alluring high-end guests (Hanley, 2015). Such high-end guests who can afford the most expensive services offered by hotels also happen to have the ability to afford electric cars. Apart from hotels, car rental services are starting to offer electric car rentals in some parts of the world. The case of Enterprise Rent-A-Car electric car rental services in Orlando city exemplifies the use of electric cars in the travel industry. The electric car rental service named as Drive Electric Orlando (DEO) programme started in 2013 as a result of collaboration between Enterprise and Electrification Coalition and has grown considerably through the strength of tie-ups with hotels, theme parks and tourism promotional organizations like Visit Orlando (Padilla, 2017). The DEO programme offers opportunities to tourists for test drives of the latest versions of the electric car model—Chevy Volt, a hybrid electric car brand of Chevrolet (Hanley, 2015). Through such an opportunity, tourists are in turn provided the experience of new low emission vehicle technologies through car rental services and motivated to be owners of electric powered low emission vehicles in the near future (Padilla, 2017). The DEO programme is facilitated by 300 EV charging stations in different tourist centres of Orlando city and 40 partner hotels that provide free parking and valet services to the tourists who are clients of the program.

Green Energy Technologies

Green energy technologies are exemplified by the existence of energy conservation and cost-saving technologies primarily in the context of urban hotels. The casino destination of Las Vegas has demonstrated how major casino and convention facilities can be majorly dependent on renewable energy sources instead of conventional energy sources.

The hotel industry has also improvised its process of energy saving by installing digitally powered rooms and energy management systems that control and minimize energy consumption in guest rooms. Though this technology is not specific to urban hotels, its applicability has been possible more in such hotels. However, one of the best developments of green energy technology in the context of urban tourism has been in the case of Gardens by the Bay in Singapore.

Gardens by the Bay, Singapore

Gardens by the Bay opened in 2012 demonstrates a model of an iconic green urban attraction that involves the harnessing of a range of different renewable energy technologies like solar energy, biomass technologies and technologies providing the basis of sustainable design (Rowell, 2017). The most striking technological element of this attraction is the manmade supertrees. These supertrees ranging between 25–50 metres in height are built with a combination of green technology features—solar energy harvesting photovoltaic cells on the canopies as well as water harvesting inbuilt tanks that enable cooling of domes in the attraction containing 90,000 plants (Rowell, 2017). The supertrees themselves support 160,000 plants and the energy produced by the photovoltaic cells on their canopies are used for spectacular light, colour and sound displays from the attraction during nighttime (Rowell, 2017). The two domes or glass biomes in the attraction—the Flower dome and the Cloud Forest—are equipped also with sustainable features like green designing which allow sunlight to provide nourishment to the plants inside as well as rainwater capturing features that allow irrigation of the different flora species that they contain. The Gardens by the Bay attracts 25 million tourists every year from all over the world since it was opened in 2012 (Algie, 2014).

Gaming and entertainment companies in Las Vegas have moved away from conventional-energy-producing agencies like NV Energy. MGM Resorts International, one of the largest gaming and entertainment companies in Las Vegas exited from NV Energy paying an exit fee of $87 million (Hernandez, 2016). It is now considering a full-fledged transformation into a renewable energy agency in terms of fulfilling its energy needs by harnessing its phenomenal rooftop solar project as well as tie-ups with energy companies investing in solar energy as well as with solar photovoltaic developers and solar plants projects outside its premises (Hernandez, 2016). MGM developed rooftop solar projects in one of its principal Las Vegas projects—Mandalay Bay—which now has 26,000 solar panels spread across 28 acres and producing over 14 megawatts of energy. Other Las Vegas resorts like the Wynn Resorts and Las Vegas Sands are planning to buy more solar energy from the open market produced in surplus volumes in the solar farms in nearby places as well as produce renewable energy particularly solar energy themselves.

Conclusion

As the drive towards the enhancement of sustainable and smarter urban tourism grows (UNWTO, 2017), the connection between technology and sustainability will get further strengthened. On one hand, as the world moves towards the era of Industry 4.0, (technologically) smarter cities will receive more impetus entailing the growth of smart tourism destinations—driven by digital technologies, widespread availability of internet technology and smarter digital devices. On the other hand, as the climate crisis becomes increasingly significant, cities will have a larger role to play in the deployment of green tech and clean energy which will, in turn, become embedded in the functioning of major urban industries like tourism. Moreover, smart mobility options are being increasingly adopted by global cities in the world which focuses on offering efficient mass transit systems and clean energy/low emission automobile and inter-modal transport options to both tourists and local residents (Arcadis, 2017).

A summary diagram highlighting the interlinkages discussed in this chapter is presented in Fig. 16.1.

The World Economic Forum (WEF) considers innovation in technology to be the key for growth in the times to come principally in the areas of digitalization, smart systems (encompassing intelligent systems in clean energy, soft mobility and man-machine interfaces), artificial

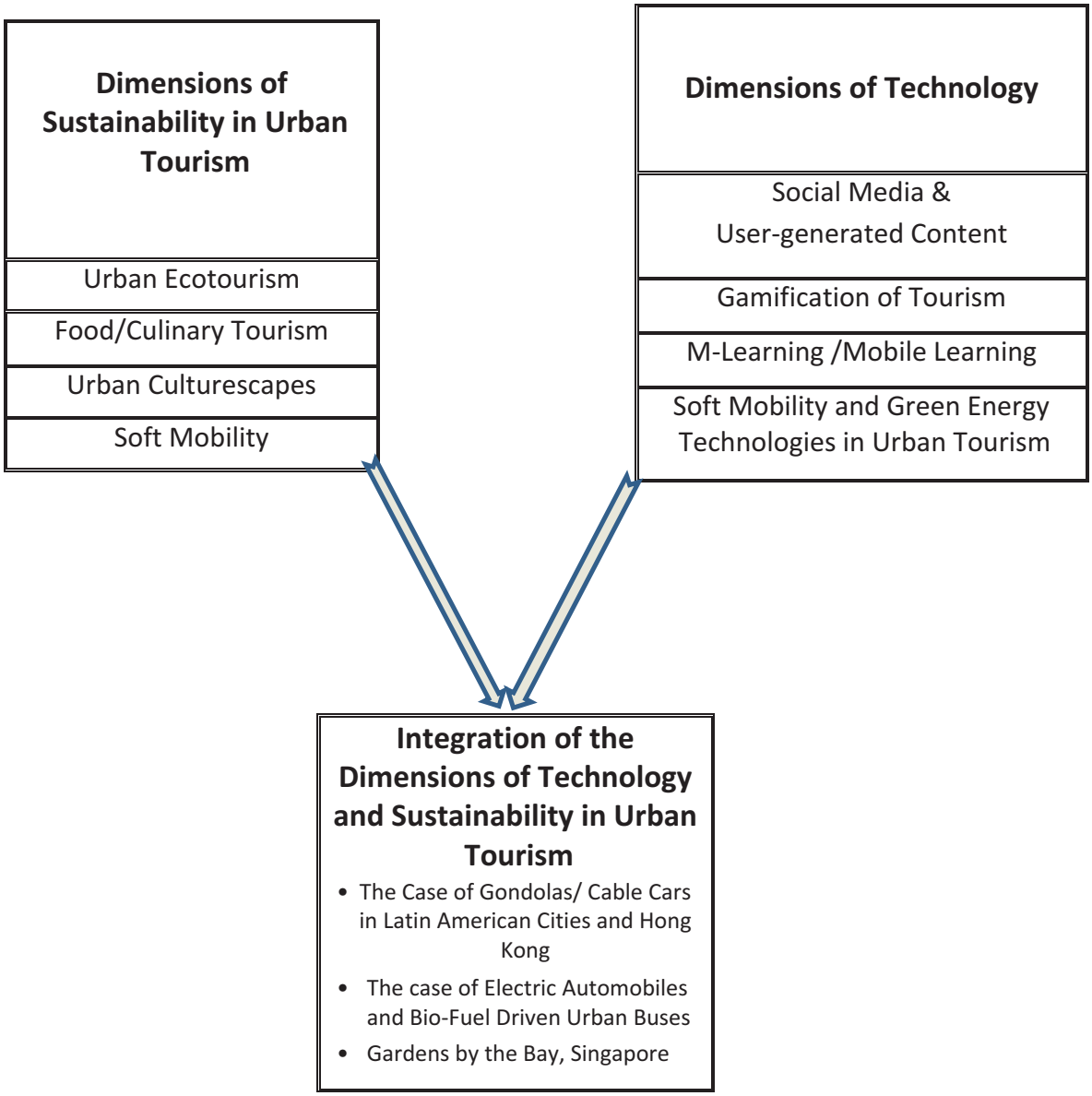


Fig. 16.1 Interlinkages between sustainability and technology in the urban tourism setting

intelligence and sustainable technologies (World Economic Forum, 2017). The UNWTO's call for a sustainable tourism for development in 2017 in the context encompasses social and economic inclusivity, cultural diversity, cultural mutualism and clean energy-driven environmental practice and (UNWTO, 2017). Therefore, the connection and integration of sustainability and technology in urban tourism contexts need to entail the social and cultural dimensions of both as their fundamental basis in addition to their environmental and innovational aspects.

The integration of sustainable forms of tourism and social and environmental dimensions of technology in urban contexts must yield benefits for all sections of visitors as well as all sections of the local resident population in urban destinations of the world.

One of the fundamental considerations for urban destination management organizations should be to put the sustainability-technology integration as the main underlying element in their destination marketing and branding process. Future studies focusing on the sustainability-technology integration may focus on the context of specific destinations—their scope in terms of sustainable forms of urban tourism attractions available as well as the available forms of digital and clean energy technology and their extent of use in such contexts.

References

Algie, J. (2014). Gardens by the Bay, Singapore: The coolest gardens in the world? Retrieved December 12, 2017, from https://jonistravelling.com/gardens-by-the-bay/.

Arcadis. (2017). *Sustainable Cities Mobility Index 2017: Bold moves*. Arcadis.

Barber, M (2017). 11 urban gondolas changing the way people move. Retrieved December 29, 2017, from https://www.curbed.com/2017/9/21/16340394/urban-gondolas-cable-cars-cities.

BBC. (2017). London buses to be powered by coffee. Retrieved December 21, 2017, from http://www.bbc.com/news/uk-england-london-42044852.

Bock, K. (2015). The changing nature of city tourism and its possible implications for the future of cities. *European Journal of Futures Research, 3*(20), 1–8.

Buffenstein, A. (2016). Icelandic music festival hosts impressive Avant-Garde Art Program. Retrieved December 13, 2017, from https://news.artnet.com/exhibitions/cycle-festival-2016-avant-garde-art-694697.

Buhalis, D., & Law, R. (2008). Progress in information technology and tourism management: 20 years on and 10 years after the internet—The state of etourism research. *Tourism Management, 29*(4), 609–623.

Burke, B. (2014). *Gamify: How gamification motivates people to do extraordinary things*. Brookline, MA: Bibliomotion.

Gebhardt, L., Krajzewicz, D., Oostendorp, R., Goletz, M., Greger, K., Klötzke, M., … Heinrichs, D. (2016). Intermodal urban mobility: Users, uses, and use cases. *Transportation Research Procedia, 14*, 1183–1192.

Geneco. (2016). Case study: Bio-Bus. Retrieved December 21, 2017, from http://www.geneco.uk.com/Case_study_bio_bus/.

Gillan, A. (2014). Taiwan, home to the best street food markets in the world. Retrieved December 30, 2017, from https://www.theguardian.com/travel/2014/may/17/taiwan-taipei-street-food-markets.

Hanharan, J. (2010). Ecotourism and sustainability in the tourism sector. In L. Leonard & J. Barry (Eds.), *Global ecological politics* (Advances in Ecopolitics) (Vol. 5, pp. 171–229). Emerald Group Publishing Limited.

Hanley, S. (2015). Hilton Hotels partners with Tesla and GE to add charging stations. Retrieved December 26, 2017, from https://ecomento.com/2015/10/12/hilton-hotels-tesla-ge-electric-car-charging-stations/.

Harris, B. (2017). Tesla's electric truck 'needs the energy of 4,000 homes to recharge. Retrieved December 28, 2017, from https://www.weforum.org/agenda/2017/12/tesla-s-electric-truck-needs-the-energy-of-4-000-homes-to-recharge-say-researchers//.

Hernandez, D. (2016). Las Vegas casinos seek to power their bright lights with renewable energy. Retrieved December 14, 2017, from https://www.theguardian.com/environment/2016/mar/07/las-vegas-casinos-solar-power-nevada-energy.

Herrera, C. F, Herranz, J. B., & Arilla, J. M. P. (2012). Gastronomy's importance in the development of tourism destinations in the world. United Nations World Tourism Organisation. Global report on Food Tourism (pp. 6–9). Madrid: UNWTO.

Hirtenstein, A. (2017). London's iconic red buses to run on coffee in bid to cut emissions. Retrieved December 29, 2017, from http://www.independent.co.uk/news/business/news/london-red-buses-run-coffee-biofuel-cut-emissions-vehicles-air-pollution-biobean-shell-a8064516.html.

Kaplan, A. M., & Haenlein, M. (2010). Users of the world, unite! The challenges and opportunities of social media. *Business Horizons, 53*(1), 59–68.

Kastelein, B. (2004). Urban ecotourism: Impossible conundrum? Mexican cities must clean up their acts to keep tourists. *Business Mexico*, pp. 36–42.

Kazakova, K. (2015). Tourism gamification examples and what is there for you? Retrieved December 22, 2017, from https://www.ojoo.com/tourism-gamification-examples-and-what-is-there-for-you/.

Kramer, M. (2009). Soft mobility—Measures for a climate-friendly transport policy in Europe. *The Greens | EFA in the European Parliament*. Brussels, Belgium.

Merrick, J (2011). Glasgow and Edinburgh: The architectural rivalry. Retrieved December 13, 2017, from http://www.independent.co.uk/arts-entertainment/architecture/glasgow-and-edinburgh-the-architectural-rivalry-2315854.html.

Negruşa, A. D., Toader, V., Sofică, A., Tutunea, M. F., & Rus, R. V. (2015). Exploring gamification techniques and applications for sustainable tourism. *Sustainability, 7*, 11160–11189. https://doi.org/10.3390/su70811160

Nualkhair, C. (2015). Bangkok's best street food: A guide to dishes and districts. Retrieved December 30, 2017, from https://www.theguardian.com/travel/2015/nov/25/bangkok-best-street-food-guide-dishes-districts-thailand.

OECD. (2012). Climate change. Retrieved July 15, 2019, from https://www.oecd.org/env/cc/Work-on-Climate-Change-2013-14_web.pdf.

Özdemir, G., & Çelebi, D. (2015). Reflections of destinations on social media. In V. Katsoni (Ed.), *Cultural tourism in a digital era* (Springer Proceedings in Business and Economics). Cham: Springer.

Padilla, M (2017). A plug for electric car rentals. Retrieved December 27, 2017, from http://www.travelweekly.com/North-America-Travel/Insights/A-plug-for-electric-car-rentals.

Peltier, D. (2015). The future of food tourism goes beyond the restaurant experience. Retrieved December 10, 2017, from https://skift.com/2015/10/23/the-future-of-food-tourism-goes-beyond-the-restaurant-experience/.

Peltier, D. (2017). Local food trend keeps farms at center of tourism strategies. Retrieved December 10, 2017, from, https://skift.com/2017/04/25/local-food-trend-keeps-farms-at-center-of-tourism-strategies/.

Rapier, R. (2017). U.S. electric vehicle sales soared in 2016. Retrieved December 20, 2017, from https://www.forbes.com/sites/rrapier/2017/02/05/u-s-electric-vehicle-sales-soared-in-2016/#697d2a22217f

Roberts, S., & Tribe, J. (2008). Sustainability indicators for small tourism enterprises—An exploratory perspective. *Journal of Sustainable Tourism, 16*(5), 575–594.

Rocca, R. A. L. (2009). Soft mobility and urban transformation: Some European case studies. *TeMALab Journal of Mobility, Land Use and Environment, 3*, 85–80.

Rokka, J., & Moisander, J. (2009). Environmental dialogue in online communities: Negotiating ecological citizenship among travellers. *International Journal of Consumer Studies, 33*(2), 199–205.

Rowell, C. (2017). Gardens by the Bay, Singapore, breaks new ground in sustainable building. Retrieved December 17, 2017, from http://www.construction-global.com/major-projects/gardens-bay-singapore-breaks-new-ground-sustainable-building.

Saarinen, J. (2006). Traditions of sustainability in tourism studies. *Annals of Tourism Research, 33*(4), 1121–1140.

Santos, S. (2011). 2012 social media and tourism industry statistics. Retrieved December 22, 2012, from http://www.stikkymedia.com/blog/2012-social-media-and-tourism-industry-statistics.

Sarkar, S. K., Au, N., & Law, R. (2013). Analyzing ecotourists' satisfaction in socialization and knowledge sharing intentions via social media. In *Information and communication technologies in tourism* (pp. 313–326). Springer International Publishing.

Seraj, M., & Ayesugul, T. (2012). Social network citizenship. In M. M. Cruz-Canha, P. Goncalves, N. Lopez, E. M. Miranda, & G. D. Putnik (Eds.), *Handbook of research on business social networking: Organizational, managerial and technological dimensions* (pp. 339–357). IGI Global.

Sever, N. S., Sever, G. S., & Kuhzady, S. (2015). The evaluation of potentials of gamification in tourism marketing communication. *International Journal of Academic Research in Business and Social Sciences, 5*(10), 188–202.

Siemens AG. (2013). Annual report on sustainability. Retrieved July 15, 2019, from https://www.siemens.com/annual/13/en/download/pdf/Siemens_AR2013.pdf.

Tate, Z. (2017). Capitalism is losing support. It is time for a new deal. Retrieved December 30, 2017, from https://www.weforum.org/agenda/2017/11/capitalism-losing-support-we-need-a-new-deal/.

Transdanube. (2014). *Transnational soft mobility and tourism marketing strategy for transdanube regions*. Danube Competence Center.

UNWTO. (2004). Sustainable development of tourism. Retrieved December 26, 2017, from http://sdt.unwto.org/content/about-us-5.

UNWTO. (2012). Global report on city tourism—Cities 2012 Project. UNWTO, Madrid.

UNWTO. (2017). The tourism sector highlights the potential of urban tourism and the need to move toward more sustainable practices. Retrieved January 5, 2018, from http://media.unwto.org/press-release/2017-05-12/tourism-sector-highlights-potential-urban-tourism-and-need-move-toward-more.

WCED. (1987). *Our common future*. World Commission on Environment and Development. Oxford: Oxford University Press.

WEF. (2017). Technology and innovation for the future of production: Accelerating value creation. *World Economic Forum White Paper*.

Index[1]

A
Aadhaar, 163, 163n2, 165
Analytics, 5, 8, 15, 18, 20, 25, 40, 45, 46, 48, 50, 51, 103–111, 146, 150–152, 161, 164, 167, 192, 207, 231, 241, 242, 244
Artificial intelligence, 3, 104, 127–138, 211–224, 237, 239–241, 260

B
Banking, 6, 16, 156, 157, 165, 230–244
Barriers, 7, 21, 84, 192, 200, 204, 253
Behavior, 49, 50, 107, 110, 114, 123, 146, 149–151, 164, 169–183, 219, 220, 224, 239, 248, 252–255, 261, 263–265, 267
Beneficiaries, 161, 165, 166
Business, 1, 16, 40, 62, 84, 104, 113, 128, 145, 164, 170, 211, 229, 247, 274
Business functions, 118
Business-to-business (B2B), 104

C
Cabibihan, J.J., 128
Capabilities, 26, 28, 32, 41, 53, 54, 67, 76, 85, 87, 92–93, 96, 104, 130–132, 135, 145, 176, 179, 213, 219, 222, 223, 249–250, 255, 258–260, 265, 266
Case study, 161–166, 192, 193, 195, 200–204, 206, 250, 251, 255
Cloud, 40, 51, 118
Code, 23, 33, 192, 231–236, 241, 244, 260
Codification, 46

[1] Note: Page numbers followed by 'n' refer to notes.

© The Author(s) 2020
B. George, J. Paul (eds.), *Digital Transformation in Business and Society*, https://doi.org/10.1007/978-3-030-08277-2

Cognition, 121, 219, 252
Consumer, 2, 4, 7, 24, 31, 40, 50, 51, 62, 64–73, 75, 84–86, 88, 89, 107, 119, 120, 127, 128, 131, 134, 137, 145–152, 163, 169–183, 213, 217, 239, 274, 284
Crimes, 8, 27, 109, 239, 287, 288
Cultural, 15, 17, 113, 133, 194, 252, 254, 266, 276–280, 282, 284–286, 292
Cyber, 6, 240
Cyberattack, 25
Cybercrimes, 239
Cybersecurity, 110, 223, 224
Cyberspace, 8, 65

D
Data, 1, 15, 40, 74, 94, 104, 120, 129, 145, 158, 179, 191, 218, 229, 258
Database, 39, 44, 46, 47, 161
Data-capture, 179
Data-driven, 4, 6, 45, 54, 113, 273–293
Debit, 156
Decentralised, 90, 91
Decision-making, 6, 15, 20, 27, 39, 44–48, 52–54, 91, 171, 175, 183, 259, 260, 275, 283, 284
Differentiate, 17
Diffusion, 66, 176
Digitalization, 120, 129, 131, 177, 275, 291
Digital marketing, 7, 170–173, 176, 182, 183
Digital natives, 3–5, 113–123
Digitize, 40, 165
Dimension, 3, 8, 91, 113, 157, 194–196, 199, 219, 259, 274, 276–277, 280, 283, 292, 293
Disruptive, 53, 109, 267
Drivers, 1, 122, 129, 150, 176, 183, 192, 200, 204

E
Eco-friendly, 163, 282
Ecology, 4, 278
E-commerce, 40, 164, 170
Economics, 5, 14–16, 18–20, 22, 24, 26–28, 31, 51–53, 86, 108, 129, 130–132, 137, 157, 158, 164, 166, 167, 169, 178, 192, 194, 200, 204, 217, 230, 232, 234, 274, 275, 286, 292
Ecosystem, 40, 41, 241, 242
E-customer, 173–183
E-government, 14, 32, 191
E-mail, 107, 115, 176, 179
Emergence, 41, 66, 84, 171
E-money, 156, 158
End-users, 30, 241
Entrepreneurship, 86, 130
Environment, 17–19, 28–30, 43, 50, 53, 105, 107, 111, 114, 118, 121, 133, 157, 158, 163, 169, 170, 172, 173, 175–177, 182, 213–214, 218, 230, 232, 248, 259, 263, 264, 266, 267, 274, 276, 277, 279, 285
E-purchasing, 176
Export, 22, 83–97

F
Flourishing, 13–33
Functionality, 25, 54, 255
Fundamentals, 1, 45, 178, 238, 254, 292, 293
Future, 3, 8, 14, 21–23, 25, 29–31, 51–54, 76–77, 104, 105, 107–110, 120, 123, 128–130, 132, 133, 135, 137, 150, 166, 167, 181, 182, 206, 207, 211–224, 230, 247, 248, 251, 255, 261, 267, 273–275, 289, 293

G
Gadget, 116
Gallup, 21
Governance, 6, 16–18, 20, 27, 28, 31, 32, 90, 91, 231–236, 238, 240, 244

H
Healthcare, 16, 49, 106, 109, 150
Holograms, 259
Hospital, 213
Hospitality, 5, 128, 211–224, 275, 286
Human-centered, 30, 32, 120
Humanoid, 215
Hyper-connected, 113

J
Jobs, 5, 31, 72, 118, 129, 130, 132–137, 217, 218, 223, 264
Just-in-time, 1

K
Kiosk, 215, 219
Know-how, 30, 31, 114, 151
Knowledge-driven, 39

L
Law, 6, 68, 110, 134, 231
Leadership, 7, 19, 53, 120, 247–267
Linear, 158, 159, 213
Logic, 18, 181
Logistics, 41, 84, 175, 214, 217

M
Mainframe, 103, 237
Mankind, 8
Mapping, 109, 206, 258
Marketing, 7, 15, 61–77, 83–97, 104, 147, 164, 170–173, 223, 240, 284, 285, 293
Matrix, 53, 93

N
Natives, 3–5, 113–123
Negotiation, 53, 122
Networks, 16, 17, 26, 27, 69, 70, 73, 75, 83, 88, 89, 91, 96, 109, 110, 131, 156, 157, 162, 174, 180, 194, 201, 207, 260, 276, 281, 282, 286

O
Openness, 252
Organisation, 4–6, 8, 39–51, 53, 54, 62, 67, 69, 71–75, 90, 92, 108, 114–118, 120, 122, 123, 155, 194, 200, 214, 216, 218, 229, 230, 233, 235, 237, 240, 247, 248, 255, 266, 274, 276, 284, 289, 293

P
Package, 201, 239
Photo-sharing, 283
Physical, 8, 30, 85, 86, 93, 95, 97, 133, 135, 136, 157, 183, 213, 222, 224, 237, 240, 258, 265
Pollution, 280, 289
Prediction, 109, 217, 240, 241, 259
Privacy, 8, 21, 25, 28, 31, 105–107, 111, 121, 145–152, 194, 201
Project, 5, 23, 25, 32, 88, 104, 116, 118, 161, 165, 181, 247–267, 274, 275, 279, 287–289, 291
Psychology, 76, 253

R
Railroads, 64
Real-data, 15, 20, 32
Research, 15–17, 19, 21, 23, 31, 42, 54, 76–77, 85, 86, 89, 90, 95–97, 104, 105, 116, 129, 156, 162, 164, 167, 192–199, 206–207, 217, 219, 222–224, 232, 233, 249, 250, 253
Retool, 135
Risk, 6, 18, 106, 116, 122, 134, 150, 156, 170, 212, 222–223, 229–244, 258
Robotics, 5, 8, 127–138, 211–224, 237

S
Scalability, 222
Scans, 116, 121, 204
Scenario, 25, 39, 40, 54, 115, 129, 213, 222
Security, 8, 16, 24, 27, 29, 32, 107, 145–152, 259, 260
Smart cities, 4, 13–33
Social media, 4, 40, 61–77, 83–97, 108, 110, 116, 119, 146, 172, 276, 283–284
Society, 1, 2, 5, 8, 17, 20, 21, 27, 33, 128, 129, 132, 138, 145, 146, 150, 157, 237, 255, 266, 274, 278, 279, 286
Stakeholder, 5, 6, 28, 41, 84, 89–91, 115, 123, 129, 155, 167, 171, 181, 191, 192, 194, 195, 200, 204, 206, 207, 233–235, 239, 260, 265
Sustainability, 14, 16, 17, 19, 30, 132, 206, 260, 274–276, 278, 280, 283, 286, 291, 292

T
Telepresence, 222, 224
Tesla, 275, 289
Theory, 68, 73, 85–87, 93, 95, 106, 130, 178, 230
Tourism, 128, 212, 214, 221, 274–280, 282–286, 289–293
Transformation, 2–5, 7, 40–41, 45, 53, 65, 114, 115, 120, 122, 123, 128, 161, 170, 171, 173, 177, 180, 182, 240, 248, 255, 258, 266, 274, 286, 291
Two-dimensional, 260

U
Uber, 16, 62, 119, 148, 172, 281
Upskilled, 133
Urban, 4, 15–20, 24, 25, 27–31, 33, 116, 275–293
User-generated, 67, 88, 171, 283–284
Utilisation, 15, 85, 235, 279
Utopia, 22

V
Visualisation, 122, 192, 201, 206, 242, 259

W
Wearables, 106, 258, 261, 265, 266
Weather, 20, 191, 260
Website, 6, 22, 50, 63, 64, 66, 67, 69, 84, 88, 109, 115, 146, 149, 152, 164, 174, 181, 192, 195, 196, 198, 199, 201, 202, 277
Wireless, 18

X
Xbox, 62, 72

Y
Yammer, 116

Australasian Accounting, Business and Finance Journal

Volume 13
Issue 2 *Women in Business in India*

Article 4

Do Women on Boards affect Firm's Financial Performance? Evidence from Indian IPO Firms

Amit Kumar Singh
University of Delhi, India, amitipo10@gmail.com

Shubham Singhania
Delhi Technological University, India.

Varda Sardana
University of Delhi, India.

Follow this and additional works at: https://ro.uow.edu.au/aabfj
Copyright ©2019 Australasian Accounting Business and Finance Journal and Authors.

Recommended Citation
Singh, Amit Kumar; Singhania, Shubham; and Sardana, Varda, Do Women on Boards affect Firm's Financial Performance? Evidence from Indian IPO Firms, *Australasian Accounting, Business and Finance Journal*, 13(2), 2019, 53-68. doi:10.14453/aabfj.v13i2.4

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au

Do Women on Boards affect Firm's Financial Performance? Evidence from Indian IPO Firms

Abstract
The objective of this paper is to study whether the presence of women on boards affects the financial performance of firms. This study builds upon other studies that have earlier tried to determine the impact of number of women directors in Indian corporates on their financial performances, in terms of market performance index, i.e. Tobin's Q. The new Companies Act, 2013 has mandated the appointment of at least one woman director in certain classes of listed companies in India as described in Section 149(1)(b), which could possibly bring about a change in the governance and thus the financial performance of firms. The focus of this study is on finding and analysing the impact of compliance with this provision on the financial performance of IPO firms. In order to test our hypotheses, we selected 41 Indian Companies that have made an Initial Public Offer (IPO) in the recent past and have been listed on Bombay Stock Exchange during the period 2012-2016. The performance of each of these firms has been measured in terms of Tobin's Q for 3 successive years after their listing (IPO) as well as by using two gender diversity index values- Blau Index and Shannon index. The study concluded that the proportion of women on board is insignificant and that it could not impact the financial performance of the firms. Neither has it caused any adverse impact on the performance of these firms.

This article is available in Australasian Accounting, Business and Finance Journal: https://ro.uow.edu.au/aabfj/vol13/iss2/4

Do Women on Boards affect Firm's Financial Performance? Evidence from Indian IPO Firms

Amit Kumar Singh[1], Shubham Singhania[2] and Varda Sardana[3]

Abstract

The objective of this paper is to study whether the presence of women on boards affects the financial performance of firms. This study builds upon other studies that have earlier tried to determine the impact of number of women directors in Indian corporates on their financial performances, in terms of market performance index, i.e. Tobin's Q. The new Companies Act, 2013 has mandated the appointment of at least one woman director in certain classes of listed companies in India as described in Section 149(1)(b), which could possibly bring about a change in the governance and thus the financial performance of firms. The focus of this study is on finding and analysing the impact of compliance with this provision on the financial performance of IPO firms. In order to test our hypotheses, we selected 41 Indian Companies that have made an Initial Public Offer (IPO) in the recent past and have been listed on Bombay Stock Exchange during the period 2012-2016. The performance of each of these firms has been measured in terms of Tobin's Q for 3 successive years after their listing (IPO) as well as by using two gender diversity index values- Blau Index and Shannon index. The study concluded that the proportion of women on board is insignificant and that it could not impact the financial performance of the firms. Neither has it caused any adverse impact on the performance of these firms.

JEL classification: G38, M14, J15, J16

Keywords: IPO Firms, Women Directors, Tobin's Q, Financial Performance, Gender Diversity.

[1] Associate Professor, Department of Commerce, Delhi School of Economics, University of Delhi, India. Email id: amitipo10@gmail.com
[2] Research Scholar, USME, East Campus, Delhi Technological University, India.
[3] Research Scholar, Department of Commerce, Delhi School of Economics, University of Delhi, India

1. INTRODUCTION

The subject of women on boards is a new and developing area of research. Many researchers have been carrying out studies to find out the impact of gender diversity on the performance of companies. This area of research has gained widespread interest after the institution of legal norms regarding the appointment of women directors on board in firms across the globe, beginning with the European corporate sector. It was in 2013, when the Companies Act came into force after revamping the old Companies Act of 1956, that the representation of women on board was paid heed to in India. It mandated the appointment of at least one woman director on boards of all the Indian listed companies and those public companies in India having a paid up capital of hundred crores or more and a turnover of three hundred crores or more in any year.

The reason for introducing this legal requirement was the low average numbers of women on the boards of Indian corporations, which is quite contrary to the principles of equality that the Indian constitution emphasizes. With the reservation of board seats for women in the European counterparts, the Ministry of Corporate Affairs, realizing the need of the hour and the importance of Indian women with respect to their leadership and managerial expertise, reserved a mandatory position for women on boards in the corporate sector.

The Gender Diversity Index reveals that in the year 2017, there has been a substantial rise in the number of board seats occupied by women in Fortune 1000 companies. It has been observed that out of 801 active Fortune 1000 companies, 55% of the companies have achieved the status of 'W' Companies, with more than 20% of their board seats held by women. The 'Z' companies, that is, those with no women directors, fell to an all time low of 7% in 2017. However, a high incidence of tokenism is observed in the Indian scenario with just one seat reserved for women on boards as against 40% reservation in Norway and Germany. If we consider the representation of women in the new firms coming up with their IPO, women hold just 9.2% of the board seats in the largest 25 IPO companies in 2017, which has increased from 8.2% of board seats in 2016.

This paper looks into the impact that women on boards could have, on the financial performance of the firms. In this respect, we have chosen 41 BSE listed Indian companies that have made Initial Public Offers (IPOs) during the period 2012-2016. We calculated the proportion of female directors in these firms to analyse the impact of gender diversity on board on the firm performance, calculated in terms of Tobin's Q, which is a proxy indicator for market performance for the 3 successive years immediately after their IPOs.

Many researchers have investigated the impact that the presence of women creates on corporate returns. Previous research has actually produced mixed results, given that the link between gender diversity and financial performance in a firm is both theoretically and empirically complex. However, most existing studies have used data from western countries rather than the developing nations. The management and leadership styles develop in synchronisation with the culture, which has a long lasting influence on the firm's financial health. This paper uses evidence from India, a developing country with a male dominated culture, like most of the Asian countries.

Our investigations focus on finding out how the presence of women directors on board would change the financial health after the institutionalisation of the women reservation norms by Companies Act 2013 under Section 149(1)(b) and Securities Exchange Board of India in Revised Clause 49 of the listing agreements. We have used updated data of the firms that

have stepped into the capital market by making new issue of shares, after the provisions of Companies Act 2013 came into effect.

The paper has been divided into five sections. Section 1 and Section 2 pertains to the introduction and literature review respectively. Section 3 deals with research gap, objectives and hypotheses. Research design is covered under Section 4 wherein the methods adopted, variables chosen and sources of data have been explained. Section 5 highlights the results and analysis. We conclude our paper with Section 6, highlighting the areas and Scope for future research and limitations of the study.

2. LITERATURE REVIEW

The corporate governance mechanism is something that has been a focus of attention in many countries in recent times. Corporate governance norms play a large role in building up the financial performance of any business concern. Lately, board gender diversity has been one of the most talked about aspect of corporate governance. It has been observed that the gender composition of the board may influence the performance of the company. It is time that productive resolutions are taken at the government and company levels for board level gender equilibrium.

2.1 Performance of IPOs

Singh and Maurya (2018) in their Indian study have shown that IPOs on National Stock Exchange (NSE) posted a very small gain of 4.63%, in short run, calculated on one month basis. About 64.21% of IPOs introduced on NSE were found to be under-priced and the rest were concluded to be overpriced. Also, except 1 firm, the issue price and the listing price was also found to be the same. Jain and Kini (1994), while investigating post IPO performance, noticed that they incurred a significant level of decline in post IPO performance, causing equity retention. It can also be said that the markets usually overreact to the news of an IPO and then slowly stabilises to normal. Coakley et al. (2007) also highlighted huge underperformance of the initial public offerings in the bubble years 1998-2000, while later the IPO performances have been more or less normal. Murthy and Singh (2016) show that for IPOs listed from the year 2000 to 2003, the long run variables have no relationship with short-run variables. Moreover, in the context of Indian new issues market, Singh and Maurya (2018) in their study concluded that it is the presence of independent directors, rather than non-executive directors, that provides for relevant information with respect to the post IPO returns.

Krishnan et al. (2011) noted a positive relationship between the image and goodwill of VC-backed IPOs and their long run performance. As per them, reputable venture capitalists are more actively involved in the governance of firms, post IPO, and thus are likely to influence long run performance positively in post IPO period. It has been found that market overreacts to the initial public offers, as is proved by the findings that majority of the IPOs have significant returns on the day of listing, but beyond that they do not give much return (Murthy & Singh, 2014).

2.2 Women and Board Diversity

Diversity refers to a variety of different characteristics put together in a common place. When it comes to board of directors, diversity can be in terms of ethnicity, gender, age, or the like. In a male dominated country like India, not much has been said about women representation

on the board of directors or top management positions in a company, until the onset of Companies Act 2013. Das and Dey (2016) have mentioned in their study that the number of women present on the board is indicative of firm's ethical behaviour and a balance on the board. Indian companies should learn how to build competitive advantage by putting up women on their board, thereby ensuring good reputation amongst the stakeholders (Kaur & Singh, 2017) which will ultimately lead to better financial health of the firms. Studies prove that women empowerment is likely to be advantageous for the corporations owing to the synergies brought in by gender diversity (Paul, 2017). Sanan (2016) reveals in her study that the number of companies with no independent women directors has reduced over time, which is good news as far as efficient corporate governance is concerned. But whether or not it is likely to improve corporate performance is a question that still needs to be answered.

2.3 Gender Diversity on Board and Positive Firm Performance

Taking into consideration the all-round benefits of gender diversity on boards, many researchers have come up with the idea that women on board have a positive impact on financial performance. Luckerath and Rovers (2013) believe that the presence of women on board will boost firm performance. Companies with women directors tend to be more ethical when it comes to dealing with their stakeholders, thus ensuring high firm value and better performance. Carter et al. (2003) also found a positive and direct relationship between firm value and fraction of women on board. In the Indian scenario, women have always faced challenges with respect to their position and appointment in the business organisations, primarily because of the glass ceiling that confronts them. Realising the importance, managerial and leadership abilities of women, there has been institutionalisation of norms across the globe with regards to appointment of women on board. These norms have forced the appointment of women on boards, and firms have given better results after such appointments. Adams and Ferreira (2008) have stated that gender diversity in top management is likely to bring a forward push in the financial health of the firms with weaker shareholder rights.

In a global study undertaken by Terjesen et al. (2016), it was found that firms having a higher proportion of female directors on board are going to have better Tobin's Q and return on assets as compared to those which have none. Another investigation that was conducted with respect to Indian context, in light of the mandated representation of women on board, also showed positive results on firm performance (Sanan, 2016). Studies have been conducted across various sectors and industries in India to find out the implications of one woman on board norm. Das and Dey (2016), while investigating large Indian corporations for the role of corporate governance in firm performance, have come across positive impact of board diversity on financial performance. However, the performance of the corporates may respond differently to board diversity in different sectors and different methodology used for analysis.

2.4 The Negative Implications of Gender Diversity on Board

While going through the available literature, we have come across some instances where gender diversity on board has worsened off the financials than before. In their study of FTSE 100 companies, Ryan and Haslam (2005) found that during periods of declining stock markets, companies who appointed women directors experienced worse performance than those with male directors. This trend has been experienced globally. McGuinness (2017), as per investigations conducted in Hong Kong, also envisages the idea of gender bias imposing a cost on stakeholders and thus affecting performance negatively. A similar study conducted in Norway by Ahern and Dittmar (2012), suggests that in companies with quota reserved for women on board, there is a decline in Tobin's Q over a period of time. Where on the one

hand, many authors have supported the idea of appointment of women on board, claiming that diversity would help in better decision making and problem solving, it has also been observed that diversity brings with it more chances of conflicts and makes decision making time consuming (Gallego et al., 2010). Wellalage and Locke (2013), in a study conducted in Sri Lankan context, have found that board heterogeneity has a negative impact due to increased chances of conflicts. In a country like India, where there are many family businesses running and operating since years, the mandate of appointing women directors has been complied with by putting the female relatives or family members on the boards.

A lot has been said about the positive and negative implications of the presence of women on board on corporate financial performance. A few authors have envisaged a direct relation between the two variables stating that gender diversity leads to more opinions and better decision making owing to different attributes of men and women. Better decision making leads to problem solving and thus, better performance. There are studies which show that the relationship between the two variables is negative, owing to more chances of conflicts and delays in working.

2.5 Women on Board and No Impact on performance

Noland et al. (2016), in a global study of 21,980 firms, found no impact of the gender quota on board on a firm's financial performance. Many studies have also been carried out in developing countries with the same results. A study conducted by Yasser (2012) in Pakistan showed no impact of gender diversity of board on the financials of the companies listed on the Karachi Stock Exchange. Gallego et al. (2010), while carrying out research in the context of Spain, have also concluded that high level of gender diversity may or may not influence performance.

However, the existing literature has raised questions on the new provision of Companies Act by calling it a mere tokenism, with just one independent woman director present on boards not making a significant difference. The number of companies with 2 or more women directors is substantially less (Sanan, 2016). In order to achieve better performance, Joshi (2017) concluded that there is a need for Indian companies to exploit the opportunity to build a strong talent pool of diverse directors.

3. RESEARCH GAP

The existing literature brings to light various instances where the presence of women brought positive results, negative results and there were instances with no impact on performance at all. The investigations conducted do not give a uniform result to throw light on the effect of gender diversity of board on the firm performance. Moreover, it is imperative to seek answer to this pertinent question whether the addition of a mandatory provision has led to any improvement or change in the performance or not. Firms coming out with an IPO in India are required to disclose their managerial board in their prospectus issued to the shareholders. Hence, it would be both essential as well as interesting to see whether having a diverse board will send a positive signal to investors so that the firms can take better informed decisions owing to access to diverse viewpoints. The impact of such perceptions would certainly be easy to measure in terms of the short term performance. In our study, we have examined 41 publicly listed IPO firms for their compliance with provision for appointment of women on board and how that impacts their performance. We seek to find answer to the question whether presence of women on board is just in letter or in spirit too, which will help firms take future decisions effectively.

4. OBJECTIVES

- To analyse the impact of the presence of women directors on boards on the short term performance of firms.
- To analyse the impact of gender diversity of boards on the short term financial performance of firms.
- To anlayse the impact of return on assets and leverage on the financial health of the firms.

5. HYPOTHESES

H_{o1}: The percentage of women on board has no effect on the short term firm performance.
H_{a1}: The percentage of women on board affects the short term firm performance.
H_{o2}: The size of a firm has no effect on the short term firm performance.
H_{a2}: The size of a firm affects the short term firm performance.
H_{o3}: Leverage in a firm has no effect on the short term firm performance.
H_{a3}: Leverage in a firm affects the short term firm performance.
H_{o4}: Return on assets of a firm has no effect on the short term firm performance.
H_{a4}: Return on assets of a firm affects the short term firm performance.

6. RESEARCH DESIGN

6.1 Sample size and Data sources

The data includes a secondary sample of 41 IPO firms listed on Bombay Stock Exchange, selected on the basis of purposive sampling technique, to investigate the impact of having women directors on the board of the firms and how they affect the performance of these firms. All the selected firms have made an initial public offer (IPO) during the period 2012-2016. Their performance for 3 financial years has been taken into consideration after the IPO listing.

The above sample has been chosen since the bill relating to the mandatory provision for the appointment of one woman on board was already in place and firms had started complying with the same to protect themselves from any unnecessary penal action. Hence, the purpose of selecting such new IPO companies is that when a firm goes public, there are various factors which are taken into consideration while making initial investments and one such element is the ability of the board to take well informed decisions.

We believe that IPO firms are the best form of sample to be chosen, as they would adequately reflect the impact of such diversity on board and its implications on financial performance. Hence, we tried to analyse how the appointment of at least one woman director as per the Companies Act 2013 affects the financial performance of firms after their listings (Welbourne, 1999). Their presence is likely to give a lot of signals to the prospective investors with regards to the functioning as well as other aspects of the firms.

The data about the IPO firms i.e. issue price, opening and closing date as well as other information related to the IPOs has been retrieved from the Chittorgarh website. The corporate governance variables have been taken from the corporate governance reports of these firms and the financial variables have been retrieved from the prowess database and cross checked with the annual reports of these companies.

6.2 Variables

Dependent Variable

Financial Performance is the major element that incentivises the investors to make investments into a firm in the expectations of gaining good returns, and it is on this basis that the performance of the board of directors is also evaluated. We measured financial performance by using a market based measure rather than the accounting based measure, since the accounting based measures are based on the past results and might not reflect the changes in the governance norms immediately (Haldar et al., 2015).

Tobin's Q

Tobin's Q is considered as one of the best market based measures due to multiple reasons: Firstly, it takes into account the risk factor, unlike other accounting measures, and is not subject to distortions. Secondly, it reflects the market's expectations of future earnings and is, thus, a good proxy for a firm's competitive advantage (Montgomery & Wernerfelt, 1988). Hence, Tobin's Q was used as a proxy for measuring the performance, based on the work of Ahern and Dittmar (2012), Haldar et al. (2015), and Campbell et al. (2008) for various years. Tobin's Q is calculated using the Chung and Pruitt (1994) formula, which is, market valuation of shares plus debt divided by the total assets. A unity value for the Tobin's Q ratio acts as a significant benchmark for the measurement of firm performance. All the firms having a value greater than 1 are construed to give better returns to the investors as compared to the firms having value less than 1. It is also a sine quo nun for efficient utilization of resources.

Independent Variables

Table I provides a brief description of the independent variables used in the study.

Table I: Independent variables

S.No	Independent Variable	Definition	Symbols
1	P-Women	Percentage of women directors on board with respect to the total directors.	P-Women
2	Blau Index	A measure of gender diversity.	SHANNON
3	Shannon Index	A measure of gender diversity.	BLAU
4	Return on Asset	Percentage of profits earned against the total assets employed by a firm.	ROA
5	Leverage	Computed as the ratio of total long term debt to the total assets.	LEVER
6	Firm Size	Measured in terms of the natural log of the total assets.	F-SIZE

Percentage of women on board (P-woman)

P-woman represents percentage of women on board with respect to the total number of directors. This variable is a proxy for measuring gender diversity at the board level. Many studies in the past have confirmed the benefits of having a diverse board, in terms of a positive relationship between women on board and organizational innovation, better decision making, different perspectives etc. (Torchia et al., 2018). Diversity on board provides various opportunities in terms of strategic alliances that firms can build, based on its relationships with other groups (Halder et al., 2015). Hence, P-woman has been used based on several

studies by Campbell and Mínguez-Vera (2008), Giraldez and Berenguer (2018), Welbourne(1999), and Halder et al. (2018).

Blau Index

The Blau Index, that has been used as a proxy for women directors on boards, is a Gibbs-Martin index of sociology, psychology and management studies which is calculated using the formula $1-\sum_{i=1}^{n} p_i^2$ where p denotes the proportion of board directors in each male and female category and n symbolises the total number of directors on board. The variable has been used to enhance the robustness of the study with respect to women on board (Campbell & Mínguez-Vera, 2008).

Shannon Index

The Shannon Index, a reputed diversity index found in the literature, aims to state in quantitative terms, the uncertainty in predicting the species identity of any person chosen randomly from a given set of data. It is calculated using the formula: $-\sum_{i=1}^{n} p_i \ln(p_i)$. This variable has also been chosen to enhance the robustness of the study in terms of gender diversity on board (Campbell & Mínguez-Vera, 2008). Here, p denotes the proportion of board directors in each male and female category and n symbolises the total number of directors on board, similar to the Blau index.

In our study, women directors on boards of companies has been selected as an independent variable, and amongst the independent variables, we have applied Blau Index and Shannon Index, which shall take into account the number of categories in terms of gender as well as the evenness in terms of the distribution of board members amongst them. These attributes can also be said to be the measure of 'variety' and 'balance' (Stirling, 1998).

Control Variables

Return on Assets (ROA)

Return on assets has been used as one of the control variables in our study, and has been measured as a percentage of profits earned against the total assets employed in a firm. Return on assets is one of the operational-based measures of financial performance and, thus, is likely to affect the market based performance of a firm (Alghifari et al., 2013; Campbell & Mínguez-Vera, 2008).

Leverage (LEVER)

The leverage of a firm is computed as the ratio of total long term debt to the total assets, indicating what percentage of total assets of a firm is financed by debt. Leverage of the firms affects the ability of the managers to take decisions, since the creditors would interfere in the working of the firm based on their own interests. (Chen & Jaggi, 2000; Hutchinson & Gul, 2004).

Firm Size

Firm size has been measured with respect to the natural log of the total assets. It has been observed in the literature that the success of women on board differs based on the size of the firm, and women achieve great success in smaller firms. This was later contradicted by Sealy et al. (2008). It is also imperative to take firm size into consideration because the firm financial performance is likely to be affected by the size of the firm, as large firms have larger abilities to influence performance (Short & Keasy, 1999). Hence, based on several

studies (Haldar et al., 2015; Campbell & Mínguez-Vera, 2008; Yasser et al., 2017), it was included as a control variable.

6.3 Empirical Model Specifications

We have used a balanced panel data in our study that provides various advantages as compared to the time series data or the cross section data. Firstly, panel data offers more flexibility in forecasting the behaviour of the cross sectional units as compared to the time series data analysis (Greene, 2008). Secondly, a panel framework allows for analysis of a larger number of observations as compared to the individual time series for comparison purposes.

Panel data is a methodology that is often used to eliminate any unobservable heterogeneity that might be present in the sample for the study. However, if any unobservable heterogeneity is found to be correlated with the explanatory variables, we undertake a conditional inference (that is, estimation by fixed effects). However, if the effects are not correlated with the independent variables, we instead carry out unconditional inference using the random effects method. We have used the Hausman test (Hausman & Taylor, 1981) to check the presence of any correlation between unobservable heterogeneity and the explanatory variables (Ahern & Dittmar, 2012). We found that the unobservable effect was not correlated with the explanatory variables; hence random effect model was used in running the OLS regression.

Three different ordinary least squares (OLS) regression analyses were run using the dependent variable as Tobin's Q and all the control variables, by changing the independent variables in each model as P-woman, Blau Index and Shannon index. Since we have used three proxy variables, i.e. P-woman, Blau Index and Shannon Index, three OLS regression models were developed and the results were obtained for each of them.

The economic model used is specified below-

Model 1: $Qit = \beta 0 + \Sigma \beta 1 \text{P-WOMAN}it + \Sigma \beta 2\ \text{LEVER}it + \Sigma \beta 3\ \text{ROA}it + \Sigma \beta 4\ \text{F-Size}it + \varepsilon it$

In Model 2, we replace P-woman with Blau Index which is another proxy variable for gender diversity on board.

Model 2: $Qit = \beta 0 + \Sigma \beta 1 \text{BLAU}it + \Sigma \beta 2\ \text{LEVER}it + \Sigma \beta 3\ \text{ROA}it + \Sigma \beta 4\ \text{F-Size}it + \varepsilon it$

Additionally, in Model 3, we replace Blau with another proxy variable i.e. Shanon Index.

Model 3: $Qit = \beta 0 + \Sigma \beta 1 \text{SHANNON}it + \Sigma \beta 2\ \text{LEVER}it + \Sigma \beta 3\ \text{ROA}it + \Sigma \beta 4\ \text{F-SIZE}it + \varepsilon it$

Where,
Qit = Tobins' Q, our proxy for financial performance of firm i in period t.
P-WOMANit = Percentage of women on board of firm i in period t.
BLAUit = Blau Index, which is a proxy for gender diversity, of firm i in period t.
SHANNONit= Shannon Index, which is a proxy for gender diversity, of firm i in period t.
ROA= Return on Assets of firm i in period t.
F-SIZE = Firm Size of firm i in period t.
LEVER= Leverage of firm i in period t.
Eit = random disturbance term

7. RESULTS & ANALYSIS

Table II: Descriptive Statistics for all the variables.

	Mean	Median	Maximum	Minimum	Standard Deviation
BLAU	0.267942	0.244898	0.46875	0.142012	0.081204
P-Women	0.166263	0.142857	0.375	0.076923	0.068923
LEVER	0.340035	0.115122	5.78	-1.3	0.985695
SHANNON	0.434102	0.410116	0.661563	0.271189	0.097206
F-SIZE	6440.087	929.42	194679.5	18.84	27437.67
ROA	2.680766	5.88	79.1	-578.23	54.08431
TOBINS-Q	3.388572	2.336391	17.37609	-0.14	3.158144

Source:Authors' own calculations

The above table provides descriptive statistics for the independent, dependent and control variables used, in the form of mean, median, standard deviation, maxima, minima etc. by using 123 observations. If we look at the mean of Blau index and Shannon index, it is approximately 0.26 and 0.43 respectively, whereas average Tobin's Q is 3.38 (approximately). This depicts that the diversity is too less to have any impact on financial performance. The average percentage of women on board is 16.62% only, amongst 123 observations. The average assets are Rs.6440.86 crores which suggests good financial health of the firms. The lower leverage ratios are considered to be good as it suggests that less assets are financed by debt. The above table also depicts a very low mean leverage of 0.34. These statistics support the results of our regression tests as shown below.

Table III: Hausman Test to check whether Random effect or Fixed effect Model shall be used.

Test Summary	Chi-Sq. Statistic	Chi-Sq. d.f.	p-value
Cross-section random	5.670113	5	0.3397

Source :Authors' own calcuations

We performed the Hausman test, which helps to differentiate between the fixed effects model and random effects model in the panel data, with random effects model being adopted on the non-rejection of the null hypothesis. From the test, we find the p-value to be equal to 0.3397 i.e. it is insignificant at 5% (Table III), which means that we do not reject the null hypothesis and therefore, choose the Random Effects Model to run our regression tests.

Table IV: Random Effects Model using the variable P-Woman and other control variables.

Variable	t-Statistic	p-value
C	1.275903	0.2053
P-WOMAN	0.089598	0.9288
F-SIZE	-0.669224	0.5051
LEVER	3.594660*	0.0005*
ROA	6.392905*	0.0000*

Source: Authors' own calculations (*at 5% significance level.)

The first multiple regression is run using P-WOMEN as an independent variable (Table IV) and three control variables, i.e., total assets, leverage and return on assets. On applying the test, we find that the p-value for P-WOMEN variable is 0.9288 which is statistically insignificant and therefore we do not reject the null hypothesis. This means that the proportion of women directors does not have any impact on Tobin's Q ratio. The p-value of total assets (0.5051) is also insignificant which means that the value of total assets of a firm will not impact its financial performance or Tobin's Q ratio. The p-value for leverage is 0.0005 which is significant, thus we reject the null hypothesis and conclude that the leverage ratio would impact Tobin's Q. A rise in leverage will help to increase the earnings per share for the shareholders, which would send a positive signal to the market about the financial status of the firm and thus affect the stock price and Tobin's Q. The variable ROA is also significant, and we conclude that ROA positively affects Tobin's Q. The rise in ROA would bring a rise in Tobin's Q and thus improve firm performance.

Table V: Random Effects Model using the variable Blau Index and other control variables

Variable	t-Statistic	P-Value
C	0.897421	0.3719
BLAU	0.213332	0.8316
LEVER	3.591204*	0.0005*
F-SIZE	-0.684969	0.4952
ROA	6.419894*	0.0000*

Source: Authors' own calculations (*at 5% significance level.)

The second regression model uses Blau index as an independent variable instead of P-WOMEN, rest all variables remaining the same. Table V shows the results. The p-value of Blau index is 0.8316 which indicates that it is insignificant at 5%. Thus, Blau index value, that depicts diversity, does not influence the Tobin's Q ratio. We also interpret the p-value (0.4952) of total assets in the second model to be statistically insignificant and hence do not reject the null hypothesis. Total assets do not impact the financial performance of firms in our sample. The p-value of leverage is 0.0005 which is statistically significant, thus leverage will impact the financial performance. Also, since ROA is significant, it positively impacts the Tobin's Q.

Table VI: Random Effects Model using the variable Shannon Index and other control variables

Variable	t-Statistic	P-Value
C	0.897421	0.3719
BLAU	0.213332	0.8316
LEVER	3.591204*	0.0005*
F-SIZE	-0.684969	0.4952
ROA	6.419894*	0.0000*

Source: Authors' own calculations (*at 5% significance level.)

The third regression model applied using the Shannon index (Table VI) gives us the same results as the previous one. The p-value of Shannon Index (0.8043) is insignificant; hence we cannot reject the null hypothesis. Therefore, Shannon index value depicting diversity does not have any impact on financial performance. The p-value of total assets is 0.4910, which is insignificant, thus we cannot reject null hypothesis and total assets have no effect on the Tobin's Q ratio. However, Leverage and ROA will positively impact the financial performance, as their p- values, 0.0005 and 0. 00 respectively, are statistically significant.

8. CONCLUSION

The study has made contributions to the extant literature and has supported the arguments that diversity on board will not impact the financial performance of the firm. The presence of women on boards does not influence Tobin's Q and hence the performance. The important aspect to be noted here is that most of the firms in our sample have appointed just one female director on their boards, where they could have appointed even more. Hence, we can say that the firms have appointed the women directors only to comply with the mandatory provisions of Companies Act 2013, which shows that this is practised in letter and not in spirit. The existence of women on boards is mere tokenism, as the appointment of a single woman director will not bring much change in a firm's overall decision making.

Examining the causes for no improvement in the performance of the firms, we could relate back to some of the major arguments, like poor women representation. A closed culture at board level, dominated by male directors, might have led to the negation of the diversity advantages. Also, male directors are likely to favour male directors over women due to the in-group favouritism (Hutchinson et al., 2015). Moreover, an essential theory in this regard from the psychological perspective, the similarity attraction theory, advocates the fact that directors feel more comfortable and trustworthy interacting with directors from same demographics (Chatman & O' Reilly, 2004). Another set of arguments in this regard is that there lies an empathetic concern towards facilitating advancement of those belonging to same class (Macdonald & Westphal, 2013). Finally, the argument based on the critical mass theory points out that the critical mass of women positively impacts board strategic tasks, but since in our sample firms, critical mass could not be achieved, there was little scope for improvement in the firm performance (Torchia et al., 2011).

The Firm size measured in terms of the total assets also shows no impact on financial performance. This is contrary to our expectations, as other studies have concluded that size of the firm has a positive effect on performance (Halder et al., 2018; Buallay, 2017; Torchia et al., 2011). This could be because of the short time period of 3 years that is studied. It is possible that in the long run, we are able to see the impact of firm size on firm performance. Also, a low debt to asset ratio and ROA is found to positively improve performance, consistent with Campbell and Mínguez-Vera (2008). This may be because of the operation of

signalling theory which says that favourable financial ratios send a positive signal to the investors. Higher leverage ratio is seen to positively affect the firm performance and is consistent with Yasser et al. (2017). The ROA also sends a positive signal to the general investors about good financial health of the company, as it may result in higher returns to the shareholders.

9. LIMITATIONS & FUTURE SCOPE

There are a few limitations to our study which open opportunities for future research. In our study, data for just three financial years were considered. Data for larger years would make the results much more robust.

The paper has just focussed on the short term performance; long term performance can also be tested after using data of five or more years in future. Also, only IPO firms were considered for the study. Other industries such as banking and financial companies could also be taken up for future studies.

Moreover, we have considered only the gender aspect of board diversity. Other variables such as education levels, experience of the women directors, directors coming from family firms, are also some crucial areas of research that could be considered in the future.

10. MANAGERIAL IMPLICATIONS

The results of the study do not present any improvement in the performance of the firms owing to the presence of women on board. The representation of women on board is very less and hence the benefits that could have been derived from the diversity were not available to the firms in our sample. It appears that the firms have not taken the mandatory clause in true spirits, but have affirmed the provision only as mere tokenism. Until and unless the proportion of women on board enhances, there is little room for improvement in firm performance. Based on the critical mass theory, it has already been highlighted that a minimum of three women directors are required to influence the decision making of the board. Hence, this paper urges firms to enhance the representation of women based on adequate qualification and experience to derive maximum benefits.

REFERENCES

Adams, R. B. & Ferreira, D. (2009). Women in boardroom and their impact on governance and performance. *Journal of Financial Economics, 94*(2), 291-309. https://doi.org/10.1016/j.jfineco.2008.10.007

Ahern, K. R. & Dittmar, A. K. (2012). The changing of the boards: The impact on firm valuation of mandated female board represenatation.. *The Quarterly Journal of Economics, 127*(1), 137-197. https://doi.org/10.1093/qje/qjr049

Alghifari, S., Triharjono, S., & Juhaeni, Y. (2013). Effect of return on assets (roa) against Tobin's q: Studies in food and beverage company in Indonesia stock exchange years 2007-2011. *International Journal Of Science and Research (IJSR)*, *2*, 108-116.

Buallay, A., Hamdan, A., & Zureigat, Q. (2017). Corporate governance and firm performance: Evidence from Saudi Arabia. *Australasian Accounting, Business and Finance Journal*, *11*(1), 78-98. https://doi.org/10.14453/aabfj.v11i1.6

Campbell, K., & Mínguez-Vera, A. (2008). Gender diversity in the boardroom and firm financial performance. *Journal of Business Ethics*, *83*(3), 435-451. https://doi.org/10.1007/s10551-007-9630-y

Carter, D. A., Simkins, B. J. & Simpson, W. G. (2003). Corporate Governance, Board Diversity and Firm Value. *Financial Review, 38*(1), 33-53. https://doi.org/10.1111/1540-6288.00034

Coakley, J., Hadass, L. & Wood, A. (2007). Post-IPO operating performance, venture capital and bubble years. *Journal of Business Finance & Accounting, 34*(9), 1423-1446. https://doi.org/10.1111/j.1468-5957.2007.02055.x

Chatman, J. A., & O'Reilly, C. A. (2004). Asymmetric reactions to work group sex diversity among men and women. *Academy of Management Journal*, *47*(2), 193-208. https://doi.org/10.2307/20159572

Chen, C. J., & Jaggi, B. (2000). Association between independent non-executive directors, family control and financial disclosures in Hong Kong. *Journal of Accounting and Public policy*, *19*(4-5), 285-310. https://doi.org/10.1016/S0278-4254(00)00015-6

Chung, K. H., & Pruitt, S. W. (1994). A simple approximation of Tobin's q. *Financial Management*, 70-74. https://doi.org/10.2307/3665623

Das, . A. & Dey, S. (2016). Role of corporate governance on firm performance : A study on large Indian corporations after implementation of Companies' Act 2013. *Asian Journal of Business Ethics, 5*(1), 49-164. https://doi.org/10.1007/s13520-016-0061-7

Gallego-Álvarez, I., García-Sánchez, I. M., & Rodríguez-Dominguez, L. (2010). The influence of gender diversity on corporate performance. *Revista de contabilidad*, *13*(1), 53-88. https://doi.org/10.1016/S1138-4891(10)70012-1

Giraldez-Puig, P., & Berenguer, E. (2018). Family female executives and firm financial performance. *Sustainability*, *10*(11), 4163. https://doi.org/10.3390/su10114163

Greene, W. H. (2008). The econometric approach to efficiency analysis. *The Measurement of Productive Efficiency and Productivity Growth*, *1*(1), 92-250. https://doi.org/10.1093/acprof:oso/9780195183528.003.0002

Haldar, A., Shah, R., & Rao, S. N. (2015). Gender diversity in large listed Indian companies. *Corporate Ownership & Control*, *12*(13), 1727-9232. https://doi.org/10.22495/cocv12i3c5p8

Haldar, A., Shah, R., Nageswara Rao, S. V. D., Stokes, P., Demirbas, D., & Dardour, A. (2018). Corporate performance: Does board independence matter?–Indian evidence. *International Journal of Organizational Analysis*, *26*(1), 185-200. https://doi.org/10.1108/IJOA-12-2017-1296

Haldar, A., Shah, R., Rao, S. V. D. N., & Bombay, I. I. T. (2014, March). Board room diversity and firm value: Evidence from india. In *Proc. of Annual Conference on Government, Corporate Governance and Growth of Capital Markets, IICM, Mumbai*.

Hausman, J. A., & Taylor, W. E. (1981). Panel data and unobservable individual effects. *Econometrica: Journal of the Econometric Society*, *49*(6),1377-1398. https://doi.org/10.2307/1911406

Hutchinson, M., Mack, J., & Plastow, K. (2015). Who selects the 'right'directors? An examination of the association between board selection, gender diversity and outcomes. *Accounting & Finance*, *55*(4), 1071-1103. https://doi.org/10.1111/acfi.12082

Hutchinson, M., & Gul, F. A. (2004). Investment opportunity set, corporate governance practices and firm performance. *Journal of Corporate Finance*, *10*(4), 595-614. https://doi.org/10.1016/S0929-1199(03)00022-1

Jain, B. A. & Kini, O. (1994). The post-issue operating performance of IPO firms. *The Journal of Finance, 49*(5), 1699-1726. https://doi.org/10.1111/j.1540-6261.1994.tb04778.x

Joshi, R. (2017). *Does gender diversity improve firm performance? Evidence from India.* JustJobs Network Inc. Retrieved from https://www.justjobsnetwork.org/wp-content/uploads/2018/02/Gender_diversity_firm_performance_web.pdf

Kalsie, A., & Shrivastav, S. M. (2016). Analysis of board size and firm performance: evidence from NSE companies using panel data approach. *Indian Journal of Corporate Governance*, *9*(2), 148-172. https://doi.org/10.1177/0974686216666456

Kaur, A. & Singh, B. (2017). Construing reputation from gender diversity on boards: Indian evidence. *Paradigm , 21*(2), 111-125. https://doi.org/10.1177/0971890717736195

Krishnan, C. N. V., Ivanov, V. I., Masulis, R. W., & Singh, A. K. (2011). Venture capital reputation, post-IPO performance, and corporate governance. *Journal of Financial and Quantitative Analysis*, *46*(5), 1295-1333. https://doi.org/10.1017/S0022109011000251

Lückerath-Rovers, M., (2013). Women on boards and firm performance. *Journal of Management & Governance, 17*(2), 491-509. https://doi.org/10.1007/s10997-011-9186-1

Singapore Management University and McGuinness, P. B. (2017). The economic benefits of women on IPO firm boards. *Perspectives@SMU*. Available at: https://ink.library.smu.edu.sg/pers/353

McDonald, M. L., & Westphal, J. D. (2013). Access denied: Low mentoring of women and minority first-time directors and its negative effects on appointments to additional boards. *Academy of Management Journal*, *56*(4), 1169-1198. https://doi.org/10.5465/amj.2011.0230

Montgomery, C. A., & Wernerfelt, B. (1988). Diversification, Ricardian rents, and Tobin's q. *The Rand Journal of Economics*, *19*(4), 623-632. https://doi.org/10.2307/2555461

Murthy, K. B., & Singh, A. K. (2014). Short-run performance of IPO market in India. *International Journal of Financial Management*, *4*(2), 11-12.

Murthy, K. V., Singh, A. K., & Gupta, L. (2016). Long-run performance of IPO market in India. *International Journal of Financial Management*, *6*(1), 17-27. https://doi.org/10.21863/ijfm/2016.6.1.024

Ryan, M. K. & Haslam, S. A., (2005). The glass cliff : Evidence that women are over - represented in precarious leadership positions. *British Journal of Management, 16*(2), 81-90. https://doi.org/10.1111/j.1467-8551.2005.00433.x

Sanan, N. K., (2016). Board gender diversity and firm performance: Evidence from India. *Asian Journal of Business Ethics, 5*(1), 1-18. https://doi.org/10.1007/s13520-016-0050-x

Sealy, R., Vinnicombe, S., & Singh, V. (2008). The female FTSE report 2008*: A decade of delay.* Cranfield: Cranfield University, School of Management. Retrieved from https://annazavaritt.blog.ilsole24ore.com/wp-content/uploads/sites/54/files/ftse-woman.pdf

Singh, A. K. & Maurya, S. (2018). Corporate governance, ownership structure, and IPO underpricing: Evidence from the Indian new issue market. *Indian Journal of Research in Capital Markets, 5*(1), 7-24. https://doi.org/10.17010/ijrcm/2018/v5/i1/122905

Stirling, A. (1998). On the economics and analysis of diversity. *Science Policy Research Unit (SPRU), Electronic Working Papers Series, Paper*, *28*, 1-156.

Short, H., & Keasey, K. (1999). Managerial ownership and the performance of firms: Evidence from the UK. *Journal of Corporate Finance*, *5*(1), 79-101. https://doi.org/10.1016/S0929-1199(98)00016-9

Terjesen, S., Couto, E. B. & Francisco, P. M. (2016). Does the presence of independent and female directors impact firm performance? A multi country study of board diversity. *Journal of Management & Governance, 20*(3), 447-483. https://doi.org/10.1007/s10997-014-9307-8

Torchia, M., Calabrò, A., & Huse, M. (2011). Women directors on corporate boards: From tokenism to critical mass. *Journal of Business Ethics*, *102*(2), 299-317. https://doi.org/10.1007/s10551-011-0815-z

Torchia, M., Calabrò, A., Gabaldon, P., & Kanadli, S. B. (2018). Women directors contribution to organizational innovation: A behavioral approach. *Scandinavian Journal of Management*, *34*(2), 215-224. https://doi.org/10.1016/j.scaman.2018.02.001

Wellalage, N. H. & Locke, S. (2013). Women on board, firm financial performance and agency costs. *Asian Journal of Business Ethics, 2*(2), 113-127. https://doi.org/10.1007/s13520-012-0020-x

Welbourne, T. M. (1999). Wall street likes its women: An examination of women in the top management teams of initial public offerings. *CAHRS Working Paper Series*, 106.

Yasser, Q. (2012). Affects of Female Directors on Firms Performance in Pakistan. *Modern Economy, 3*, 817-82 https://doi.org/10.4236/me.2012.37104

Yasser, Q. R., Mamun, A. A., & Rodrigs, M. (2017). Impact of board structure on firm performance: evidence from an emerging economy. *Journal of Asia Business Studies*, *11*(2), 210-228. https://doi.org/10.1108/JABS-06-2015-0067

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/327948653

Efficient Fluorescence Quenching of 5-Aminoquinoline: Silver ion recognition study

Article *in* Journal of Luminescence · September 2018

DOI: 10.1016/j.jlumin.2018.09.062

CITATIONS
0

READS
111

4 authors, including:

Nupur Pandey
DSB Campus Kumaun University Nainital
1 PUBLICATION **0** CITATIONS
SEE PROFILE

Mohan Singh Mehata
Delhi Technological University
65 PUBLICATIONS **681** CITATIONS
SEE PROFILE

Nisha Fatima
Kumaun University
1 PUBLICATION **0** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Synthesis and Photophysics of Nanoparticles/Quantum Dots using ps Time-resolved Techniques View project

Photoinduced Charge Transfer Dynamics of Synthesized Semiconducor Quantum Dots and Organic Molecules under the Influence of External Perturbation View project

All content following this page was uploaded by Mohan Singh Mehata on 18 June 2019.

The user has requested enhancement of the downloaded file.

Journal of Luminescence 205 (2019) 475–481

Contents lists available at ScienceDirect

Journal of Luminescence

journal homepage: www.elsevier.com/locate/jlumin

Efficient fluorescence quenching of 5-aminoquinoline: Silver ion recognition study

Nupur Pandey[a], Mohan Singh Mehata[b,*], Nisha Fatma[a], Sanjay Pant[a,*]

[a] Photophysics Laboratory, Department of Physics, Centre of Advance Study, DSB Campus, Kumaun University, Nainital 263002, India
[b] Laser Spectroscopy Laboratory, Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi 110042, India

ARTICLE INFO

Keywords:
5-aminoquinoline
Ag^+ ions
Fluorescence quenching
Charge transfer
Optical sensor

ABSTRACT

Fluorescence (FL) quenching of 5-aminoquinoline (5AQ) by Ag^+ ions has been studied in acetonitrile (ACN) using steady-state and time-resolved fluorescence techniques. In the absence of Ag^+ ions, FL shows a single band assigned to the intramolecular charge transfer (ICT) state. The FL intensity of 5AQ gradually reduces on the successive addition of Ag^+ ions to the solution containing 5AQ. The decrease of FL intensity is nearly ~ 17 fold in the presence of 60 mM Ag^+ ions. Together with the decrease in FL intensity, the FL band maximum is red-shifted. The Stern-Volmer plot has been employed to determine the quenching constant. The nature of quenching has been stated based on the decrease of FL intensity and lifetime in the presence of metal ions. Also, the Benesi-Hildebrand relation has been used to explore the binding mechanism between 5AQ and Ag^+ ions. Thus, 5AQ displays high affinity for Ag^+ ions over other competitive metal ions (Cu^{2+}, Ca^{2+}, Pb^{2+}, Fe^{2+}, Li^+, K^+). Therefore, the present system emerges to be quite important in sensing of Ag^+ ions in an aprotic environment in a concentration range of 4 μM to 60 mM.

1. Introduction

Quinoline and its derivatives have been an intriguing class of molecules in the field of organic and pharmaceutical chemistry and widely used as a prototype compound to many drugs. Multifarious experimental and theoretical researches in photophysics and photochemistry of quinoline-based compounds have found potential applications in preparation of various drugs and dyes. Photoexcited relaxation processes of variously substituted quinolines especially amino- and hydroxyl-quinolines have been considered broadly over the past few years [1–18]. In recent literature, quinoline fused heterocyclic derivatives have been developed that possess diverse biological activities, i.e., anticancer, antimicrobial, antimycobacterial, anticonvulsant, anti-inflammatory and cardiovascular activities [19].

Aminoquinoline (AQ) is an imperative group of fluorophore that plays a momentous role in supramolecular chemistry, electron transfer process, charge transfer, host-guest systems and sensing applications. Photoexcitation to the lowest excited state is usually associated with intramolecular charge transfer (ICT) from amino group to the aromatic ring in these compounds hence producing the ICT state. Enol to keto conversion occurs via excited state proton transfer (ESPT) reaction following photoexcitation. The position of NH_2 group in the quinoline ring and the surrounded solvent molecules/polymers influences the excited state properties of aminoquinolines [5,8–15].

The FL quenching-based studies have been an area of fundamental interest from decades since this can be valuable in elucidating the localization of fluorophores in protein and membrane and their absorptivity to quenchers. Quenching can occur through various mechanisms and become progressively significant due to its applications in fluorescence sensing and drug delivery systems. An array of substances like, metal ions, halides, molecular oxygen, amines etc. can act as quencher [17,18]. Quinoline based fluorescent sensors are extensively used to determine the fluorescence intensity variations and shift of fluorescence wavelength [19–30]. As an example, when halide ions come in contact with excited state 6-methoxyquinoline and 8-methylquinoline, the FL intensity of these probe molecules gets quenched regularly and follows the diffusional quenching mechanism [16–18], and acts as an optical sensor.

In heterocyclic quinoline, the nitrogen atom can act as a chelating site towards metal ions [20]. Photoinduced electron transfer (PET), intramolecular charge transfer (ICT), twisted intramolecular charge transfer (TICT), fluorescence resonance energy transfer (FRET), chelation enhancement quenching effect (CHEQ) or chelation-enhanced fluorescence effect (CEFE) are important mechanisms of fluorescence sensing in the design of quinoline-based fluorescence sensors [20–30]. The development and design of a fluorescent sensor for the variety of

* Corresponding authors.
E-mail addresses: msmehata@dtu.ac.in (M. Singh Mehata), sanjaypant@kunainital.ac.in (S. Pant).

https://doi.org/10.1016/j.jlumin.2018.09.062
Received 8 July 2018; Received in revised form 23 September 2018; Accepted 27 September 2018
Available online 28 September 2018
0022-2313/ © 2018 Elsevier B.V. All rights reserved.

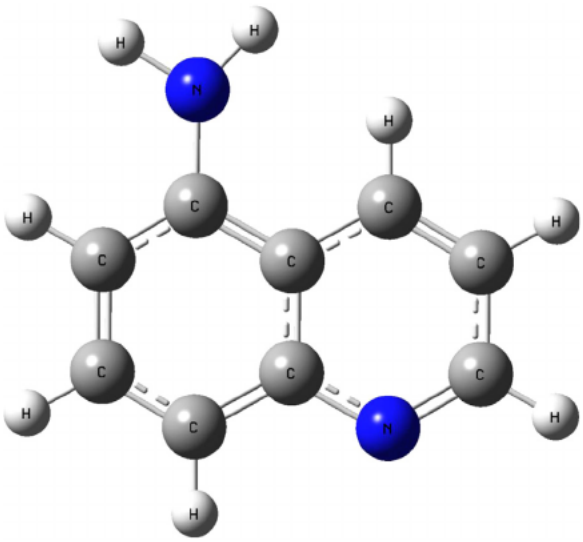
metal ions has acquired attention due to their essential role in medical, environmental and biological applications [22–30]. The FL of ligand may be quenched or enhanced by coordination depending on the influence of the metal ions on the non-radiative processes competing with the radiative process, i.e., FL. The FL resulting from the complex formation is a consequence of numerous factors; the relative position and behavior of the lower electronic excited state of organic ligands and those of metal ions, the heavy atom effects, the structural rigidity and the paramagnetic character of ions [31].

Silver ion, like other heavy metal ions; copper, zinc, lead, mercury, etc., is one of the indispensable toxic metal and has been broadly employed in chemistry, pharmacy, electrical industry, photographic and imaging industry [32–40]. The bioactive silver has gained interest in the domain of biology, such as antimicrobial activity, disinfectant, antibiotics and sterilization even at low concentration level [34]. However, in few cases, rapid absorption of Ag^+ ions causes bioaccumulation, toxicity and irreversible darkening of skin along with disruptive renal and neurological function [34,35]. Furthermore, Ag^+ ions can disable sulphydryl enzymes and bind with imidazole, amine, and carboxyl groups of different metabolites [35,37,41,42]. Since the d^{10} electronic configuration of Ag^+ ions is generally known as a fluorescence quencher following electron transfer and intersystem crossing processes, thus Ag^+ ions could be detected by chemosensors via fluorescence quenching mechanism [28]. Recent reports on fluorescent sensors for detection of Ag^+ ions comprise different fluorophores such as pyrene, oxadiazole, phenanthroline, coumarin, quinoline, gold nanoparticles etc [32–40].

In the present work, we have commenced a detailed study of FL quenching of 5AQ by Ag^+ ions, to comprehend the nature of possible quenching mechanism. Steady-state and time-resolved measurements confirm the quenching mechanism of probe 5AQ in the presence of Ag^+ ions. With the addition of Ag^+ ions, the FL maximum is red-shifted ($\sim$ 46 nm) with $\sim$ 17 fold quenching of fluorescence intensity. Interestingly, no fluorescence quenching and wavelength shift are noticed on addition of variant metal ions (Ca^{2+}, Cu^{2+}, Li^+, K^+, Fe^{2+}, Pb^{2+}). The molecular structure of 5AQ is given in Scheme 1.

2. Experimental section

5-Aminoquinoline (5AQ) purchased from Aldrich chemical company Inc., U.S.A. (with 99% purity) was used as received and matched with the reported FL spectrum to confirm its fluorescence purity (to make sure that there is no impurity). Silver nitrate and all other metal salts were obtained from Loba Chemie and SRL Pvt. Ltd. (with 98–99% of purity) and used as received. Spectroscopic or HPLC grade solvents were used for the present study. The purity of compounds was established by agreement of corrected excitation spectra with the electronic absorption spectra. The stock solution of $AgNO_3$ (0.1 M) was prepared from the stock solution of 5-aminoquinoline (10^{-4} M). From this solution, diluted solutions of $AgNO_3$ were prepared. This ensured that the probe concentration did not change on the addition of $AgNO_3$. Freshly prepared samples at room temperature were used in the present study. A similar method was used during preparation of diluted solutions of other metal ions.



Scheme 1. Molecular structure of 5-aminoquinoline.

2.1. Instrumentation

Steady-state absorption spectra were recorded by UV–visible Double beam spectrophotometer (Systronics Model no: 2203). Fluorolog-3 spectrofluorometer (Horiba Jobin Yvon Inc.) accompanied by available analysis software have been used to record and analyze the excitation and emission spectra. 450 W xenon lamp was used as the excitation source. Transparent quartz cuvette was used for the absorption and emission measurements and frontal geometry was used in recording emission to avoid inner filter effect. For all the measurements, excitation and emission slit widths were chosen about 2.0 nm, and the excitation wavelength was 345 nm. Time-correlated single photon counting (TCSPC) system (DeltaFlex-01-DD, Horiba Jobin Yvon IBH Ltd.) coupled with DeltaDiode (340 nm) and PMT (PPD 850) was used for fluorescence decay measurements and least square fitting and reconvolution methods were used to analyze the data. Instrumental response function was measured using Ludox scatterer. The numbers of counts in the peak channel were at least 10,000.

3. Results and discussion

Absorption, fluorescence (FL), FL-excitation spectra and FL-decays were measured for 5AQ in acetonitrile (ACN) in the absence and presence of varying concentration of silver ions (Ag^+) at room temperature.

Fig. 1 shows absorption spectra of 5AQ in ACN in absence and presence of various concentrations of Ag^+ ions. The absorption maximum appeared at 343 nm, and is ascribed to the $\pi \rightarrow \pi^*$ transition of quinoline group. In presence of a small concentration of Ag^+ ions, keeping the solute concentration fixed the optical density (absorbance) of 5AQ decreases, which is accompanied by broadening of the absorption spectrum. The corresponding optical density decreases from 0.65 to 0.45 with increasing concentration of Ag^+ ions from 0.0 to 7.0 mM. The change in spectral broadening and absorption maximum

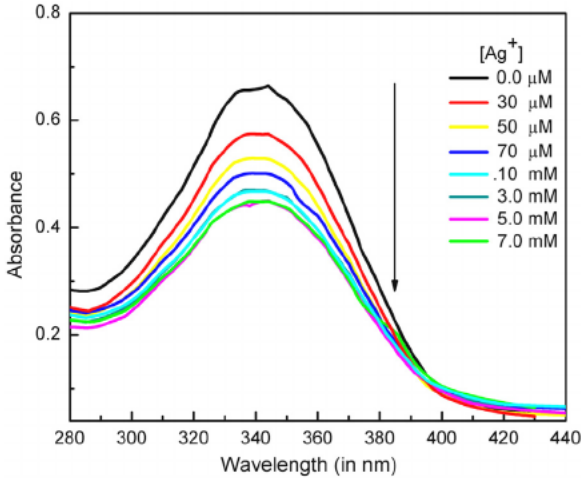


Fig. 1. Absorption spectra of 5AQ (1×10^{-4}M) in ACN in the presence of various concentration of Ag^+ ions.

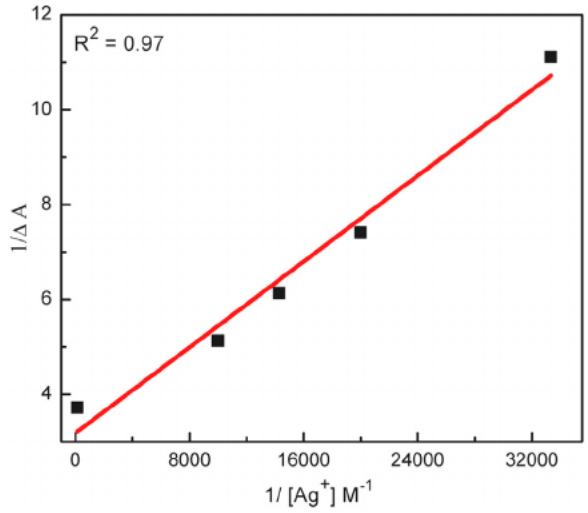


Fig. 2. Benesi-Hildebrand plot of 5AQ in ACN with increasing concentration of Ag^+ ions.

upon addition of Ag^+ ions suggests that there is a ground state complexation between 5AQ and Ag^+ ions. To understand the formation of complex between the probe and the quencher and to determine the binding constant, the well-known Benesi-Hilderband [32] Eq. (1) was applied.

$$\frac{1}{(A - A_0)} = \frac{1}{(A_{max} - A_0)} + \left[\frac{1}{K[M^{n+}]} + 1\right] \quad (1)$$

where A is the absorbance of 5AQ in presence of various concentration of Ag^+ ions, A_0 is the absorbance of 5AQ, A_{max} is the absorbance of 5AQ in the presence of the maximum concentration of Ag^+ ions, $[M^{n+}]$ is the concentration of metal ion and K is the binding constant. A plot between $1/\Delta A$ vs. the inverse of Ag^+ ions concentration is linear (Fig. 2) and demonstrates the 1:1 complex of 5AQ: Ag^+. The binding constant (K) determined from the ratio of intercept/slope of the plot comes out to be $1.407\times10^4 M^{-1}$.

Fig. 3(a,b) shows the fluorescence spectra of 5AQ in ACN in the absence and presence of varying concentration of Ag^+ ions at room temperature. The fluorescence maximum is around 482 nm upon excitation at 345 nm and is independent of excitation wavelength. The fluorescence of 5AQ reduced upon addition of small amount (about 4 μM) of Ag^+ ions. While increasing the concentration of Ag^+ ions from 4 μM, the fluorescence intensity is gradually decreased, as shown in Fig. 3(a). The FL intensity diminished nearly 17 folds in the presence of 60 mM concentration of Ag^+ ions. Thus, the FL of 5AQ is quenched significantly in the presence of Ag^+ ions in ACN. In addition to quenching, the FL spectrum is shifted towards red with the addition of Ag^+ ions. The red shift in the FL spectrum is about 46 nm for the highest concentration of Ag^+ ions taken in the study (60 mM), as shown in (Fig. 4). The excitation spectrum of 5AQ in ACN was found to be independent of monitored emission wavelength and exhibits well defined L_a and L_b bands and resembles to the absorption spectrum [10,11,13]. Fig. S1(a) shows the overlapped excitation spectrum of 5AQ in ACN on successive addition of Ag^+ ions. The excitation spectrum shifted from 345 nm to 360 nm with the addition of Ag^+ ions as shown in Fig. S1(b). Together with the spectral shift, the intensity of excitation spectrum gradually decreases with increasing concentration of Ag^+ ions. The change in the excitation spectrum is in agreement with that observed in absorption spectrum and demonstrates the ground state complex formation between Ag^+ ions and 5AQ in ACN.

The FL maximum shifted from 482 nm to 528 nm while changing the concentration of Ag^+ ions from 0.0 μ M to 60 mM. A linear relationship has been observed between the ratio of FL intensity at two

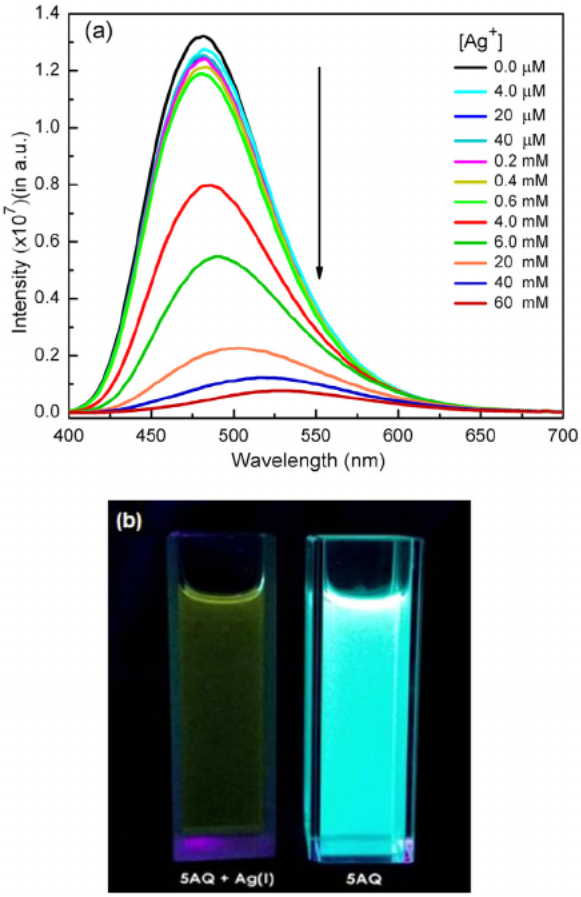


Fig. 3. (a) Fluorescence spectra of 5AQ $(1 \times 10^{-4}\ M)$ in ACN upon addition of various concentration of Ag^+ ions at excitation wavelength of 345 nm. (b) Fluorescence of 5AQ with UV light irradiation in the presence (left) and absence (right) of Ag^+ ions.

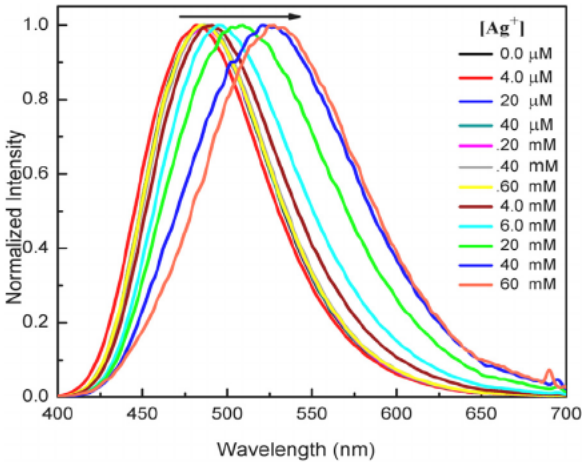


Fig. 4. Normalized fluorescence spectra of 5AQ in ACN in presence of various concentrations of Ag^+ ions.

wavelengths (I_{530}/I_{480}) and Ag^+ ion concentration (Fig. 5). In pure ACN, 5AQ shows a Stokes shift of ~ 8322 cm^{-1} and increases upto 10,058 cm^{-1} upon addition of Ag^+ ions (60 mM). The relation between Stoke's shift (Table 1) and Ag^+ ions concentration is shown in Fig. S2. This further demonstrates formation of complex in the ground state.

To understand the nature of FL quenching, the decrease of FL intensity is plotted as a function Ag^+ ions concentration following the so-called Stern-Volmer (S-V) relation Eq. (2) and modified S-V relation Eq.

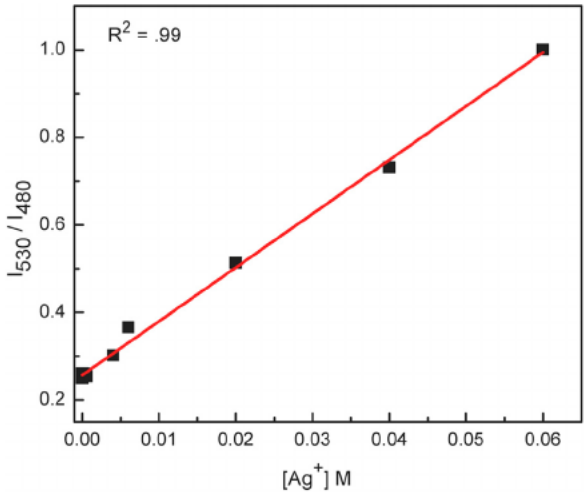


Fig. 5. Plot of the ratio of the intensity (I_{530}/I_{480}) of 5AQ in the presence of various concentration of Ag^{+}ions.

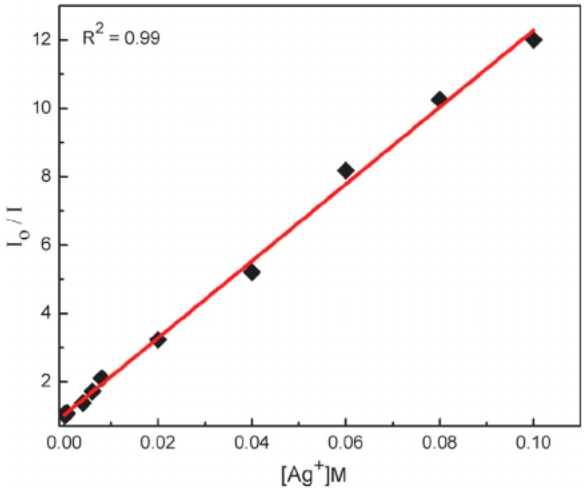


Fig. 6. The Stern-Volmer plot between the relative fluorescence intensity of 5AQ and Ag^{+} ions concentration.

Table 1
Steady-state parameters of 5AQ in acetonitrile in the presence of various concentration of Ag^{+} ions.

Concentration of Ag^{+} ions	$^{max}\lambda_{abs}$ (nm)	$^{max}\lambda_{em}$ (nm)	Stokes shift Δν (cm^{-1})
0.0 μM	344	482	8322
4.0 μM	344	482	8322
40 μM	344	483	8366
0.4 mM	344	484	8408
0.6 mM	344	484	8408
4.0 mM	344	485	8451
6.0 mM	344	490	8661
8.0 mM	344	495	8868
20 mM	344	502	9149
40 mM	344	525	10,019
60 mM	344	528	10,058

(3):

$$\frac{I_0}{I} = \frac{\tau_0}{\tau} = 1 + k_q\tau_0[Q] = 1 + K_{SV}[Q] \quad (2)$$

$$\frac{\left[\frac{I_0}{I} - 1\right]}{[Q]} = (K_{SV} + K_g) + K_{SV}K_g[Q] \quad (3)$$

where I_0 and I are the fluorescence intensities in the absence and presence of the quencher, respectively. [Q] is the quencher concentration and K_{SV} ($k_q\tau_0$) is the S-V constant, which represents the efficiency of fluorescence quenching. k_q is the bimolecular constant, τ_0 is the decay time in absence of quencher and K_g is ground state association constant. The value of K_{SV} can be determined from the slope of the graph displaying dependence of intensity or decay time on [Q], as expressed by Eq. (2). The obtained S-V plot for 5AQ in the presence of Ag^{+} ions is linear with Adj R^2 = 0.99, as shown in Fig. 6. The FL intensity may reduce either due to dynamic or static quenching, even with the linear relation of S-V plot. Thus, the steady-state spectral analysis only is not sufficient to confirm the nature of quenching. When decay time does not change with a decrease in fluorescent intensity, only static quenching is present i.e., K_{SV} = 0. A non-fluorescent complex is formed between the fluorophore and the quencher in case of quenching being static. The other case is that of dynamic quenching (i.e., K_g = 0) in which quenching response competes with emission in timescale and can be governed by the diffusion of quencher in any medium and its collision with excited probe molecule. These two quenching effects i.e., dynamic and static quenching, involve molecular interaction between the fluorophore and quencher. Simultaneous occurrence of static and dynamic quenching results in the positive deviation of plot I_0/I versus [Q] as Eq. (2) follows quadratic dependence with quencher concentration. The relation, $[I_0/I - 1]/[Q] = (K_{SV} + K_g) + K_{SV}K_g[Q]$ gives K_g after substituting the K_{SV} value retrieved from the data [1,17,18].

To ascertain the fluorescence quenching mechanism of 5AQ by Ag^{+} ions, time-resolved FL experiments were performed for various concentrations of Ag^{+} ions. FL decays were collected at λ_{ex} = 340 nm and λ_{em}= 530 nm. The FL decay of 5AQ fits to a bi-exponential function $\alpha_1 \exp(-t/\tau_1) + \alpha_2 \exp(-t/\tau_2)$ in an aprotic solvent (ACN). The term τ_1 and τ_2 are the lifetimes and α_1 and α_2 are the corresponding pre-exponential factors. The fitted lifetime are 1.26 ± 0.2 ns (0.93 %) and 11.60 ± 0.01 ns (99.07 %). The contribution of smaller decay constant is very small (< 1%). As an approximation, the decay fitted single exponential with a lifetime of 11.50± 0.01. Upon increasing the concentration of Ag^{+} ions, the lifetime decreases monotonically. It is interesting to note that, the lifetime component decreases from 11.50 ns to 3.14 ns on addition of 40 mM of Ag^{+} ions. The fitted overlapped decay curves obtained at varying concentrations of quencher are shown in Fig. 7 and the analyzed parameters are summarized in (Table 2). The decrease in FL lifetime plotted as a function of Ag^{+} ions concentration follow the S-V relation Eq. (2) again linearly, as shown in Fig. 8. Thus, the decrease in lifetime and the linear S-V plots demonstrate the dynamic quenching as a predominant mechanism. The value of K_{SV} constant for 5AQ obtained from the slope of the plot (Fig. 6) comes out to be 112.7 M^{-1} and k_q is around 9×10^{9} M^{-1} s^{-1}, respectively.

Furthermore, the absorption and fluorescence spectra of 5AQ in ACN in the presence and absence of different metal ions (Cu^{2+}, Ca^{2+}, Pb^{2+}, Fe^{2+}, Mg^{2+}, Li^{+}, K^{+}) were examined. The fluorescence response of 5AQ in the presence of various metal ions under UV is shown in Fig. S3 of the Supporting information. The FL of 5AQ upon addition of 40 mM concentration of various metal ions is illustrated in Fig. 9. As discussed, the fluorescence intensity of 5AQ is quenched significantly on addition of Ag^{+} ions whereas for other metal ions it shows minimal change. To investigate the interference from other metal ions we further conducted the competition experiment in the presence of 40 mM of various metal ions (Cu^{2+}, Ca^{2+}, Pb^{2+}, Fe^{2+}, Mg^{2+}, Li^{+}, K^{+}) with subsequent addition of 40 mM of Ag^{+} ions, as shown in Fig. 10. These results reveal that the recognition of Ag^{+} ion by probe 5AQ is not significantly influenced by other competitive metal ions. Thus, it is observed that the probe molecule 5AQ has good selectivity for Ag^{+} ions (as compared to the other metal ions under the same condition and concentration) and hence can be used to detect Ag^{+} ions in polar aprotic medium, which

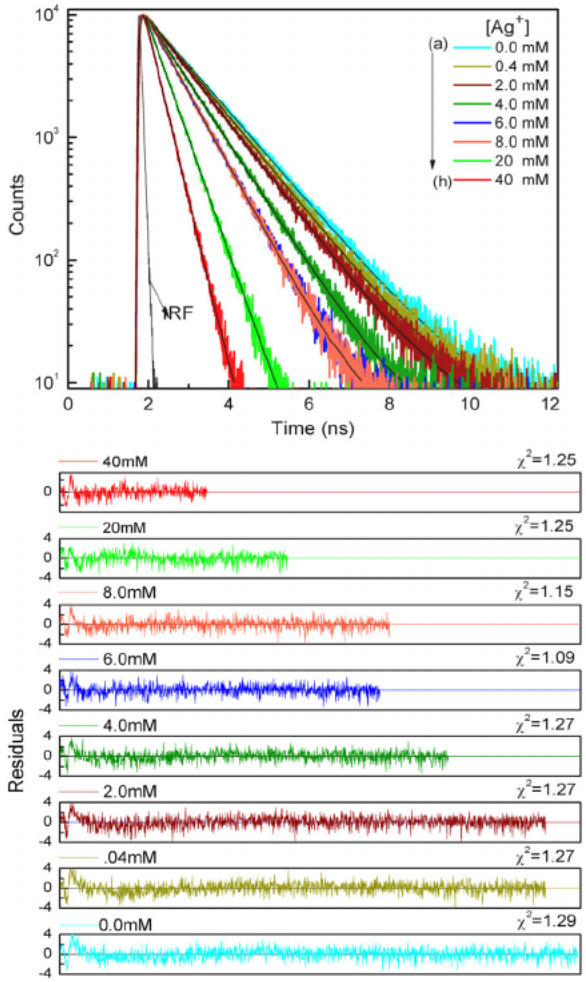


Fig. 7. Overlapped fitted fluorescence decay curves of 5AQ in presence of various concentration of Ag^+ ions and corresponding residuals at λ_{ex} = 340 nm and λ_{em} = 530 nm.

Table 2
Fluorescence lifetime parameters of 5AQ with various concentration of Ag^+ ions at λ_{ex} = 340 nm and λ_{em} = 530 nm.

Concentration of Ag^+ ions	τ (ns)	χ^2	τ_0/τ
0.0 mM	11.50 ± 0.01	1.29	1.0
0.4 mM	10.84 ± 0.01	1.27	1.060
2.0 mM	10.02 ± 0.009	1.27	1.148
4.0 mM	8.64 ± 0.008	1.27	1.331
6.0 mM	7.40 ± 0.09	1.09	1.554
8.0 mM	7.31 ± 0.007	1.15	1.573
20.0 mM	4.65 ± 0.005	1.25	2.473
40.0 mM	3.14 ± 0.004	1.25	3.662

might be due to high-affinity of Ag^+ ions to 5AQ as compared to other metal ions.

As discussed, in the absence of quencher, the FL decay of 5AQ is fitted with a bi-exponential function. The contribution of the shorter component is very small (< 1%) and contribution of the longer component is dominant in the overall decay function. The smaller lifetime component is assigned to the locally excited (LE) state while the longer lifetime component is assigned to the intramolecular charge transfer (ICT) state [10,11,13]. Thus, the ICT is probably responsible for the overall FL. However, there is no rise time component in the FL decay, indicating that ICT may be faster than the resolution of our experimental setup.

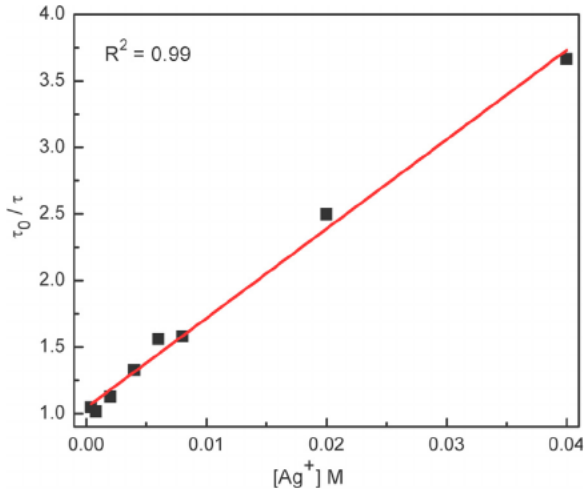


Fig. 8. Stern-Volmer plot between the lifetime of 5AQ and Ag^+ ions concentration.

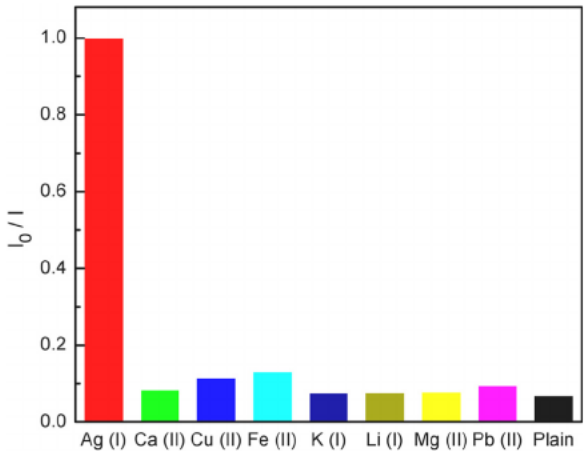


Fig. 9. Fluorescence response of 5AQ in ACN in the presence of various metal ions. The excitation wavelength was 345 nm.

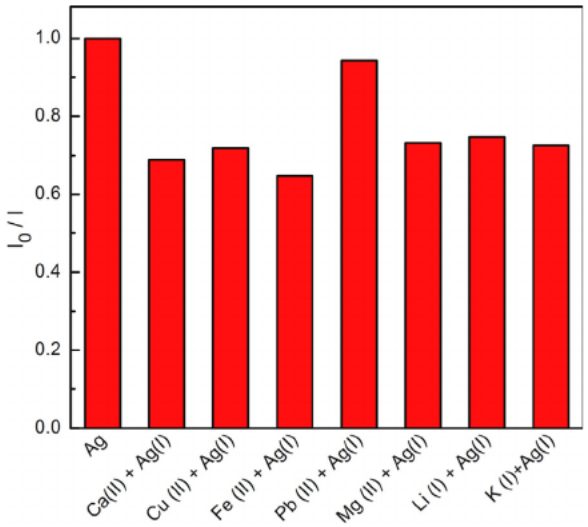


Fig. 10. Fluorescence response of 5AQ (10^{-4} M) in presence of 40 mM of various metal ions with subsequent addition 40 mM of Ag^+ ions to the solution at λ_{ex} = 345 nm.

In the presence of Ag^+ ions, the decay trend is very similar to that in the absence of quencher. However, the observed decay fits well with mono-exponential function with an approximation. The lifetime decreases with increasing the concentration of Ag^+ ions, which indicates that the FL corresponding to the ICT is quenched. This is due to the complexation between 5AQ and Ag^+ ions, as evident from the absorption and excitation spectra. While increasing the concentration of Ag^+ ions, the absorption/excitation intensity decreases and the spectra shift towards red. This may be due to the formation of a complex between 5AQ and Ag^+ ions, which has minimal oscillator strength/absorption intensity. The shift in the FL spectrum may probably be ascribed to the different ICT states formed by the interaction of 5AQ with Ag^+ ions. The decrease of FL lifetime also indicates the decrease of the FL of ICT state formed from the LE state.

Furthermore, it has been classified that metal ions, such as Ag^+ ions, being soft cations, prefer soft donors for complexation [32]. The nitrogen atom of 5AQ is involved in the coordination with Ag^+ ions. The interaction between Ag^+ ions and the nitrogen atom in 5AQ moiety enhances the electron-withdrawing ability of 5AQ that leads to a different state and results in a large red-shifted FL. The donation of the $-C = N$ nitrogen lone pair electrons to the Ag^+ may decrease the fluorescent intensity [21]. Generally, the FL of luminescent ligands is quenched by complexation of the ligand with transition metal ion due to paramagnetic quenching. In this case, the unpaired electrons of the metal ions actively interact with the π electrons of the ligand, providing a path way for inter system crossing from singlet excited state of the ligand to the higher multiplicity states (i.e., triplet state) due to interaction with metal ions [31].

It is important to mention that the present study has been performed in some other solvents also. Since 5AQ shows weak fluorescence in polar protic solvents (water, methanol, ethanol), therefore enhancement in FL intensity is noticed by adding different concentrations of Ag^+ ions. However, a small amount of quenching is observed in DMSO (dimethyl sulfoxide) as compared to that in ACN (acetonitrile) in the presence of Ag^+ ions. Thus, this makes the system attractive for FL quenching/sensing in ACN by Ag^+ ions.

4. Conclusions

In summary, we have studied fluorescence-quenching behavior of 5AQ for Ag^+ ions in ACN, which results in ~ 17-fold decrement in FL intensity upon addition of Ag^+ ions. The lifetime is reduced from 11.5 to 3.14 ns. The large Stokes shifted FL, which produces from ICT state is advantageous to avoid self-absorption and interference from the light source in fluorescence sensing technology. We demonstrate dynamic quenching for all concentration of silver ions used in the study. Stern-Volmer quenching constant (K_{SV}) as $112.7\ M^{-1}$, biomolecular quenching constant (k_q) as $9.8 \times 10^9\ M^{-1}s^{-1}$ and binding constant (K) as $1.407 \times 10^4\ M^{-1}$ have been determined for Ag^+ ions. In addition, 5AQ shows high affinity for Ag^+ ions over other competitive metal ions. Therefore, the proposed system emerges to be quite interesting in sensing/recognition of Ag^+ ions in protic environment in a concentration range of $40\ \mu M$ to $60\ mM$.

Acknowledgments

Authors are thankful to Prof. A.B. Melkani, Department of Chemistry, Kumaun University for providing a facility for absorption measurement. Author NP is thankful to **University Grants Commission** (UGC) (Sanction No. F, 5–53/2007(BSR) New Delhi India for providing research fellowships under the scheme for the meritorious students in sciences (RFSMS) and Dr. J.P. Bridhkoti for his valuable discussions. Author MSM thanks Science and Engineering Research Board (SERB Department of Science & Technology) (EMR/2016/001110) (DST), Govt. of India for financial support.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jlumin.2018.09.062

References

[1] J.R. Lakowicz, Principles of Fluorescence Spectroscopy, third ed., Springer, 2006, https://doi.org/10.1007/978-0-387-46312-4.
[2] H. Mishra, S. Pant, H.B. Tripathi, Temperature-dependent time-resolved fluorescence study of cinchonine alkaloid dication, J. Fluoresc. 18 (2008) 17–27, https://doi.org/10.1007/s10895-007-0229-6.
[3] M.S. Mehata, H.C. Joshi, H.B. Tripathi, Fluorescence characteristics of protonated form of 6-hydroxyquinoline in Nafion® film, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 59 (2003) 559–567, https://doi.org/10.1016/S1386-1425(02)00206-8.
[4] Y.H. Kim, J.S. Youk, S.H. Kim, S.K. Chang, Solvatochromic fluorescence behavior of 8-aminoquinoline-benzothiazole: a sensitive probe for water composition in binary aqueous solutions, Bull. Korean Chem. Soc. 26 (2005) 47–50, https://doi.org/10.5012/bkcs.2005.26.1.047.
[5] D. Panda, D. Ghosh, A. Datta, Acid-base behavior of 3-aminoquinoline in its ground and excited states, J. Photochem. Photobiol. A Chem. 207 (2009) 254–259, https://doi.org/10.1016/j.jphotochem.2009.07.017.
[6] R. Rautela, N.K. Joshi, H.C. Joshi, N. Tewari, S. Pant, Solvatochromic study of 2-hydroxy-4-methylquinoline for the determination of dipole moments and solute-solvent interactions, J. Mol. Liq. 154 (2010) 47–51, https://doi.org/10.1016/j.molliq.2010.04.004.
[7] N.K. Joshi, H.C. Joshi, R. Gahlaut, N. Tewari, R. Rautela, S. Pant, Steady State and time-resolved fluorescence study of isoquinoline: reinvestigation of excited state proton transfer, J. Phys. Chem. A 116 (2012) 7272–7278, https://doi.org/10.1021/jp210516t.
[8] A.K. Singh, S. Ghosh, R. Kancherla, A. Datta, Engineering the excited-state dynamics of 3-aminoquinoline by chemical modification and temperature variation, J. Phys. Chem. B 120 (2016) 12920–12927, https://doi.org/10.1021/acs.jpcb.6b09939.
[9] S.G. Schulman, L.B. Sanders, Fluorescence and phosphorescence of 5- and 8-aminoquinoline, Anal. Chim. Acta 56 (1971) 83–89, https://doi.org/10.1016/S0003-2670(01)80111-2.
[10] J.P. Bridhkoti, H.C. Joshi, S. Pant, Fluorescence characteristics of 5-aminoquinoline in acetonitrile: water, J. Mol. Liq. 164 (2011) 197–200, https://doi.org/10.1016/j.molliq.2011.09.014.
[11] J.P. Bridhkoti, H. Mishra, H.C. Joshi, S. Pant, Photochemistry of 5-aminoquinoline in protic and aprotic solvents, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 79 (2011) 412–417, https://doi.org/10.1016/j.saa.2011.02.044.
[12] D. Panda, A. Datta, The role of the ring nitrogen and the amino group in the solvent dependence of the excited-state, dynamics of 3-aminoquinoline, J. Chem. Phys. 125 (2006) 1–10, https://doi.org/10.1063/1.2232199.
[13] J.P. Bridhkoti, R. Gahlaut, H.C. Joshi, S. Pant, Effect of positional substitution of amino group on excited state dipole moments of quinoline, J. Lumin. 131 (2011) 1869–1873, https://doi.org/10.1016/j.jlumin.2011.04.038.
[14] S. Kumar, S. Bawa, H. Gupta, Biological activities of quinoline derivatives, Mini-Rev. Med. Chem. 9 (2009) 1648–1654, https://doi.org/10.2174/138955709791012247.
[15] K. Tiwari, P. Arora, N. Pandey, P. Pandey, H.C. Joshi, S. Pant, Experimental and computational approaches on dipole moment of 5-aminoisoquinoline, J. Mol. Liq. 200 (2014) 460–464, https://doi.org/10.1016/j.molliq.2014.10.027.
[16] M.S. Mehata, H.B. Tripathi, Fluorescence quenching of 6-methoxyquinoline: an indicator for sensing chloride ion in aqueous media, J. Lumin. 99 (2002) 47–52, https://doi.org/10.1016/S0022-2313(02)00294-6.
[17] R. Rautela, P. Arora, N.K. Joshi, S. Pant, H.C. Joshi, Fluorescence quenching of 8-methyl quinolinium: an efficient halide indicator mechanism, J. Mol. Liq. 218 (2016) 632–636, https://doi.org/10.1016/j.molliq.2016.03.002.
[18] N.K. Joshi, N. Tewari, P. Arora, R. Rautela, S. Pant, H.C. Joshi, Photophysical behavior and fluorescence quenching by halides of quinidine dication: steady state and time resolved study, J. Lumin. 158 (2015) 412–416, https://doi.org/10.1016/j.jlumin.2014.10.041.
[19] Z. Liu, W. He, Z. Guo, Metal coordination in photoluminescent sensing, Chem. Soc. Rev. 42 (2013) 1568, https://doi.org/10.1039/c2cs35363f.
[20] B. Zhang, H. Liu, F. Wu, G.F. Hao, Y. Chen, C. Tan, Y. Tan, Y. Jiang, A dual-response quinoline-based fluorescent sensor for the detection of Copper (II) and Iron(III) ions in aqueous medium, Sens. Actuators B Chem. 243 (2017) 765–774, https://doi.org/10.1016/j.snb.2016.12.067.
[21] J. Nie, N. Li, Z. Ni, Y. Zhao, L. Zhang, A sensitive tetraphenylethene-based fluorescent probe for Zn^{2+} ion involving ESIPT and CHEF processes, Tetrahedron Lett. 58 (2017) 1980–1984, https://doi.org/10.1016/j.tetlet.2017.04.027.
[22] M. Formica, V. Fusi, L. Giorgi, M. Micheloni, New fluorescent chemosensors for metal ions insolution, Coord. Chem. Rev. 256 (2012) 170–192, https://doi.org/10.1016/j.ccr.2011.09.010.
[23] Xiangming Meng, ShuXin Wang, Manzhou Zhu. Quinoline-Based Fluorescence Sensors, Molecular Photochemistry Satyen Saha, IntechOpen, doi:10.5772/31771, 2012. Available from: ⟨https://www.intechopen.com/books/molecular-photochemistry-various-aspects/quinoline-based-fluorescence-sensors⟩.
[24] Y. Hu, F. Zhao, S. Hu, Y. Dong, D. Li, Z. Su, A novel turn-on colorimetric and fluorescent sensor for Fe^{3+} and its application in living cells, J. Photochem.

Photobiol. A Chem. 332 (2017) 351–356, https://doi.org/10.1016/j.jphotochem.2016.09.006.
[25] J. Wu, W. Liu, J. Ge, H. Zhang, P. Wang, New sensing mechanisms for design of fluorescent chemosensors emerging in recent years, Chem. Soc. Rev. 40 (2011) 3483, https://doi.org/10.1039/c0cs00224k.
[26] P. Arora, K. Suyal, N.K. Joshi, H.C. Joshi, S. Pant, Fluorescence characteristics of 5-amino salicylic acid: an iodide recognition study, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 94 (2012) 119–125, https://doi.org/10.1016/j.saa.2012.03.056.
[27] Y. Li, H. Yu, G. Shao, F. Gan, A tetraphenylethylene-based “turn on” fluorescent sensor for the rapid detection of Ag^+ ions with high selectivity, J. Photochem. Photobiol. A Chem. 301 (2015) 14–19, https://doi.org/10.1016/j.jphotochem.2014.12.013.
[28] C. Liu, S. Huang, H. Yao, S. He, Y. Lu, L. Zhao, X. Zeng, Preparation of fluorescein-based chemosensors and their sensing behaviors toward silver ions, RSC Adv. 4 (2014) 16109, https://doi.org/10.1039/C3RA47392A.
[29] N.J. Wang, C.M. Sun, W.S. Chung, A highly selective fluorescent chemosensor for Ag^+ based on calix − 4-arene with lower-rim proximal triazolylpyrenes, Sens. Actuators B Chem. 171 (2012) 984–993, https://doi.org/10.1016/j.snb.2012.06.014.
[30] A. Affrose, S. Dawood, S. Parveen, B.S. Kumar, Selective sensing of silver ion using berberine, a naturally occurring plant alkaloid, Sens. Actuators B. Chem. 206 (2015) 170–175, https://doi.org/10.1016/j.snb.2014.09.042.
[31] S.N. Sathyanarayana, Electronic Absorption Spectroscopy and Related Techniques, Universities Press, 2001 (ISBN 8173713715,9788173713712).
[32] A.D. Arulraj, R. Devasenathipathy, S.M. Chen, V.S. Vasantha, S.F. Wang, Highly selective and sensitive fluorescent chemosensor for femtomolar detection of silver ion in aqueous medium, Sens. Bio-Sens. Res. 6 (2015) 19–24, https://doi.org/10.1016/j.sbsr.2015.10.004.
[33] L. Chen, Q. Zhao, X.Y. Zhang, G.H. Tao, Determination of silver ion based on the redshift of emission wavelength of quantum dots functionalized with rhodanine, Chin. Chem. Lett. 25 (2014) 261–264, https://doi.org/10.1016/j.cclet.2013.10.029.
[34] J. Prabhu, K. Velmurugan, Q. Zhang, S. Radhakrishnan, L. Tang, R. Nandhakumar, Symmetric fluorescent probes for the selective recognition of Ag-ion via restricted C=N isomerization and on-site visual sensing applications, J. Photochem. Photobiol. A Chem. 337 (2017) 6–18, https://doi.org/10.1016/j.jphotochem.2017.01.006.
[35] S.K. Sheet, B. Sen, R. Thounaojam, K. Aguan, S. Khatua, Ruthenium(II) complex-based luminescent bifunctional probe for Ag^+ and phosphate ions: Ag^+-assisted detection and imaging of rRNA, Inorg. Chem. 56 (2017) 1249–1263, https://doi.org/10.1021/acs.inorgchem.6b02343.
[36] L.Q. Li, L.J. Gao, A novel rosamine based fluorescent sensor for Ag^+ recognition, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 152 (2016) 426–430, https://doi.org/10.1016/j.saa.2015.07.097.
[37] A.S. Al-Kady, M. Gaber, M.M. Hussein, E.Z.M. Ebeid, Fluorescence enhancement of coumarin thiourea derivatives by Hg^{2+}, Ag^+ and silver nanoparticles, J. Phys. Chem. A 113 (2009) 9474–9484, https://doi.org/10.1021/jp905566z.
[38] C.J. Chen, A.J. Zhang, J.B. Hu, Y.H. Miao, X. Liu, A simple quinoline-based fluorescent “turn on” probe with high selectivity for Ag^+ in aqueous media, Inorg. Chim. Acta 435 (2015) 137–141, https://doi.org/10.1016/j.ica.2015.06.014.
[39] G. Ambrosi, E. Borgogelli, M. Formica, V. Fusi, L. Giorgi, M. Micheloni, E. Rampazzo, M. Sgarzi, N. Zaccheroni, L. Prodi, PluS Nanoparticles as a tool to control the metal complex stoichiometry of a new thio-aza macrocyclic chemosensor for Ag(I) and Hg(II) in water, Sens. Actuators B Chem. 207 (2015) 1035–1044, https://doi.org/10.1016/j.snb.2014.07.107.
[40] Y. bao Jiang, C. Gao, X. Zhang, J. shui Yao, Q. ze Liu, X. xia Cai, A highly selective and sensitive fluorescence probe with A-π-D-π-A structure for detection of Ag^+, J. Mol. Struct. 1163 (2018) 33–40, https://doi.org/10.1016/j.molstruc.2018.01.058.
[41] T. Anand, G. Sivaraman, A. Mahesh, D. Chellappa, Analytica Chimica Acta Aminoquinoline based highly sensitive fluorescent sensor for lead (II) and aluminum (III) and its application in live cell imaging, Anal. Chim. Acta 853 (2015) 596–601, https://doi.org/10.1016/j.aca.2014.11.011.
[42] Y.W. Choi, J.J. Lee, C. Kim, A highly selective fluorescent chemosensor based on a quinoline derivative for zinc ions in pure water, RSC Adv. 5 (2015) 60796–60803, https://doi.org/10.1039/C5RA09954D.

Evolving Systems
https://doi.org/10.1007/s12530-019-09304-6

ORIGINAL PAPER

Edge detection based on type-1 fuzzy logic and guided smoothening

Sahil Raheja[1] · **Akshi Kumar**[2]

Received: 17 May 2019 / Accepted: 30 September 2019
© Springer-Verlag GmbH Germany, part of Springer Nature 2019

Abstract
Edge detection is an important phenomenon in computer vision. Edge detection is helpful in contour detection and thus helpful in obtaining the important information. Edge detection process heavily depends on chosen technique. Soft computing techniques are considered as powerful edge detection methods due to their adaptability. This paper presents a fuzzy logic based edge detection method where the quality of edges is controlled using sharpening guided filter and noise due to the sharpening is controlled using Gaussian filter. The accuracy of the method is judged using a variety of statistical measures. It has been found that by proper selecting the smoothening parameters a significant improvement in the detected edges can be obtained.

Keywords Edge detection · Image enhancement · Fuzzy logic · Guided filtering

1 Introduction

In numerous computer vision systems, edge detectors are very important. The process of edge detection is valuable in the analysis of images by decreasing the processed data (Shah 1997). The technique of edge detection is among prime methods applied in various field where processing of digital image is required. The purpose of using edge detection in the processing of an image is to get rid of useless data and preserve the important image characteristics. It includes certain methods like feature extraction, registration, interpretation and image segmentation which are very helpful in keeping the useful data only (Goswami and Chan 2011).

An edge could be defined as the object between boundary and image background (Gevers et al. 2012; Shih 2010). Soft computing is a developing research area that comprises of integral components of machine learning techniques, for example, fuzzy logic, artificial neural networks, adaptive neuro-fuzzy inference system, particle swarm optimization and genetic algorithms etc. These strategies are generally utilized for edge detection and image segmentation in the greater part of the critical applications. Out of these methods, fuzzy rule based classifiers have been used in various classes of engineering problems such as self localization and landmark recognition (Zhou and Angelov 2007), data density estimation (Angelov and Zhou 2008), real time object detection (Angelov et al. 2011) and human activity recognition (Sargano et al. 2017) etc.

It is important to note that a universal definition for edges cannot be created as each human has its own way of judging the edges in images. Therefore, edge is a subjective phenomenon, and outcome depends on individual perspective. Therefore, in this paper, we have discussed image sharpening and fuzzy logic based edge detection, and to control sharpening noise Gaussian kernel based filtering is proposed. The results are compared with notable statistical quality measures.

This paper is organized as follows, in Sect. 2 of the paper related works is presented. Edge detection steps are detailed in Sect. 3 of the paper. Proposed method is discussed in Sect. 4 of the paper. In Sect. 5 results of the paper are discussed, and finally in Sect. 6 of the paper major conclusions are discussed.

✉ Sahil Raheja
sahilrahejaphd@gmail.com

Akshi Kumar
akshikumar@dce.ac.in

[1] Department of Information Technology, Delhi Technological University, Delhi, India

[2] Department of Computer Science and Engineering, DTU, Delhi, India

Published online: 09 October 2019

2 Related works

In this part of the article, we have examined various edge detection strategies and concentrate on digital images' edge detection. Taking about the standard images, the edges are used to define the boundaries of the subject and helpful for division, enlistment, identification. It removes the irrelevant data and interestingly of protecting the essential properties of a picture. There are different conventional systems useful in edges detection mechanism. Some of the popular techniques are Laplacian Roberts, Sobel and gradient (Shih 2010). With the help of masks, edge detection is possible by linear operators. These masks are representing the ideal edge steps in different ways both for intensity and color. Shih utilized similar technique for identification. The comparison of various types of edge detection methods are detailed in Table 1.

Roberts (Rogowska 2000), Prewit and Sobel (Rangarajan 2005) introduced edge detection on the basis of the principle of gradient which demonstrate the impact of these filters on the images which is based on the first derivative. A progressively successful operator is the Laplacian, which utilizes the second derivative in locating the edge (Rangarajan 2005).

Alshennawy and Aly (2009), Aborisade (2011) and Begol and Maghooli (2011) proposed soft computing based methodologies, for example, fuzzy logic. Mehrara et al. (2009) proposed another methodology that works on the concept of back propagation neural network. Zhang et al. (2009) introduced a hybrid technique works on the principle of adaptive neuro-fuzzy inference system (ANFIS) for the detection of edges.

Mathur and Ahlawat (2008) introduced a latest algorithm based on fuzzy relative pixel value. In this methodology, the relative pixel values are examined and subsequently reduce the image processing by the application of artificial intelligence. Table 2 demonstrates edge detection methods work on the concepts of fuzzy logic, neural network and ANFIS.

The other recent notable mechanisms in edge detection are: deep convolutional neural network (Marmanis et al. 2018), fuzzy cellular automata (Farbod et al. 2018), particle swarm optimization (Setayesh et al. 2013), cuckoo search optimization (Gonzalez et al. 2016), Anisotropic Gaussian Kernels (Zhang et al. 2017), convolution neural networks (CNNs) (Jeong et al. 2019; Liu and Lew 2016). Edge detection based on single pixel imaging was proposed in Ren et al. (2018). In recent published work; it is detailed that artificial intelligence and CNN based edge detection methods have shown that these methods fail in presence of small perturbations (Cosgrove and Yuille 2019).

This paper, proposes a technique based on type 1 fuzzy logic and guided image filtering is used for the sharpening of the images. Using sharpening, strength of the edges can be enhanced to detect them easily. Even in case of perturbation sharpening can be used to control edges. The proposed method is effective as it considers the advantages of both fuzzy logic design and guided image smoothening.

Table 1 Comparison of edge detection techniques

Edge-detection techniques	Methods
First-order derivative/Gradient methods	Sobel, Roberts, Prewitt
Second-order derivative/zero crossing	Laplacian of Gaussian
Optimal image detection	Canny edge detector
Soft computing	ANN, PSO, SVM, GA, Ant colony optimization

Table 2 Soft computing based notable techniques

Authors	Techniques
Borji and Hamidi	Fuzzy logic
Alshennawy	Fuzzy logic
Zhang et al.	Adaptive neuro-fuzzy
Mehrara et al.	Neural network
Aborisade	Fuzzy logic
Begol	Fuzzy logic

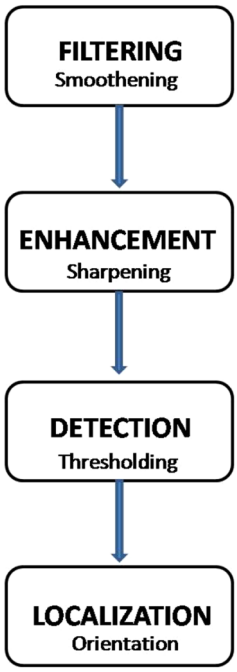


Fig. 1 Edge detection notable steps

3 Edge detection steps

The main aim of edge detection is to successfully locate and identify genuine edges. Edge detection is difficult in low resolution which is mostly found in colour images. In edge detection the main steps are filtering, enhancement, detection and localization which are utilized for detection of correct edges (Fig. 1).

We cannot use images straightway in light of the fact that it might be corrupted by arbitrary varieties in intensities, varieties in brightening, or uneven contrast. This irregular variety in intensity is known as noise. Image sharpening is needed to enhance the quality of image. In enhancing the quality of image, high pass filtering plays an important role. The sharpened image (I_s) signal is proportional to the high pass filtering of original signal and mathematically written as

$$I_S = I + \gamma\left[I - \underbrace{I \otimes H}_{\text{Smoothening}}\right] \tag{1}$$

where, I: original image, H: high pass filter and 'γ' is scaling factor (Fig. 2).

3.1 Canny edge detection

Canny edge detection is a well known method in edge detection; it is a multi-step algorithm and can detect edges in noisy environment by suppressing noise (Canny 1986). The main steps of the algorithms are as under

1. Use Gaussian filter to reduce noise and other redundant information:

$$f(i,j) = H(i,j) \otimes I(i,j)$$
$$\text{Where,} \quad H(i,j) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{i^2+j^2}{2\sigma^2}\right) \tag{2}$$

2. Compute gradient of $(f_i(i,j), f_j(i,j))$ of $f(i,j)$ using any of the gradient operators (Roberts, Sobel, Prewitt, etc.) to get:

$$M(i,j) = \sqrt{f_i^2(i,j) + f_j^2(i,j)} \quad \text{and} \quad \theta = \tan^{-1}\frac{f_j(i,j)}{f_i(i,j)} \tag{3}$$

3. Threshold M:

$$M_T(i,j) = \begin{cases} M(i,j) & \text{if} \quad M(i,j) > \gamma \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where, γ is threshold and is chosen carefully such that that all edges are preserved while major portion of the noise is suppressed.

4. Use non-maximal suppression in the pixels edges in M_T obtained above to thin the edge ridges.
5. Using two different thresholds τ_1 and τ_2 (where $\tau_1 < \tau_2$) obtain two binary images I_1 and I_2. Note that I_2 with greater τ_2 has less noise and fewer false edges but

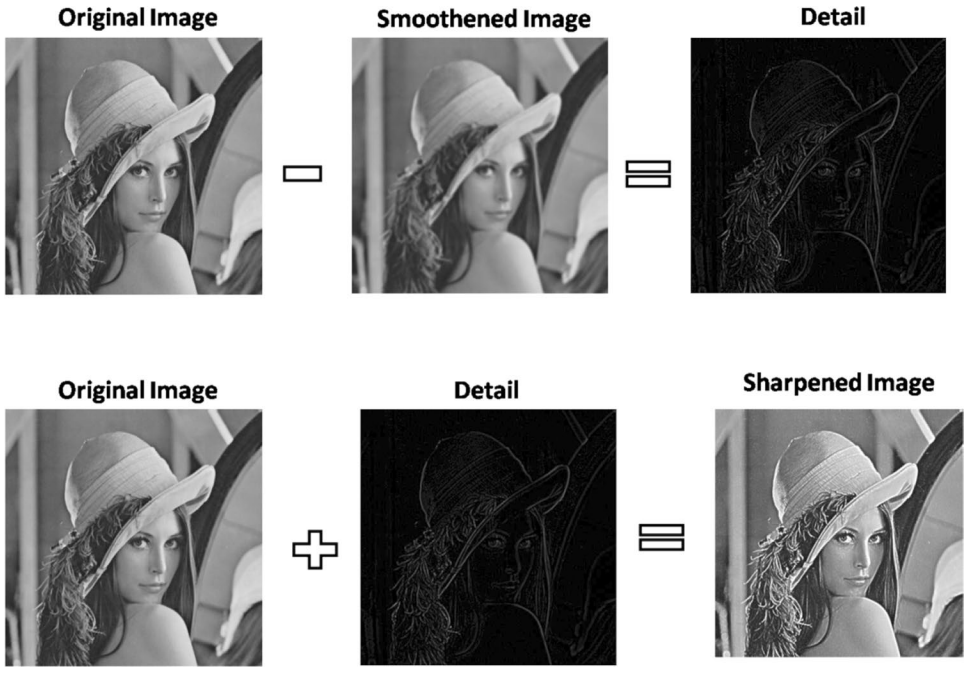


Fig. 2 Demonstration of image sharpening (Lena)

Fig. 3 Edge detected image using Canny method (Lena)

greater gaps between edge segments, when compared to I_1 with smaller τ_1.

6. Link edge segments in I_2 to form continuous edges (Fig. 3).

The canny edge detection is considered as one of the better methods in edge detection. However, without post processing of detected edges performance of canny edge detection is not good, and suffers from large number of falsely accepted edges.

4 Proposed method

In this section proposed method is detailed. The various steps are detailed and utility of various processes are discussed.

4.1 Fuzzy expert system

Figure 4 depicts the basic architectural layout of a fuzzy expert system. In this work three methods defined as M_1, M_2 and M_3 are presented. In method M_1 on the input image fuzzy logic is directly applied to detect edges. In method M_2 the input image is first passes through the sharpening filter and after this fuzzy logic is used to detect edges. In method M_3 the input image is first passes through the sharpening filter and then for Gaussian filter and finally fuzzy logic is applied to detect edges. In fuzzy logic based edge detection first input image is fuzzified using fuzzy input and output membership function and then apply IF-ELSE rules using Mamdani fuzzy inference engine and at the output defuzzification is done to obtain crisp values to obtain desired results (Fig. 5). The detailed description of the fuzzy based system, image sharpening using guided filter and finally use of Gaussian kernel for noise removal is discussed.

In this work, we use triangular membership function for both input and output and defined as: $triangle(x;a,b,c) = \max\left(\min\left(\frac{x-a}{b-a}, \frac{c-x}{c-b}\right), 0\right)$ (Figs. 6, 7).

4.2 Fuzzy rules

Fuzzy rules are developed using human expert system, and edge and non-edge pixels are decided on the basis of neighbourhood pixels. For the edge detection a total of 30 rules are defined. The rules are developed on a 3×3 mask as shown in Fig. 1. In the rule designing Img (x, y) represents pixel position, '1' represents white pixel and '0' represents black pixel value. For 3 black and 5 white pixels in neighbourhood 4 rules are defined, while for 4 black and 4 white pixels in neighbourhood, 8 rules are defined, for 5 black and 3 white pixels in the neighbourhood and for 6 black and 2

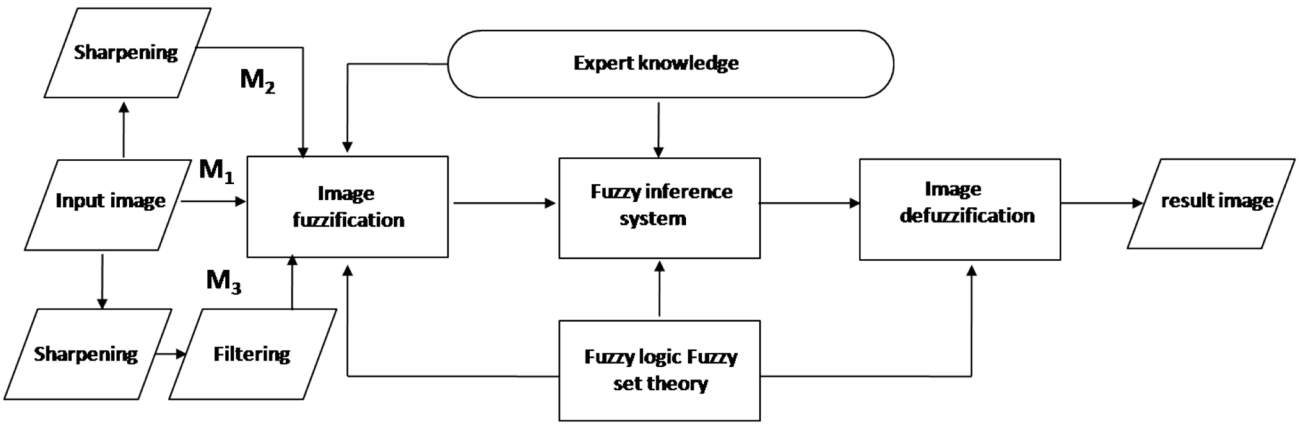


Fig. 4 Schematic of proposed edge detection mechanism (M_1: edge detection using fuzzy logic only, M_2: edge detection using fuzzy logic and sharpening filter and M_3: edge detection using fuzzy logic and sharpening filter)

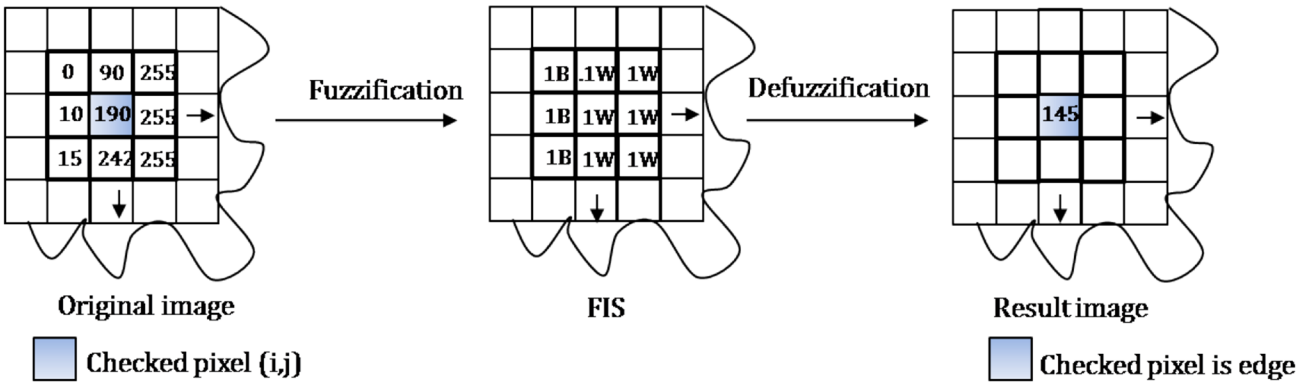


Fig. 5 Schematic view of fuzzy logic based edge detection

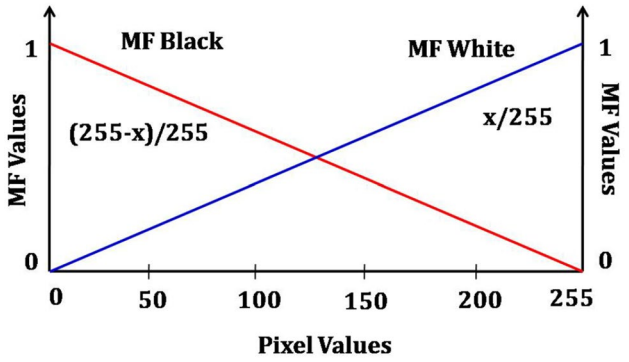


Fig. 6 Membership function for black and white pixels values

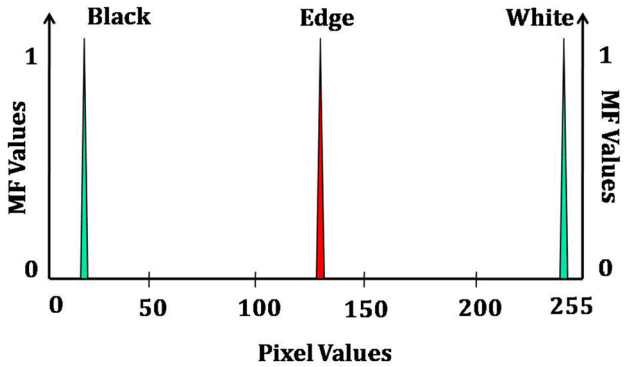


Fig. 7 Output membership function for black, white and edge

white pixels in the neighbourhood again for both the cases 8 rules are defined and finally for all black or white pixels in the neighbourhood 2 rules are defined. Various rules are detailed below (Fig. 8).

4.3 Guided image filtering and sharpening

In guided image filtering, the output image is linear transformation of guided image 'G' and is expressed as (He et al. 2013)

$$\hat{I}_n = a_i G_n + b_i, \quad \forall n \in w_i. \tag{5}$$

a_i and b_i are the co-efficient in window w_i. The guided image filtering problem can be formulated as the minimization of difference in input and output images, and ε is smoothness parameter and decides the degree of smoothness:

$$E(a_i, b_i) = \sum_{n \in w_i} ((a_i G_n + b_i - I_n)^2 + \varepsilon a_i^2). \tag{6}$$

The co-efficient are evaluated as

$$a_i = \frac{(1/|w|)\sum_{p \in w_i} G_n I_n - \bar{G}_i \bar{I}_i}{\sigma_i^2 + \varepsilon}$$
$$b_i = \bar{I}_i - a_i \bar{G}_i \tag{7}$$

where, bar represents the mean values of corresponding parameter. $|w|$ denotes the total number of pixel w_i:

$$\hat{I}_n = \left(\frac{1}{|w|}\sum_{i \in w_n} a_i\right) G_n + \left(\frac{1}{|w|}\sum_{i \in w_n} b_i\right). \tag{8}$$

In He et al. (2010) it is proved that GF can also be expressed as

$$W_{nm}^{GF}(G) = \frac{1}{|w|^2}\sum_{i:(n,m) \in w_i}\left(1 + \frac{(G_n - \bar{G}_i)(G_m - \bar{G}_i)}{\sigma_i^2 + \varepsilon}\right) \tag{9}$$

In general we have

$$\hat{I}_n = \sum_{m \in w_n} W_{nm}^{GF}(G) I_m. \tag{10}$$

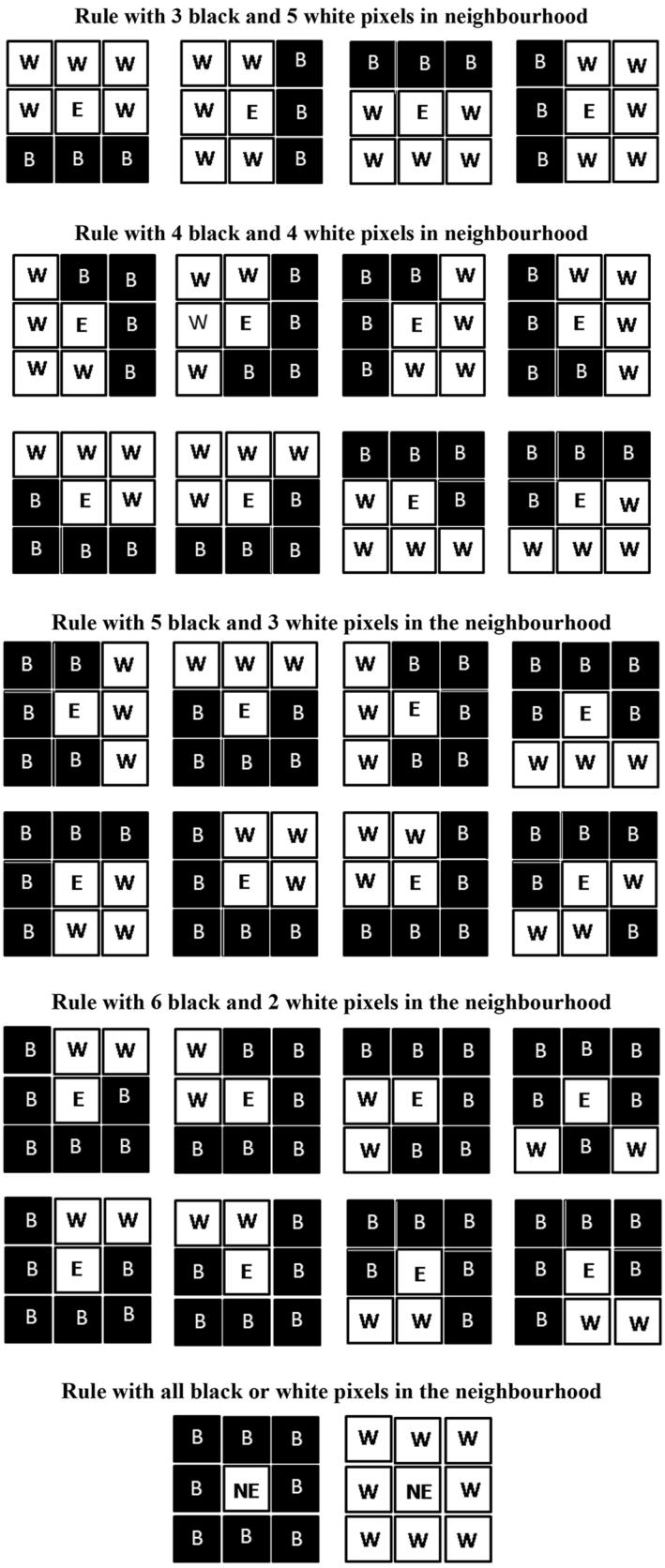


Fig. 8 Fuzzy rules black (B), white (W) and edge (E) and non-edge (NE)

The enhanced image can be written as

$$\hat{I}_s = \gamma(\hat{I}_n - G_n) + G_n \tag{11}$$

$$\hat{I}_s = \gamma[a_i G_n + b_i - G_n] + G_n \tag{12}$$

$$\hat{I}_s = [\gamma(a_i - 1) + 1]G_n + kb_i \tag{13}$$

The Gaussian 5×5 kernel is applied as filter for the removal of noise, with kernel as

$$\frac{1}{159}\begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix}. \tag{14}$$

The Gaussian kernel is used in method M_3 only.

5 Results

This section presents the edge detection results under the considered method. The simulation is performed in MATLAB software. For performance evaluation various edge detection measures are considered as described below:

5.1 Edge detection performance measures

It is not easy to define the general-purpose evaluation for edge detection, as some edges are missed, some are falsely accepted, and some alters their positions. Before going into detail of edge detection method, let us consider basic properties and concept of edge detection. Throughout the paper we have defined, ground truth image as (I_{gt}) and edge detected image as (I_{ed}). Let the quality measure is denoted by Q and property as (P) then following properties must be satisfied:

Symmetry (P_1): $Q(I_{gt}, I_{ed}) = Q(I_{ed}, I_{gt})$

Ideal solution (P_2): $I_{gt} = I_{ed}$

Sensitivity to noise (P_3): if any pixel (p) does not belong to ground truth (I_{gt}) or edge detected image (I_{ed}) i.e., $p \notin (I_{gt} \cup I_{ed})$, then

$$Q(I_{gt}, I_{ed}) < Q(I_{ed} \cup \{p\}, I_{gt})$$

Sensitivity to improvement (P_4): if $p \in I_{gt}$ and $p \notin I_{ed}$, then

$$Q(I_{gt}, I_{ed}) < Q(I_{ed} \cup \{p\}, I_{gt})$$

First property is self explanatory, second property state that there is only one optimal solution. Third and fourth properties indicate that inclusion of correct and in-correct pixels will decrease or increase error respectively. However, these four properties are not good enough to describe edge detection as to include falsely excluded important edge sometime extra pixels are accepted. As considering $p \notin I_{gt}$ and to include p in I_{ed} will increase un-avoidable error. In edge detection the main requirements are the correct identification of edge pixels and also the correct location of edge pixels.

We classify the positives pixels as edges when the classification is binary. We can classify pixels edge into four unique classifications with the condition that ground truth is present. These four classes are: true positive (TP), false positive (FP), true negative (TN) and false negative (FN). Due to the fact that not a large portion of the pixels are edges, there will be an issue of imbalanced binary classification (Chawla et al. 2004), where the dominating class will be the negative class.

The issues related with the position of pixel can be solved with the help of the spatial tolerance during the matching of the edge pixels as the factor, which decides whether the pixel classification is accurate or not, is the least alteration in the position of the pixel. Normally, we cannot assume an edge pixel as fully accurate or not, relying upon the way that it is found at t or t+1 pixels from the true edge. According to the Liuand Haralick, a quality measure is exact in the event of small changes in value reflects small variations in the detect (Liu and Haralick 2002). The measures such as F-score, and χ^2 test, are precise when pixels positions are exact however fall flat when pixels position changes. These issues prompted penalizing an edge pixel relying upon its separation to a true edge, which encourages the concept of distance based Edge Measures (EMs).

5.2 Correlation coefficient

The correlation coefficient is obtained by means of the equation (Van Heel 1987)

$$|\rho| = \frac{\text{cov}[I_{gt}, I_{ed}]}{\sqrt{Var[I_{gt}]Var[I_{ed}]}}. \tag{15}$$

Under perfect matching $|\rho|$ is one theoretically lowest value is zero. This metric satisfies properties P_1 and P_2.

5.3 Pratt's figure of merit (FoM)

The Pratt's figure of merit (FoM) evaluates edge location exactness in edge detected image in comparison to ground truth image, by measuring the displacement of edge points that are detected from an ideal edge. The FoM is characterized by (Wesolkowski et al. 2000)

$$FoM = \frac{1}{\max(I_{gt}, I_{ed})} \sum_{i=1}^{I_{ed}} \frac{1}{1 + \mu d^2(p, I_{gt})} \tag{16}$$

Here, I_{gt} is ideal edge points (ground truth), I_{ed} is edge points detected, d is displacement of detected edges from ideal edges, μ is scaling constant.

It is essential to note that these measurements binarize information before assessing images; this implies assessment is done over images that have lost data. The above mentioned metrics return esteems in the vicinity of 0 and 1, where 0 would imply that we have no similarity between detected image and reference image, and 1 implying that high closeness was detected, or in other words, each of the pixels present in one image edges are recognized at the same place in other image.

5.4 Structure similarity image metrics (SSIM)

SSIM completes a greatly improved activity at measuring subjective image quality in comparison to MSE or PSNR. At a high state, SSIM endeavors to estimate the adjustment in luminance, contrast, and structure in a picture. The SSIM is given by (Hore and Ziou 2010)

$$SSIM(I_{gt}, I_{ed}) = \frac{(2\mu_{gt}\mu_{ed} + k_1)(2\sigma_{gted} + k_2)}{(\mu_{gt}^2 + \mu_{ed}^2 + k_1)(\sigma_{gt}^2 + \sigma_{gt}^2 + k_2)} \tag{17}$$

where, μ is mean, σ^2 is variance, σ is cross-correlation term and rest terms are fixed constants.

5.5 Hausdorff distance (HoD)

Considering two images $I_{gt}=\{a_1, \ldots,a_n\}$ and $I_{ed}=\{b_1, \ldots,b_n\}$, the Hausdoff distance calculated as (Ma and Grimson 2005):

$$\begin{aligned} &H(I_{gt}, I_{ed}) = \max(d(I_{gt}, I_{ed}), d(I_{ed}, I_{gt})) \\ &\text{Where} \quad d(I_{gt}, I_{ed}) = \max_{a \in I_{gt}} \min_{b \in I_{ed}} ||a - b|| \end{aligned} \tag{18}$$

The function $d(I_{gt}, I_{ed})$ is the directed Hausdorff distance from I_{gt} to I_{ed}. This method is based on distance among the points, and lesser distance means more closeness between the images.

5.6 Euclidiean distance (E_D)

The Euclidian distance between two images is evaluated as (Nadernejad et al. 2008)

$$E_D(I_{gt}, I_{ed}) = \frac{1}{ij} \sum_{m=0}^{i-1} \sum_{n=0}^{j-1} \left| I_{gt}(m,n) - I_{ed}(m,n) \right|^2 \tag{19}$$

5.7 Average point-to-set distances (D_K)

The average distance from the edge pixels in the image under consideration to those in the ground truth is obtained as (Peli and Malah 1982)

$$D_K = \frac{1}{I_{ed}} \sqrt[K]{\sum_{p \in I_{ed}} d^K(p, I_{gt})} \tag{20}$$

We have considered only one image Lena along-with the ground truth image which contain ideal edge to cover vast varieties of results.

5.8 Baddeley's delta metric (BDM)

Baddeley's delta metric (BDM) is modified form of the Hausdorff distance (Ma and Grimson 2005). It is based on the distance between the elements of each of the sets, and expressed as (Lopez-Molina et al. 2017)

$$\mathrm{B}_f^K = \left[\frac{1}{|P|} \sum_{p \in P} \left| fd(p, I_{gt}) - fd(p, I_{ed}) \right|^K \right]^{1/K}, \tag{21}$$

f is a function which is concave in nature, and modulates the point to point distance.

In Fig. 9a original Lena image is shown, while varying patch radius from r = 16,32,64 and 128, while keeping ε to a fixed level of 0.01 and γ equals 5 the obtained results are shown in Fig. 9b–e. the size of local patch is considered to be $(2r+1)\times(2r+1)$ For smaller radius, block size is smaller, and sharpening takes place in local patches, therefore edges becomes sharper but discontinuity points increases thus sharpness quality is not good, as we increase the radius, the sharpness in image is relatively more uniform, and for $r=64$ and 128 the sharpness in images is of good quality.

The above experiment is repeated again with $\varepsilon = 0.001$, while keeping other parameters fixed, the effect of regularization parameter can be observed the mean variance of Lena image under consideration is 0.0256 therefore the effect of regularization parameter (ε) is negligible and image quality does not alter significantly.

In Fig. 10, image enhancement for various values of γ is shown, here as γ increases the sharpness increases, but for larger values of γ the quality of the images changes considerably as compared to original image. Moreover a significant variation in the colours is also observed. It can be observed (marked region in original image) from the figures that the first with the rise in γ edges becomes sharper thereafter they start to diminish, therefore it can also be concluded that both ε and γ have significant effect on the overall quality of the sharpened images.

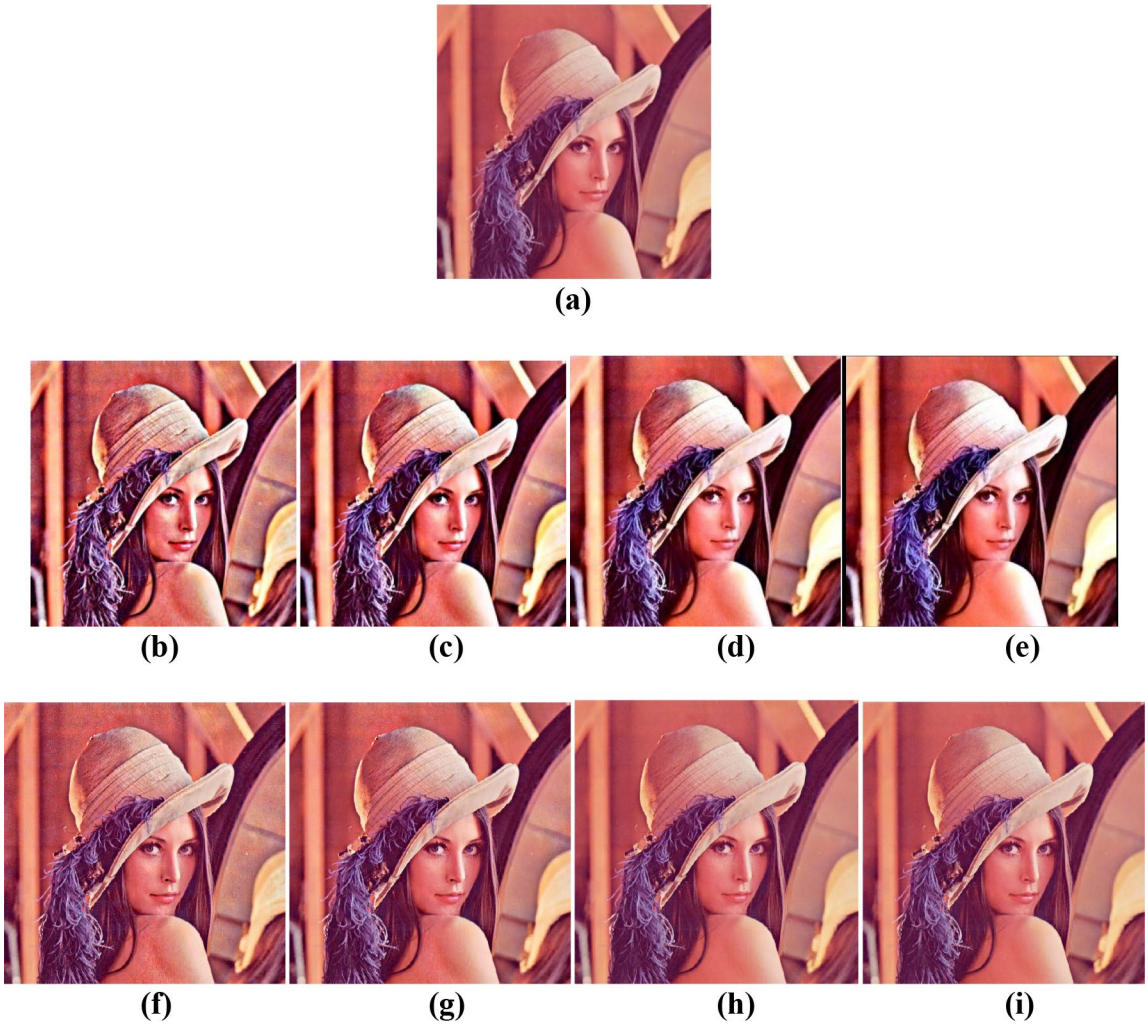


Fig. 9 Image Sharpening using guided filtering (k constant). **a** Original image. **b** $r=16$, $\varepsilon=0.01$, $\gamma=5$. **c** $r=32$, $\varepsilon=0.01$, $\gamma=5$. **d** $r=64$, $\varepsilon=0.01$, $\gamma=5$. **e** $r=128$, $\varepsilon=0.01$, $k=5$. **f** $r=16$,$\varepsilon=0.001$, $\gamma=5$. **g** $r=32$, $\varepsilon=0.001$, $\gamma=5$. **h** $r=64$, $\varepsilon=0.001$, $\gamma=5$. **i** $r=128$, $\varepsilon=0.001$, $\gamma=5$

In Fig. 11, results are presented for methods M_1 and M_2. In Fig. 11a edge detected image using fuzzy logic is shown for Lena image. In Fig. 11b–h results are shown for various values of γ as in Fig. 10. Consider the mark region in Fig. 10a, it is clear from Fig. 11a that fuzzy logic method in alone fail to capture edges in this region. Similarly around hat area noise is also added. As we increases the value of γ from 5 to 100, the noise around hat area diminishes but just above the mark area noise start to increases. These changes are shown in Fig. 12.

In Fig. 13, results are presented for methods M_3. It is clear from the figure that using filtering noise is suppressed significantly, and detected edges are much better as compared to shown in Fig. 11.

The discussion made above is based on human visualization. However, shifting in pixel positions cannot be identified using human vision system. Moreover, the regularization term also lead to the shifting of pixels. Therefore, distance based parameters for quality measurements provide the correctness of the chosen methods. In our results we have not done any further processing on detected image to clearly visualize the effect of each method.

In Table 3, listing of parameters results are shown, FoM is least for canny detection with value of 0.297 and it is best for $\gamma=5$, equals 0.483, SSIM for Canny is 0.451 and for $\gamma=20$, equals 0.560, while HoD is best for $\gamma=50$, and equals to 4.44/7, the Euclidian distance is minimum for canny method, which is an expected results as large number of edges are detected by canny method. BDM is least for $\gamma=5$, while it is at maximum for fuzzy method. D_K is best for Canny method. Finally the co-relation co-efficient is best for $\gamma=30, 50$. Therefore it can be inferred from the table that the performance of the fuzzy method in alone is poorest and among the chosen parameters for $\gamma=5$, the obtained results are better than other considered methods. In Table 4, results are compared for methods M_1 and M_3. Obtained results have

Fig. 10 Image sharpening using guided filtering (γ varying). **a** Original Lena image. **b** $r=64$, $\varepsilon=0.01$, $\gamma=5$. **c** $r=64$, $\varepsilon=0.01$, $\gamma=10$. **d** $r=64$, $\varepsilon=0.01$, $\gamma=20$. **e** $r=64$, $\varepsilon=0.01$, $\gamma=30$. **f** $r=64$, $\varepsilon=0.01$, $\gamma=50$. **g** $r=64$, $\varepsilon=0.01$, $\gamma=75$. **g** $r=64$, $\varepsilon=0.01$, $\gamma=100$

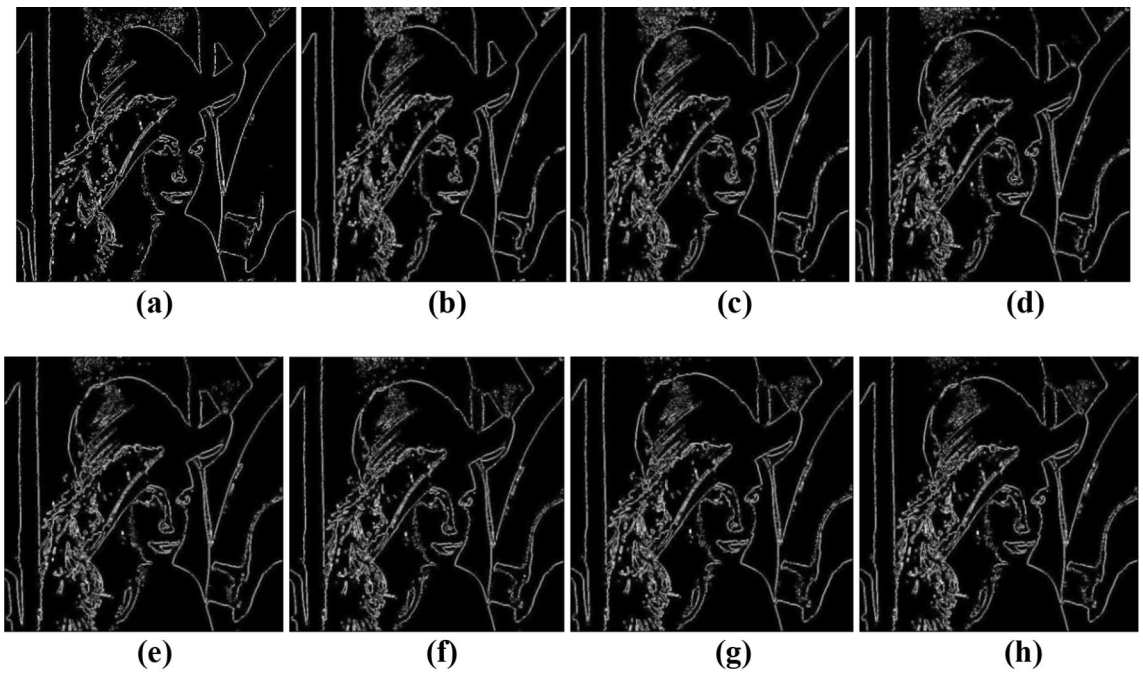


Fig. 11 Edge detected images **a** fuzzy logic, **b** M_2, $\gamma=5$, **c** M_2, $\gamma=10$, **d** M_2, $\gamma=20$, **e** M_2, $\gamma=30$, **f** M_2, $\gamma=50$, **g** M_2, $\gamma=75$, **h** M_2, $\gamma=100$

shown similar trend but obtain results are much better in comparison to the results in Table 3. In Table 4, FoM is maximum for $\gamma=20$, SSIM is maximum for $\gamma=30, 50$. HoD is maximum for $\gamma=50$. Still these performance measures are indicative, and do not provide clear information about exact edge detection mechanism. Therefore, in most of the recent research (Begol and Maghooli 2011; Mehrara et al. 2009; Zhang et al. 2009; Mathur and Ahlawat 2008; Marmanis

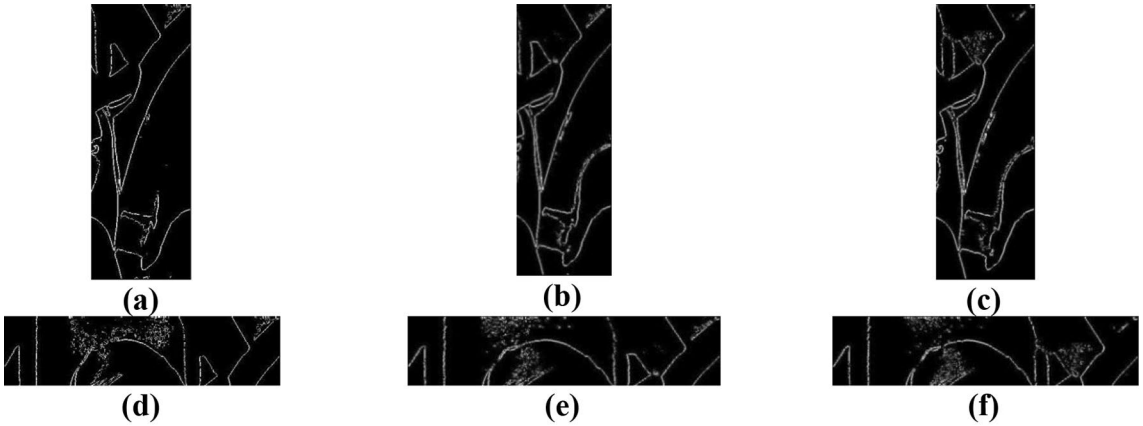


Fig. 12 Image segment with visible variations

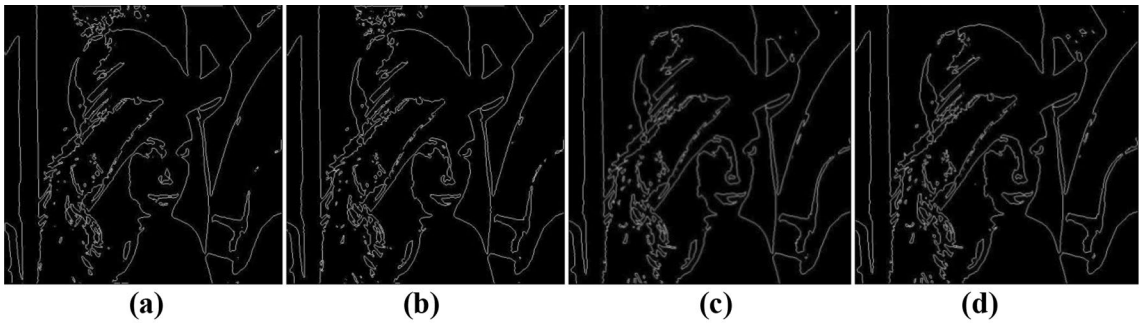


Fig. 13 Edge detected images **a** M_3, $\gamma = 5$, **b** M_3, $\gamma = 10$, **c** M_3, $\gamma = 50$, **d** M_3, $\gamma = 100$

Table 3 Comparison of parameters for methods M_1 and M_2

Methods	Performance metrics						
	FoM	SSIM	HoD (Avg./max)	E_D	BDM	D_K	ρ
Canny	0.297	0.451	6.69/9.21	**0.0048**	10.74	**0.06**	0.225
Fuzzy	0.398	0.519	5.31/8.12	0.0059	12.24	0.08	0.146
$\gamma = 5$	**0.483**	0.559	4.67/7	0.0056	**9.40**	0.111	0.21
$\gamma = 10$	0.472	0.557	4.59/6.78	0.0056	9.48	0.113	0.218
$\gamma = 20$	0.469	**0.560**	4.47/6.86	0.0055	10.03	0.11	0.229
$\gamma = 30$	0.467	0.559	4.45/7.42	0.0055	10.02	0.112	**0.232**
$\gamma = 50$	0.471	0.556	**4.44/7**	0.0055	9.98	0.102	**0.232**
$\gamma = 75$	0.467	0.557	4.46/7.42	0.0055	9.85	0.102	0.231
$\gamma = 100$	0.469	0.555	4.47/7.21	0.0055	9.93	0.106	0.231

Best obtained values are marked in bold

et al. 2018; Farbod et al. 2018; Setayesh et al. 2013) human visual system (HVS) is used to characterize edge and non-edge pixels.

The above results are obtained on Lena image, detected edges are compared with ground truth image and results are compared in terms of various performance measures. However, to prove usefulness of proposed method results are obtained other database images. Image shows in Fig. 14, is taken from Berkley Segmentation Database (https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/BSDS300/html/dataset/images.html), while in image in Fig. 15, and is from USC-SIPI Image Database (http://sipi.usc.edu/database/database.php). Considering, Fig. 14, fuzzy logic based method fail to detect edge on west-south corner of the image.

Table 4 Comparison of parameters for methods M_1 and M_3

Methods	Performance metrics						
	FoM	SSIM	HoD (Avg/max)	E_D	BDM	D_K	ρ
Canny	0.297	0.451	6.69/9.21	**0.0048**	10.74	**0.06**	**0.225**
Fuzzy	0.398	0.519	5.31/8.12	0.0059	12.24	0.08	0.146
γ = 5	0.470	0.581	4.15/10.44	0.0056	9.92	0.120	0.186
γ = 10	0.475	0.582	4.17/9.64	0.0056	9.69	0.110	0.198
γ = 20	**0.497**	0.590	4.13/8.43	0.0055	9.05	0.093	0.211
γ = 30	0.496	**0.595**	4.09/7.09	0.0055	8.81	0.088	0.217
γ = 50	0.495	**0.595**	**4.09/6.86**	0.0057	8.59	0.079	0.216
γ = 75	0.489	0.593	4.11/6.63	0.0057	8.65	0.087	0.217
γ = 100	0.493	0.593	4.12/6.56	0.0057	**8.57**	0.081	0.219

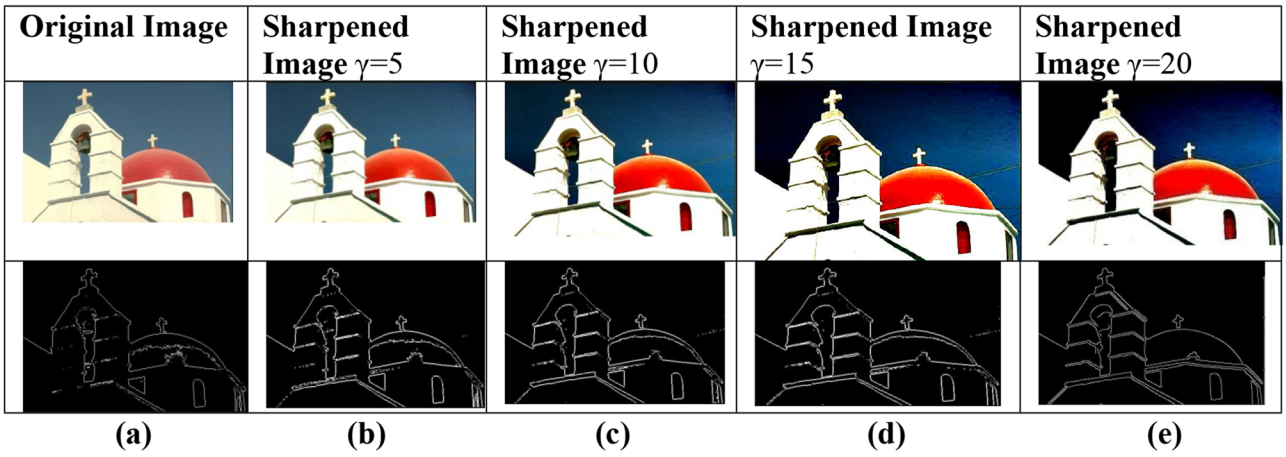


Fig. 14 BSD edge detected images **a** fuzzy logic, **b** M_3, γ = 5, **c** M_3, γ = 10, **d** M_3, γ = 15, **e** M_3, γ = 20

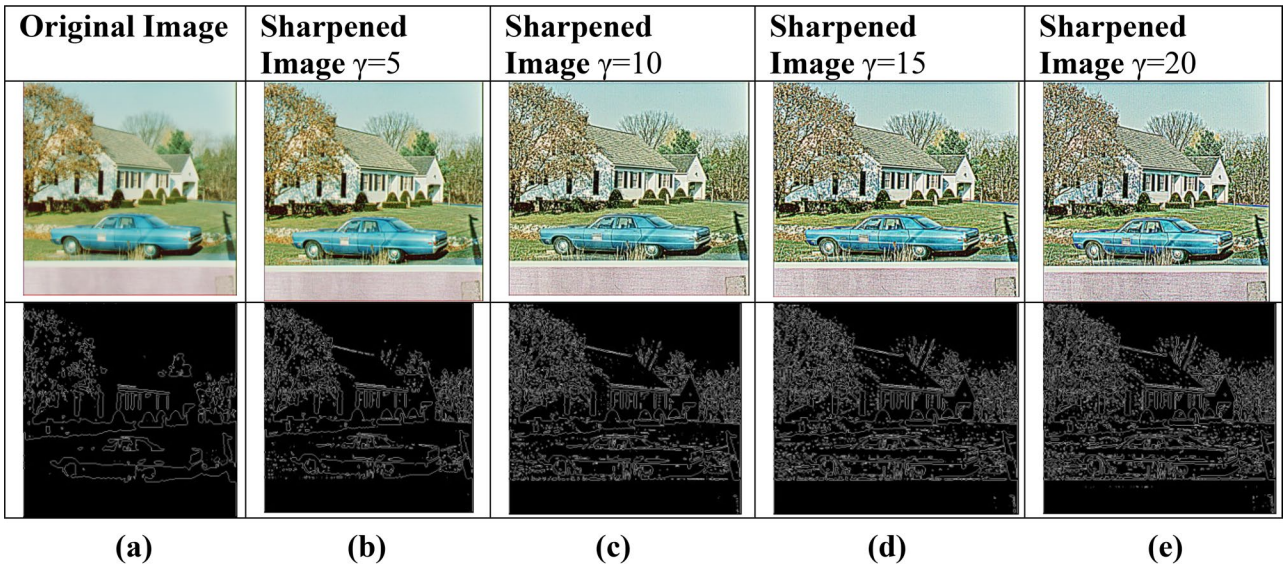


Fig. 15 USC-SIPI edge detected images **a** fuzzy logic, **b** M_3, γ = 5, **c** M_3, γ = 10, **d** M_3, γ = 15, **e** M_3, γ = 20

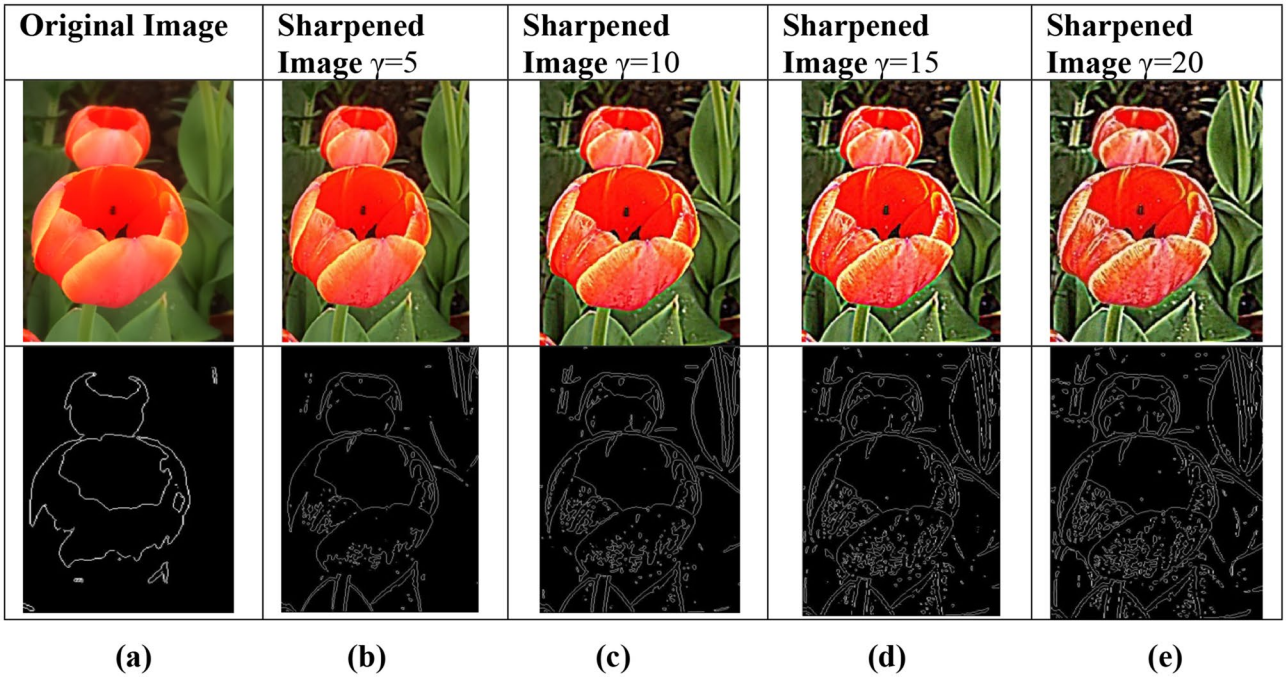


Fig. 16 Tulip edge detected images **a** fuzzy logic, **b** M_3, $\gamma = 5$, **c** M_3, $\gamma = 10$, **d** M_3, $\gamma = 15$, **e** M_3, $\gamma = 20$

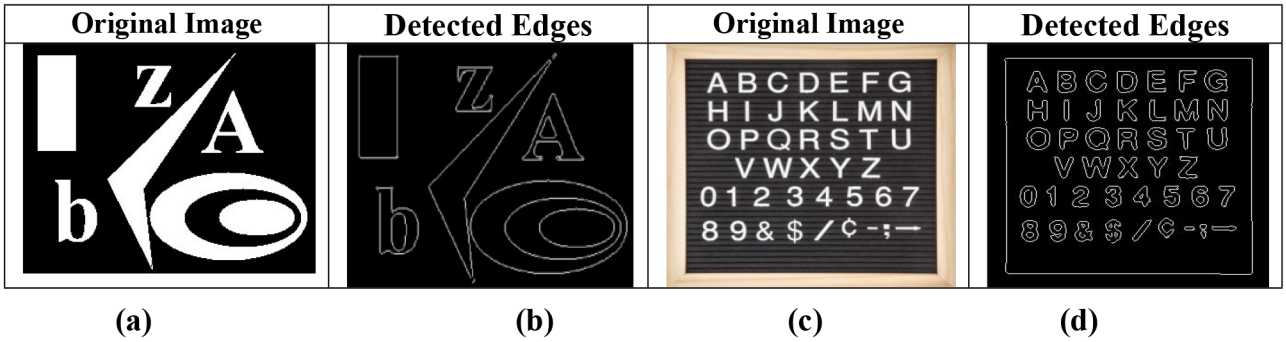


Fig. 17 Edge detection using fuzzy logic **a** original image, **b** detected edges, **c** original image, **d** detected edges

However, from (b) to (e) it is clear that as we increase sharpness the edges can be detected more correctly. From Fig. 15, it can be visualized that as sharpness increases; image detail is more clearly visible. Same effect is also obtained in detected edges.

In Fig. 16, results are obtained for tulip images and obtained response show similar trend as in Figs. 14 and 15. It is also clear from Figs. 15 and 16, sometime more sharpness leads to the generation of noise; therefore image should be sharpened to a level where effect of noise is minimal. However, if this noise is dominant, than this additional noise can be suppressed using filters as used in other edge detection methods (Shah 1997).

In the above discussed results it is found that fuzzy logic based method fails to detect some of the edges. Therefore, to check the validity of the considered fuzzy logic structure, in Fig. 17 results are generated for two images [(a) and (c)] where, edges are clearly visible, and it is found that considered fuzzy logic structure correctly detect all the edges in both the images. Thus, in case of clear edges images fuzzy logic structure gives accurate results.

In Fig. 18, animal alphabet image is taken which has more complex edges as compared to images considered in Fig. 17. Fuzzy logic based method fails to detect 'Tiger' and 'Orangutan' shapes and also letter mark on the animals are not detected. As we increase sharpness edges and letters are detected more clearly. In Fig. 19, zoomed version of fuzzy logic and sharpened image $\gamma = 20$ is shown, the name of animal is not clearly detected with fuzzy design, while with proposed method edges of both animal name and letter mark

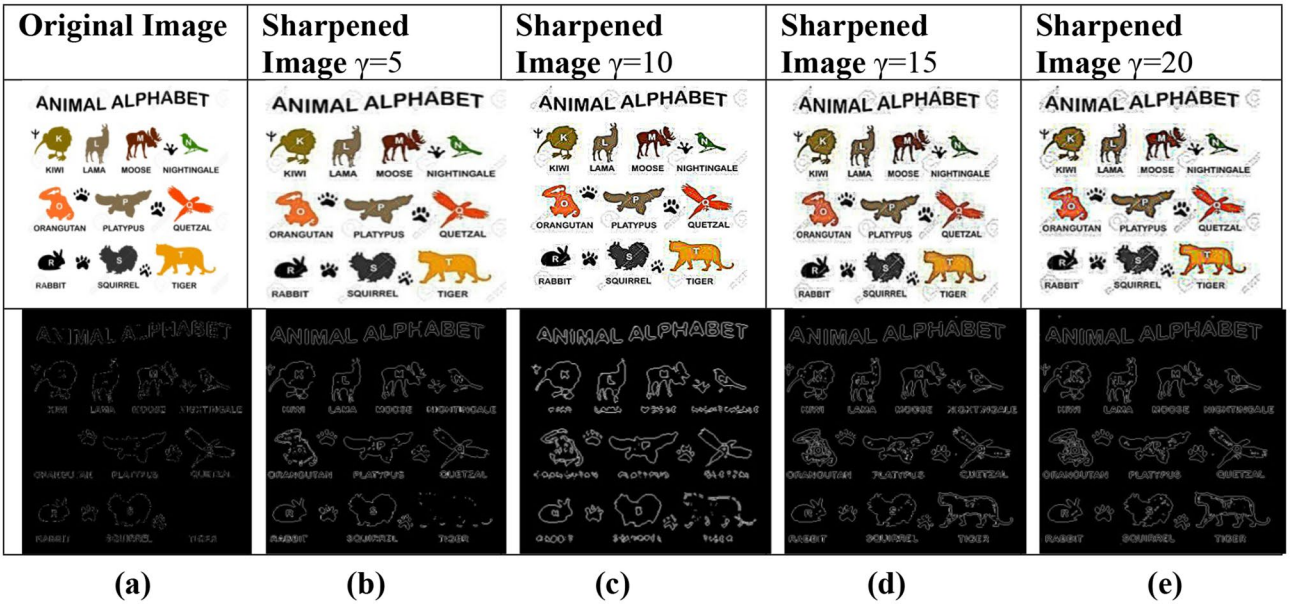


Fig. 18 Animal alphabet image **a** fuzzy logic, **b** M_3, $\gamma = 5$, **c** M_3, $\gamma = 10$, **d** M_3, $\gamma = 15$, **e** M_3, $\gamma = 20$

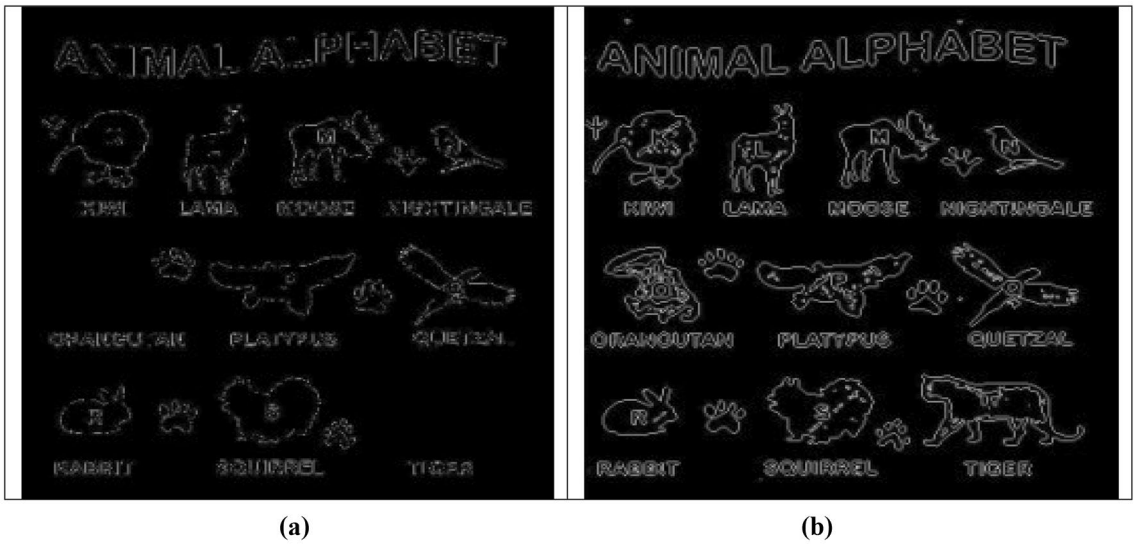


Fig. 19 Zoomed version of **a** Fig. 18a, **b** Fig. 18e

on the animals are detected. With our design claws and paws edges are clearly visible.

Finally, we can conclude that the proposed method is superior to fuzzy logic based method, and it overcomes the limitations of fuzzy logic structure. The membership function considered in this work is same as in previous works (Angelov et al. 2011; Sargano et al. 2017; Rogowska 2000; Rangarajan 2005; Alshennawy and Aly 2009; Aborisade 2011). However, to overcome the limitations of fuzzy logic method, further research can be carried out in developing fuzzy rules which are developed using human expert system and in choosing appropriate membership function.

6 Conclusions

This paper presents an edge detection method based on fuzzy logic, sharpening and filtering. The main aim of the paper is to design an edge detection method where edges can be controlled without deteriorating the considered image. It has been found that fuzzy logic method in alone falsely rejected some of the edges, noise is also added. To deal with this we have shown that sharpening of image can be done, which improves the results significantly. It is also shown that sharpening itself depends on parameters r, ε and γ, and these parameters should be chosen efficiently to get desired results. It is also notable that the regularization parameter (ε) should be kept within the sub-range of image variance as regularization parameter shift the pixel positions. It is also shown that noise generated due to fuzzy process can be significantly brought down by using Gaussian filter. The obtained results are compared with various statistical measures and it has been found that proposed methods M_2 and M_3 performs better in comparison to method M_1.

References

Aborisade DO (2011) Novel fuzzy logic based edge detection technique. Int J Adv Sci Technol 29(1):75–82

Alshennawy AA, Aly AA (2009) Edge detection in digital images using fuzzy logic technique. World Acad Sci Eng Technol 51:178–186

Angelov P, Zhou X (2008) On line learning fuzzy rule-based system structure from data streams. In: 2008 IEEE international conference on fuzzy systems (IEEE world congress on computational intelligence). IEEE, pp 915–922

Angelov P, Sadeghi-Tehran P, Ramezani R (2011) An approach to automatic real-time novelty detection, object identification, and tracking in video streams based on recursive density estimation and evolving Takagi-Sugeno fuzzy systems. Int J Intell Syst 26(3):189–205

Begol M, Maghooli K (2011) Improving digital image edge detection by fuzzy systems. World Acad Sci Eng Technol 81:76–79

Canny J (1986) A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 6:679–698

Chawla NV, Japkowicz N, Drive P (2004) Editorial: special issue on learning from imbalanced data sets. ACM SIGKDD Explor Newsl 6(1):1–6

Cosgrove C, Yuille AL (2019) Adversarial examples for edge detection: they exist, and they transfer. arXiv:1906.00335

Farbod M, Akbarizadeh G, Kosarian A, Rangzan K (2018) Optimized fuzzy cellular automata for synthetic aperture radar image edge detection. J Electron Imaging 27(1):013030

Gevers T, Gijsenij A, Van de Weijer J, Geusebroek JM (2012) Color in computer vision: fundamentals and applications, vol 23. Wiley, New York

Gonzalez CI, Melin P, Castro JR, Castillo O, Mendoza O (2016) Optimization of interval type-2 fuzzy systems for image edge detection. Appl Soft Comput 47:631–643

Goswami JC, Chan AK (2011) Fundamentals of wavelets: theory, algorithms, and applications, vol 233. Wiley, New York

He K, Sun J, Tang X (2010) Guided image filtering. In: European conference on computer vision. Springer, Berlin, pp 1–14

He K, Sun J, Tang X (2013) Guided image filtering. IEEE Trans Pattern Anal Mach Intell 6:1397–1409

Hore A, Ziou D (2010) Image quality metrics: PSNR vs. SSIM. In: Pattern recognition (ICPR), 2010 20th international conference on. IEEE, pp 2366–2369

Jeong C, Yang HS, Moon K (2019) A novel approach for detecting the horizon using a convolutional neural network and multi-scale edge detection. Multidimens Syst Signal Process 30(3):1187–1204

Liu G, Haralick RM (2002) Optimal matching problem in detection and recognition performance evaluation. Pattern Recognit 35(10):2125–2139

Liu Y, Lew MS (2016) Learning relaxed deep supervision for better edge detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 231–240

Lopez-Molina C, Ayala-Martini D, Lopez-Maestresalas A, Bustince H (2017) Baddeley's delta metric for local contrast computation in hyperspectral imagery. Prog Artif Intell 6(2):121–132

Ma X, Grimson WEL (2005) Edge-based rich representation for vehicle classification. In: Computer vision, 2005. ICCV 2005. Tenth IEEE international conference on, vol 2. IEEE, pp 1185–1192

Marmanis D, Schindler K, Wegner JD, Galliani S, Datcu M, Stilla U (2018) Classification with an edge: improving semantic image segmentation with boundary detection. ISPRS J Photogramm Remote Sens 135:158–172

Mathur S, Ahlawat A (2008) Application of fuzzy logic in image edge detection. In: International Conference "Intelligent Information and Engineering Systems"

Mehrara H, Zahedinejad M, Pourmohammad A (2009) Novel edge detection using BP neural network based on threshold binarization. In: Computer and electrical engineering, 2009. ICCEE'09. Second international conference on, vol 2. IEEE, pp 408–412

Nadernejad E, Sharifzadeh S, Hassanpour H (2008) Edge detection techniques: evaluations and comparisons. Appl Math Sci 2(31):1507–1520

Peli T, Malah D (1982) A study of edge detection algorithms. Comput Gr Image Process 20(1):1–21

Rangarajan S (2005) Algorithms for edge detection. Stony Brook University. http://www.ee.sunysb.edu/~cvl/ese558/s2005/Reports/Srikanth%20Rangarajan/submission.doc. Accessed 10 Apr 2019

Ren H, Zhao S, Gruska J (2018) Edge detection based on single-pixel imaging. Opt Express 26(5):5501–5511

Rogowska J (2000) Overview and fundamentals of medical image segmentation. In: Handbook of medical imaging, processing and analysis. Academic Press, Inc., Orlando, FL, USA, pp 69–85

Sargano A, Angelov P, Habib Z (2017) A comprehensive review on handcrafted and learning-based action representation approaches for human activity recognition. Appl Sci 7(1):110

Setayesh M, Zhang M, Johnston M (2013) A novel particle swarm optimisation approach to detecting continuous, thin and smooth edges in noisy images. Inf Sci 246:28–51

Shah M (1997) Fundamentals of computer vision. University of Central Florida, Orlando

Shih FY (2010) Image processing and pattern recognition: fundamentals and techniques. Wiley, New York

Van Heel M (1987) Similarity measures between images. Ultramicroscopy 21(1):95–100

Wesolkowski S, Jernigan ME, Dony RD (2000) Comparison of color image edge detectors in multiple color spaces. In: Image processing, 2000. Proceedings. 2000 international conference on, vol 2, IEEE, pp 796–799

Zhang L, Xiao M, Ma J, Song H (2009) Edge detection by adaptive neuro-fuzzy inference system. In: Image and signal processing, 2009. CISP'09. 2nd international congress on. IEEE, pp 1–4

Zhang W, Zhao Y, Breckon TP, Chen L (2017) Noise robust image edge detection based upon the automatic anisotropic Gaussian kernels. Pattern Recognit 63:193–205

Zhou X, Angelov P (2007) Autonomous visual self-localization in completely unknown environment using evolving fuzzy rule-based classifier. In: 2007 IEEE symposium on computational intelligence in security and defense applications. IEEE, pp 131–138

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/327948653

Efficient Fluorescence Quenching of 5-Aminoquinoline: Silver ion recognition study

Article *in* Journal of Luminescence · September 2018

DOI: 10.1016/j.jlumin.2018.09.062

CITATIONS
0

READS
111

4 authors, including:

Nupur Pandey
DSB Campus Kumaun University Nainital
1 PUBLICATION **0** CITATIONS
SEE PROFILE

Mohan Singh Mehata
Delhi Technological University
65 PUBLICATIONS **681** CITATIONS
SEE PROFILE

Nisha Fatima
Kumaun University
1 PUBLICATION **0** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Synthesis and Photophysics of Nanoparticles/Quantum Dots using ps Time-resolved Techniques View project

Photoinduced Charge Transfer Dynamics of Synthesized Semiconducor Quantum Dots and Organic Molecules under the Influence of External Perturbation View project

All content following this page was uploaded by Mohan Singh Mehata on 18 June 2019.

The user has requested enhancement of the downloaded file.

Journal of Luminescence 205 (2019) 475–481

Contents lists available at ScienceDirect

Journal of Luminescence

journal homepage: www.elsevier.com/locate/jlumin

Efficient fluorescence quenching of 5-aminoquinoline: Silver ion recognition study

Nupur Pandey[a], Mohan Singh Mehata[b,*], Nisha Fatma[a], Sanjay Pant[a,*]

[a] Photophysics Laboratory, Department of Physics, Centre of Advance Study, DSB Campus, Kumaun University, Nainital 263002, India
[b] Laser Spectroscopy Laboratory, Department of Applied Physics, Delhi Technological University, Bawana Road, Delhi 110042, India

ARTICLE INFO

Keywords:
5-aminoquinoline
Ag^+ ions
Fluorescence quenching
Charge transfer
Optical sensor

ABSTRACT

Fluorescence (FL) quenching of 5-aminoquinoline (5AQ) by Ag^+ ions has been studied in acetonitrile (ACN) using steady-state and time-resolved fluorescence techniques. In the absence of Ag^+ ions, FL shows a single band assigned to the intramolecular charge transfer (ICT) state. The FL intensity of 5AQ gradually reduces on the successive addition of Ag^+ ions to the solution containing 5AQ. The decrease of FL intensity is nearly ~ 17 fold in the presence of 60 mM Ag^+ ions. Together with the decrease in FL intensity, the FL band maximum is red-shifted. The Stern-Volmer plot has been employed to determine the quenching constant. The nature of quenching has been stated based on the decrease of FL intensity and lifetime in the presence of metal ions. Also, the Benesi-Hildebrand relation has been used to explore the binding mechanism between 5AQ and Ag^+ ions. Thus, 5AQ displays high affinity for Ag^+ ions over other competitive metal ions (Cu^{2+}, Ca^{2+}, Pb^{2+}, Fe^{2+}, Li^+, K^+). Therefore, the present system emerges to be quite important in sensing of Ag^+ ions in an aprotic environment in a concentration range of 4 μM to 60 mM.

1. Introduction

Quinoline and its derivatives have been an intriguing class of molecules in the field of organic and pharmaceutical chemistry and widely used as a prototype compound to many drugs. Multifarious experimental and theoretical researches in photophysics and photochemistry of quinoline-based compounds have found potential applications in preparation of various drugs and dyes. Photoexcited relaxation processes of variously substituted quinolines especially amino- and hydroxyl-quinolines have been considered broadly over the past few years [1–18]. In recent literature, quinoline fused heterocyclic derivatives have been developed that possess diverse biological activities, i.e., anticancer, antimicrobial, antimycobacterial, anticonvulsant, anti-inflammatory and cardiovascular activities [19].

Aminoquinoline (AQ) is an imperative group of fluorophore that plays a momentous role in supramolecular chemistry, electron transfer process, charge transfer, host-guest systems and sensing applications. Photoexcitation to the lowest excited state is usually associated with intramolecular charge transfer (ICT) from amino group to the aromatic ring in these compounds hence producing the ICT state. Enol to keto conversion occurs via excited state proton transfer (ESPT) reaction following photoexcitation. The position of NH_2 group in the quinoline ring and the surrounded solvent molecules/polymers influences the excited state properties of aminoquinolines [5,8–15].

The FL quenching-based studies have been an area of fundamental interest from decades since this can be valuable in elucidating the localization of fluorophores in protein and membrane and their absorptivity to quenchers. Quenching can occur through various mechanisms and become progressively significant due to its applications in fluorescence sensing and drug delivery systems. An array of substances like, metal ions, halides, molecular oxygen, amines etc. can act as quencher [17,18]. Quinoline based fluorescent sensors are extensively used to determine the fluorescence intensity variations and shift of fluorescence wavelength [19–30]. As an example, when halide ions come in contact with excited state 6-methoxyquinoline and 8-methylquinoline, the FL intensity of these probe molecules gets quenched regularly and follows the diffusional quenching mechanism [16–18], and acts as an optical sensor.

In heterocyclic quinoline, the nitrogen atom can act as a chelating site towards metal ions [20]. Photoinduced electron transfer (PET), intramolecular charge transfer (ICT), twisted intramolecular charge transfer (TICT), fluorescence resonance energy transfer (FRET), chelation enhancement quenching effect (CHEQ) or chelation-enhanced fluorescence effect (CEFE) are important mechanisms of fluorescence sensing in the design of quinoline-based fluorescence sensors [20–30]. The development and design of a fluorescent sensor for the variety of

* Corresponding authors.
E-mail addresses: msmehata@dtu.ac.in (M. Singh Mehata), sanjaypant@kunainital.ac.in (S. Pant).

https://doi.org/10.1016/j.jlumin.2018.09.062
Received 8 July 2018; Received in revised form 23 September 2018; Accepted 27 September 2018
Available online 28 September 2018
0022-2313/ © 2018 Elsevier B.V. All rights reserved.

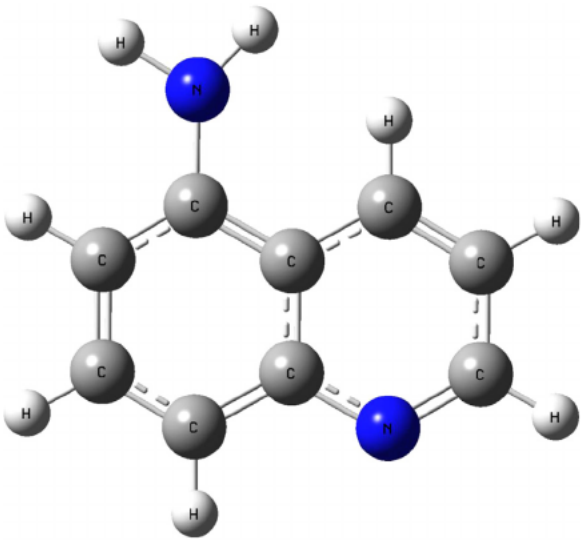
metal ions has acquired attention due to their essential role in medical, environmental and biological applications [22–30]. The FL of ligand may be quenched or enhanced by coordination depending on the influence of the metal ions on the non-radiative processes competing with the radiative process, i.e., FL. The FL resulting from the complex formation is a consequence of numerous factors; the relative position and behavior of the lower electronic excited state of organic ligands and those of metal ions, the heavy atom effects, the structural rigidity and the paramagnetic character of ions [31].

Silver ion, like other heavy metal ions; copper, zinc, lead, mercury, etc., is one of the indispensable toxic metal and has been broadly employed in chemistry, pharmacy, electrical industry, photographic and imaging industry [32–40]. The bioactive silver has gained interest in the domain of biology, such as antimicrobial activity, disinfectant, antibiotics and sterilization even at low concentration level [34]. However, in few cases, rapid absorption of Ag^+ ions causes bioaccumulation, toxicity and irreversible darkening of skin along with disruptive renal and neurological function [34,35]. Furthermore, Ag^+ ions can disable sulphydryl enzymes and bind with imidazole, amine, and carboxyl groups of different metabolites [35,37,41,42]. Since the d^{10} electronic configuration of Ag^+ ions is generally known as a fluorescence quencher following electron transfer and intersystem crossing processes, thus Ag^+ ions could be detected by chemosensors via fluorescence quenching mechanism [28]. Recent reports on fluorescent sensors for detection of Ag^+ ions comprise different fluorophores such as pyrene, oxadiazole, phenanthroline, coumarin, quinoline, gold nanoparticles etc [32–40].

In the present work, we have commenced a detailed study of FL quenching of 5AQ by Ag^+ ions, to comprehend the nature of possible quenching mechanism. Steady-state and time-resolved measurements confirm the quenching mechanism of probe 5AQ in the presence of Ag^+ ions. With the addition of Ag^+ ions, the FL maximum is red-shifted ($\sim$ 46 nm) with $\sim$ 17 fold quenching of fluorescence intensity. Interestingly, no fluorescence quenching and wavelength shift are noticed on addition of variant metal ions (Ca^{2+}, Cu^{2+}, Li^+, K^+, Fe^{2+}, Pb^{2+}). The molecular structure of 5AQ is given in Scheme 1.

2. Experimental section

5-Aminoquinoline (5AQ) purchased from Aldrich chemical company Inc., U.S.A. (with 99% purity) was used as received and matched with the reported FL spectrum to confirm its fluorescence purity (to make sure that there is no impurity). Silver nitrate and all other metal salts were obtained from Loba Chemie and SRL Pvt. Ltd. (with 98–99% of purity) and used as received. Spectroscopic or HPLC grade solvents were used for the present study. The purity of compounds was established by agreement of corrected excitation spectra with the electronic absorption spectra. The stock solution of $AgNO_3$ (0.1 M) was prepared from the stock solution of 5-aminoquinoline (10^{-4} M). From this solution, diluted solutions of $AgNO_3$ were prepared. This ensured that the probe concentration did not change on the addition of $AgNO_3$. Freshly prepared samples at room temperature were used in the present study. A similar method was used during preparation of diluted solutions of other metal ions.



Scheme 1. Molecular structure of 5-aminoquinoline.

2.1. Instrumentation

Steady-state absorption spectra were recorded by UV–visible Double beam spectrophotometer (Systronics Model no: 2203). Fluorolog-3 spectrofluorometer (Horiba Jobin Yvon Inc.) accompanied by available analysis software have been used to record and analyze the excitation and emission spectra. 450 W xenon lamp was used as the excitation source. Transparent quartz cuvette was used for the absorption and emission measurements and frontal geometry was used in recording emission to avoid inner filter effect. For all the measurements, excitation and emission slit widths were chosen about 2.0 nm, and the excitation wavelength was 345 nm. Time-correlated single photon counting (TCSPC) system (DeltaFlex-01-DD, Horiba Jobin Yvon IBH Ltd.) coupled with DeltaDiode (340 nm) and PMT (PPD 850) was used for fluorescence decay measurements and least square fitting and reconvolution methods were used to analyze the data. Instrumental response function was measured using Ludox scatterer. The numbers of counts in the peak channel were at least 10,000.

3. Results and discussion

Absorption, fluorescence (FL), FL-excitation spectra and FL-decays were measured for 5AQ in acetonitrile (ACN) in the absence and presence of varying concentration of silver ions (Ag^+) at room temperature.

Fig. 1 shows absorption spectra of 5AQ in ACN in absence and presence of various concentrations of Ag^+ ions. The absorption maximum appeared at 343 nm, and is ascribed to the $\pi \rightarrow \pi^*$ transition of quinoline group. In presence of a small concentration of Ag^+ ions, keeping the solute concentration fixed the optical density (absorbance) of 5AQ decreases, which is accompanied by broadening of the absorption spectrum. The corresponding optical density decreases from 0.65 to 0.45 with increasing concentration of Ag^+ ions from 0.0 to 7.0 mM. The change in spectral broadening and absorption maximum

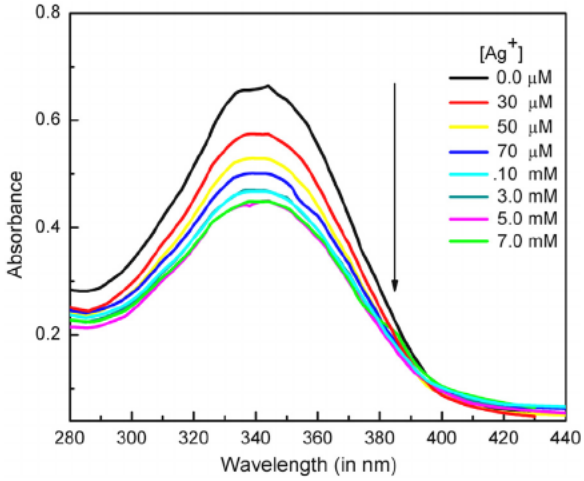


Fig. 1. Absorption spectra of 5AQ (1×10^{-4}M) in ACN in the presence of various concentration of Ag^+ ions.

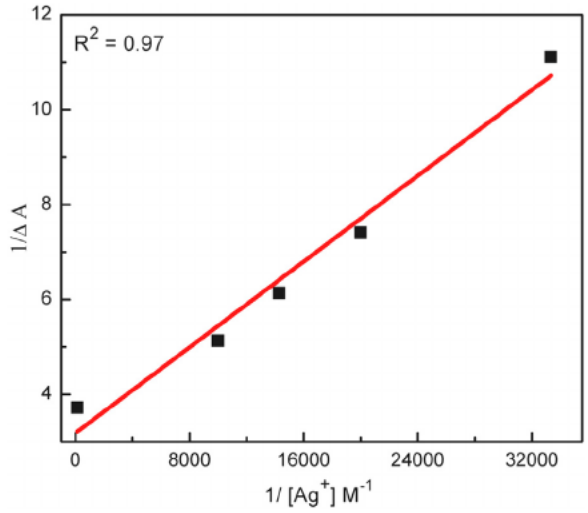


Fig. 2. Benesi-Hildebrand plot of 5AQ in ACN with increasing concentration of Ag^+ ions.

upon addition of Ag^+ ions suggests that there is a ground state complexation between 5AQ and Ag^+ ions. To understand the formation of complex between the probe and the quencher and to determine the binding constant, the well-known Benesi-Hilderband [32] Eq. (1) was applied.

$$\frac{1}{(A - A_0)} = \frac{1}{(A_{max} - A_0)} + \left[\frac{1}{K[M^{n+}]} + 1\right] \quad (1)$$

where A is the absorbance of 5AQ in presence of various concentration of Ag^+ ions, A_0 is the absorbance of 5AQ, A_{max} is the absorbance of 5AQ in the presence of the maximum concentration of Ag^+ ions, $[M^{n+}]$ is the concentration of metal ion and K is the binding constant. A plot between $1/\Delta A$ vs. the inverse of Ag^+ ions concentration is linear (Fig. 2) and demonstrates the 1:1 complex of 5AQ: Ag^+. The binding constant (K) determined from the ratio of intercept/slope of the plot comes out to be $1.407\times10^4 M^{-1}$.

Fig. 3(a,b) shows the fluorescence spectra of 5AQ in ACN in the absence and presence of varying concentration of Ag^+ ions at room temperature. The fluorescence maximum is around 482 nm upon excitation at 345 nm and is independent of excitation wavelength. The fluorescence of 5AQ reduced upon addition of small amount (about 4 μM) of Ag^+ ions. While increasing the concentration of Ag^+ ions from 4 μM, the fluorescence intensity is gradually decreased, as shown in Fig. 3(a). The FL intensity diminished nearly 17 folds in the presence of 60 mM concentration of Ag^+ ions. Thus, the FL of 5AQ is quenched significantly in the presence of Ag^+ ions in ACN. In addition to quenching, the FL spectrum is shifted towards red with the addition of Ag^+ ions. The red shift in the FL spectrum is about 46 nm for the highest concentration of Ag^+ ions taken in the study (60 mM), as shown in (Fig. 4). The excitation spectrum of 5AQ in ACN was found to be independent of monitored emission wavelength and exhibits well defined L_a and L_b bands and resembles to the absorption spectrum [10,11,13]. Fig. S1(a) shows the overlapped excitation spectrum of 5AQ in ACN on successive addition of Ag^+ ions. The excitation spectrum shifted from 345 nm to 360 nm with the addition of Ag^+ ions as shown in Fig. S1(b). Together with the spectral shift, the intensity of excitation spectrum gradually decreases with increasing concentration of Ag^+ ions. The change in the excitation spectrum is in agreement with that observed in absorption spectrum and demonstrates the ground state complex formation between Ag^+ ions and 5AQ in ACN.

The FL maximum shifted from 482 nm to 528 nm while changing the concentration of Ag^+ ions from 0.0 μ M to 60 mM. A linear relationship has been observed between the ratio of FL intensity at two

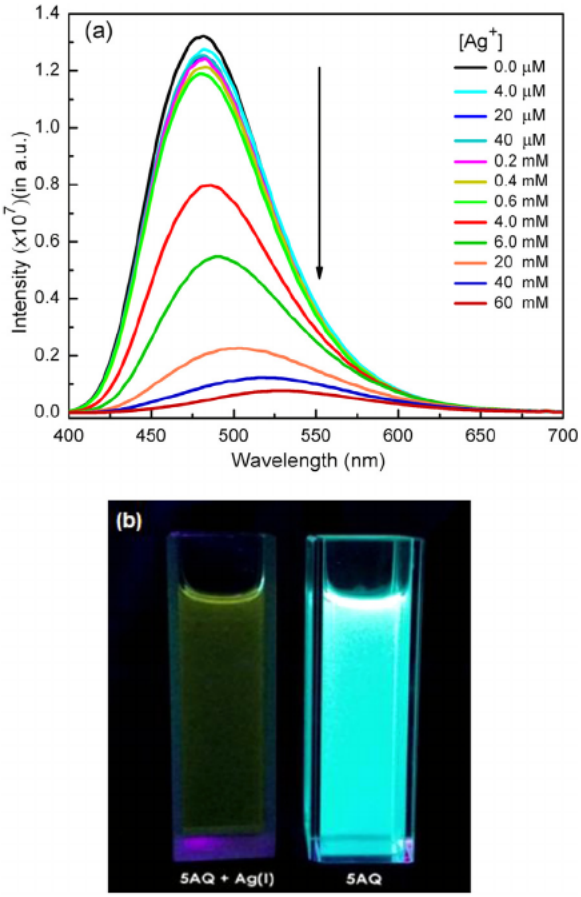


Fig. 3. (a) Fluorescence spectra of 5AQ $(1 \times 10^{-4}\ M)$ in ACN upon addition of various concentration of Ag^+ ions at excitation wavelength of 345 nm. (b) Fluorescence of 5AQ with UV light irradiation in the presence (left) and absence (right) of Ag^+ ions.

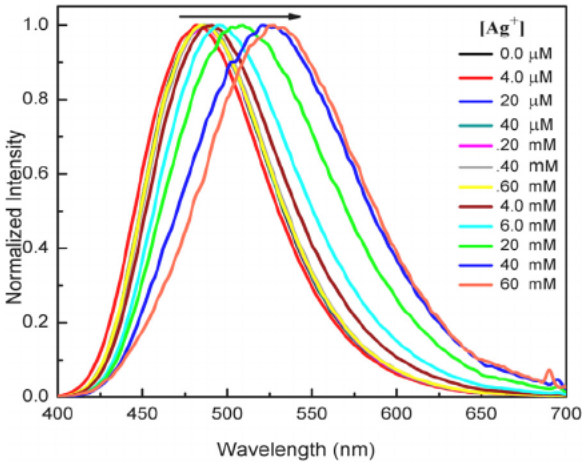


Fig. 4. Normalized fluorescence spectra of 5AQ in ACN in presence of various concentrations of Ag^+ ions.

wavelengths (I_{530}/I_{480}) and Ag^+ ion concentration (Fig. 5). In pure ACN, 5AQ shows a Stokes shift of ~ 8322 cm^{-1} and increases upto 10,058 cm^{-1} upon addition of Ag^+ ions (60 mM). The relation between Stoke's shift (Table 1) and Ag^+ ions concentration is shown in Fig. S2. This further demonstrates formation of complex in the ground state.

To understand the nature of FL quenching, the decrease of FL intensity is plotted as a function Ag^+ ions concentration following the so-called Stern-Volmer (S-V) relation Eq. (2) and modified S-V relation Eq.

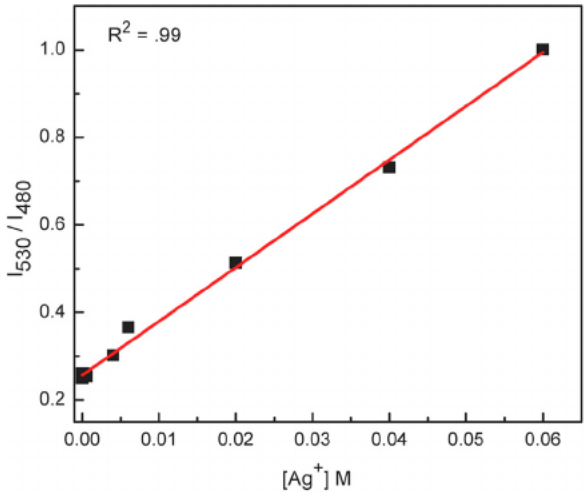


Fig. 5. Plot of the ratio of the intensity (I_{530}/I_{480}) of 5AQ in the presence of various concentration of Ag^{+}ions.

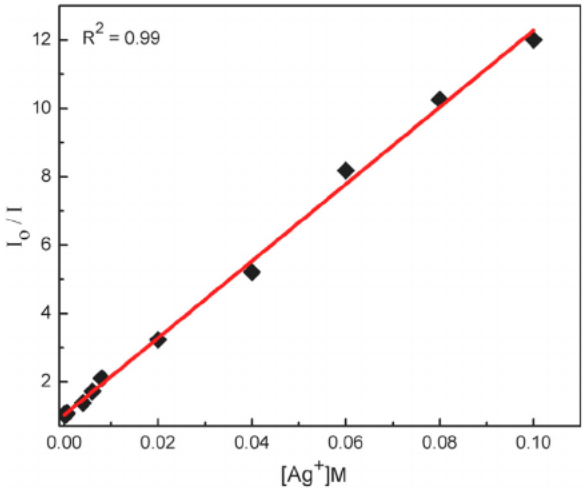


Fig. 6. The Stern-Volmer plot between the relative fluorescence intensity of 5AQ and Ag^{+} ions concentration.

Table 1
Steady-state parameters of 5AQ in acetonitrile in the presence of various concentration of Ag^{+} ions.

Concentration of Ag^{+} ions	$^{max}\lambda_{abs}$ (nm)	$^{max}\lambda_{em}$ (nm)	Stokes shift Δν (cm^{-1})
0.0 μM	344	482	8322
4.0 μM	344	482	8322
40 μM	344	483	8366
0.4 mM	344	484	8408
0.6 mM	344	484	8408
4.0 mM	344	485	8451
6.0 mM	344	490	8661
8.0 mM	344	495	8868
20 mM	344	502	9149
40 mM	344	525	10,019
60 mM	344	528	10,058

(3):

$$\frac{I_0}{I} = \frac{\tau_0}{\tau} = 1 + k_q\tau_0[Q] = 1 + K_{SV}[Q] \quad (2)$$

$$\frac{\left[\frac{I_0}{I} - 1\right]}{[Q]} = (K_{SV} + K_g) + K_{SV}K_g[Q] \quad (3)$$

where I_0 and I are the fluorescence intensities in the absence and presence of the quencher, respectively. [Q] is the quencher concentration and K_{SV} ($k_q\tau_0$) is the S-V constant, which represents the efficiency of fluorescence quenching. k_q is the bimolecular constant, τ_0 is the decay time in absence of quencher and K_g is ground state association constant. The value of K_{SV} can be determined from the slope of the graph displaying dependence of intensity or decay time on [Q], as expressed by Eq. (2). The obtained S-V plot for 5AQ in the presence of Ag^{+} ions is linear with Adj R^2 = 0.99, as shown in Fig. 6. The FL intensity may reduce either due to dynamic or static quenching, even with the linear relation of S-V plot. Thus, the steady-state spectral analysis only is not sufficient to confirm the nature of quenching. When decay time does not change with a decrease in fluorescent intensity, only static quenching is present i.e., K_{SV} = 0. A non-fluorescent complex is formed between the fluorophore and the quencher in case of quenching being static. The other case is that of dynamic quenching (i.e., K_g = 0) in which quenching response competes with emission in timescale and can be governed by the diffusion of quencher in any medium and its collision with excited probe molecule. These two quenching effects i.e., dynamic and static quenching, involve molecular interaction between the fluorophore and quencher. Simultaneous occurrence of static and dynamic quenching results in the positive deviation of plot I_0/I versus [Q] as Eq. (2) follows quadratic dependence with quencher concentration. The relation, $[I_0/I - 1]/[Q] = (K_{SV} + K_g) + K_{SV}K_g[Q]$ gives K_g after substituting the K_{SV} value retrieved from the data [1,17,18].

To ascertain the fluorescence quenching mechanism of 5AQ by Ag^{+} ions, time-resolved FL experiments were performed for various concentrations of Ag^{+} ions. FL decays were collected at λ_{ex} = 340 nm and λ_{em}= 530 nm. The FL decay of 5AQ fits to a bi-exponential function $\alpha_1 \exp(-t/\tau_1) + \alpha_2 \exp(-t/\tau_2)$ in an aprotic solvent (ACN). The term τ_1 and τ_2 are the lifetimes and α_1 and α_2 are the corresponding pre-exponential factors. The fitted lifetime are 1.26 ± 0.2 ns (0.93 %) and 11.60 ± 0.01 ns (99.07 %). The contribution of smaller decay constant is very small (< 1%). As an approximation, the decay fitted single exponential with a lifetime of 11.50± 0.01. Upon increasing the concentration of Ag^{+} ions, the lifetime decreases monotonically. It is interesting to note that, the lifetime component decreases from 11.50 ns to 3.14 ns on addition of 40 mM of Ag^{+} ions. The fitted overlapped decay curves obtained at varying concentrations of quencher are shown in Fig. 7 and the analyzed parameters are summarized in (Table 2). The decrease in FL lifetime plotted as a function of Ag^{+} ions concentration follow the S-V relation Eq. (2) again linearly, as shown in Fig. 8. Thus, the decrease in lifetime and the linear S-V plots demonstrate the dynamic quenching as a predominant mechanism. The value of K_{SV} constant for 5AQ obtained from the slope of the plot (Fig. 6) comes out to be 112.7 M^{-1} and k_q is around 9×10^{9} M^{-1} s^{-1}, respectively.

Furthermore, the absorption and fluorescence spectra of 5AQ in ACN in the presence and absence of different metal ions (Cu^{2+}, Ca^{2+}, Pb^{2+}, Fe^{2+}, Mg^{2+}, Li^{+}, K^{+}) were examined. The fluorescence response of 5AQ in the presence of various metal ions under UV is shown in Fig. S3 of the Supporting information. The FL of 5AQ upon addition of 40 mM concentration of various metal ions is illustrated in Fig. 9. As discussed, the fluorescence intensity of 5AQ is quenched significantly on addition of Ag^{+} ions whereas for other metal ions it shows minimal change. To investigate the interference from other metal ions we further conducted the competition experiment in the presence of 40 mM of various metal ions (Cu^{2+}, Ca^{2+}, Pb^{2+}, Fe^{2+}, Mg^{2+}, Li^{+}, K^{+}) with subsequent addition of 40 mM of Ag^{+} ions, as shown in Fig. 10. These results reveal that the recognition of Ag^{+} ion by probe 5AQ is not significantly influenced by other competitive metal ions. Thus, it is observed that the probe molecule 5AQ has good selectivity for Ag^{+} ions (as compared to the other metal ions under the same condition and concentration) and hence can be used to detect Ag^{+} ions in polar aprotic medium, which

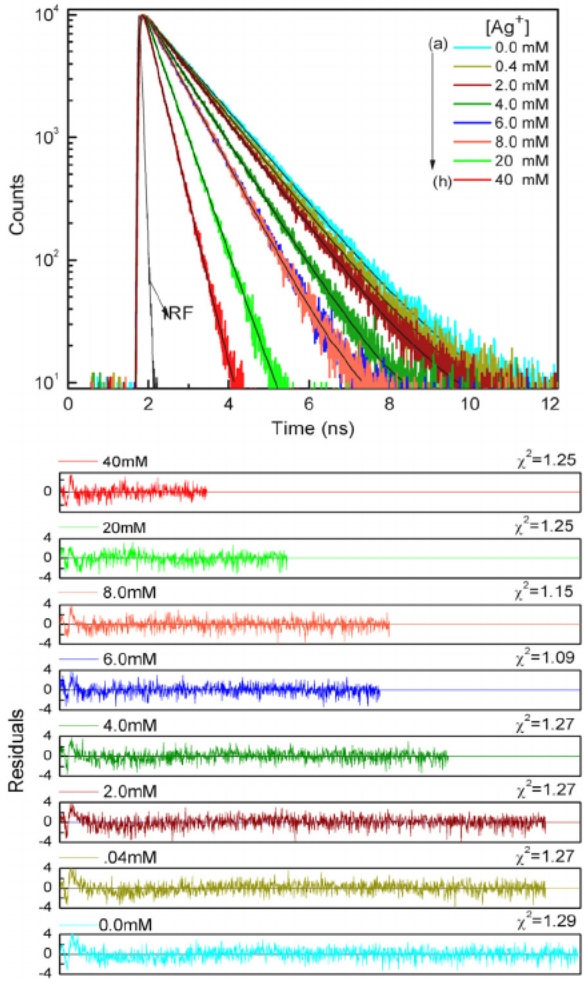


Fig. 7. Overlapped fitted fluorescence decay curves of 5AQ in presence of various concentration of Ag^+ ions and corresponding residuals at λ_{ex} = 340 nm and λ_{em} = 530 nm.

Table 2
Fluorescence lifetime parameters of 5AQ with various concentration of Ag^+ ions at λ_{ex} = 340 nm and λ_{em} = 530 nm.

Concentration of Ag^+ ions	τ (ns)	χ^2	τ_0/τ
0.0 mM	11.50 ± 0.01	1.29	1.0
0.4 mM	10.84 ± 0.01	1.27	1.060
2.0 mM	10.02 ± 0.009	1.27	1.148
4.0 mM	8.64 ± 0.008	1.27	1.331
6.0 mM	7.40 ± 0.09	1.09	1.554
8.0 mM	7.31 ± 0.007	1.15	1.573
20.0 mM	4.65 ± 0.005	1.25	2.473
40.0 mM	3.14 ± 0.004	1.25	3.662

might be due to high-affinity of Ag^+ ions to 5AQ as compared to other metal ions.

As discussed, in the absence of quencher, the FL decay of 5AQ is fitted with a bi-exponential function. The contribution of the shorter component is very small (< 1%) and contribution of the longer component is dominant in the overall decay function. The smaller lifetime component is assigned to the locally excited (LE) state while the longer lifetime component is assigned to the intramolecular charge transfer (ICT) state [10,11,13]. Thus, the ICT is probably responsible for the overall FL. However, there is no rise time component in the FL decay, indicating that ICT may be faster than the resolution of our experimental setup.

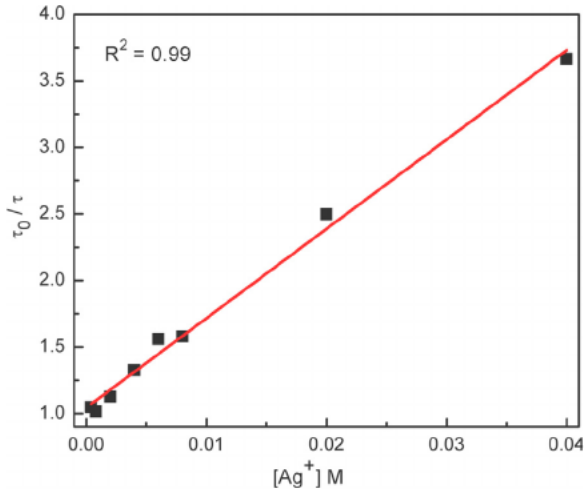


Fig. 8. Stern-Volmer plot between the lifetime of 5AQ and Ag^+ ions concentration.

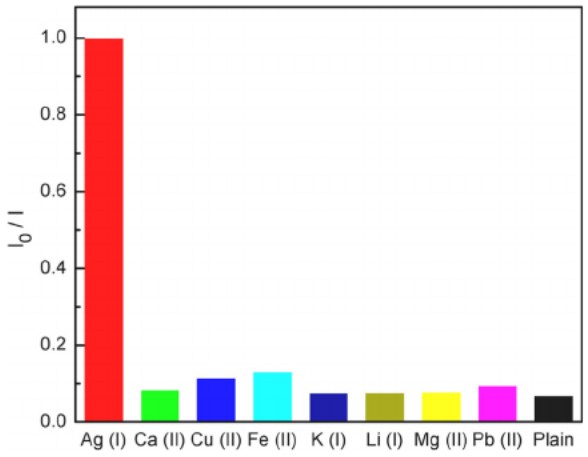


Fig. 9. Fluorescence response of 5AQ in ACN in the presence of various metal ions. The excitation wavelength was 345 nm.

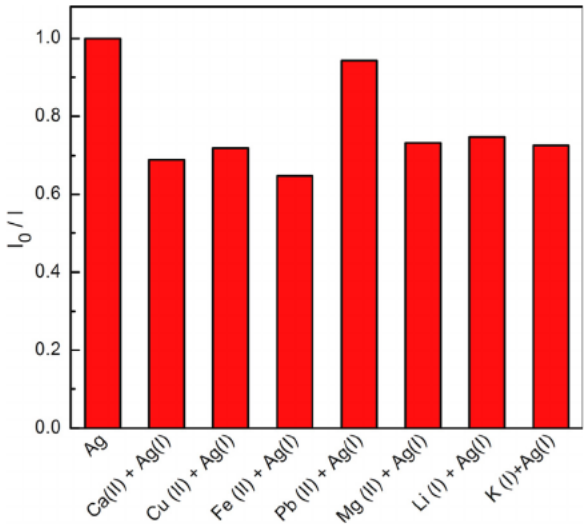


Fig. 10. Fluorescence response of 5AQ (10^{-4} M) in presence of 40 mM of various metal ions with subsequent addition 40 mM of Ag^+ ions to the solution at λ_{ex} = 345 nm.

In the presence of Ag^+ ions, the decay trend is very similar to that in the absence of quencher. However, the observed decay fits well with mono-exponential function with an approximation. The lifetime decreases with increasing the concentration of Ag^+ ions, which indicates that the FL corresponding to the ICT is quenched. This is due to the complexation between 5AQ and Ag^+ ions, as evident from the absorption and excitation spectra. While increasing the concentration of Ag^+ ions, the absorption/excitation intensity decreases and the spectra shift towards red. This may be due to the formation of a complex between 5AQ and Ag^+ ions, which has minimal oscillator strength/absorption intensity. The shift in the FL spectrum may probably be ascribed to the different ICT states formed by the interaction of 5AQ with Ag^+ ions. The decrease of FL lifetime also indicates the decrease of the FL of ICT state formed from the LE state.

Furthermore, it has been classified that metal ions, such as Ag^+ ions, being soft cations, prefer soft donors for complexation [32]. The nitrogen atom of 5AQ is involved in the coordination with Ag^+ ions. The interaction between Ag^+ ions and the nitrogen atom in 5AQ moiety enhances the electron-withdrawing ability of 5AQ that leads to a different state and results in a large red-shifted FL. The donation of the $-C = N$ nitrogen lone pair electrons to the Ag^+ may decrease the fluorescent intensity [21]. Generally, the FL of luminescent ligands is quenched by complexation of the ligand with transition metal ion due to paramagnetic quenching. In this case, the unpaired electrons of the metal ions actively interact with the π electrons of the ligand, providing a path way for inter system crossing from singlet excited state of the ligand to the higher multiplicity states (i.e., triplet state) due to interaction with metal ions [31].

It is important to mention that the present study has been performed in some other solvents also. Since 5AQ shows weak fluorescence in polar protic solvents (water, methanol, ethanol), therefore enhancement in FL intensity is noticed by adding different concentrations of Ag^+ ions. However, a small amount of quenching is observed in DMSO (dimethyl sulfoxide) as compared to that in ACN (acetonitrile) in the presence of Ag^+ ions. Thus, this makes the system attractive for FL quenching/sensing in ACN by Ag^+ ions.

4. Conclusions

In summary, we have studied fluorescence-quenching behavior of 5AQ for Ag^+ ions in ACN, which results in ~ 17-fold decrement in FL intensity upon addition of Ag^+ ions. The lifetime is reduced from 11.5 to 3.14 ns. The large Stokes shifted FL, which produces from ICT state is advantageous to avoid self-absorption and interference from the light source in fluorescence sensing technology. We demonstrate dynamic quenching for all concentration of silver ions used in the study. Stern-Volmer quenching constant (K_{SV}) as $112.7\ M^{-1}$, biomolecular quenching constant (k_q) as $9.8 \times 10^9\ M^{-1}s^{-1}$ and binding constant (K) as $1.407 \times 10^4\ M^{-1}$ have been determined for Ag^+ ions. In addition, 5AQ shows high affinity for Ag^+ ions over other competitive metal ions. Therefore, the proposed system emerges to be quite interesting in sensing/recognition of Ag^+ ions in protic environment in a concentration range of $40\ \mu M$ to $60\ mM$.

Acknowledgments

Authors are thankful to Prof. A.B. Melkani, Department of Chemistry, Kumaun University for providing a facility for absorption measurement. Author NP is thankful to **University Grants Commission** (UGC) (Sanction No. F, 5–53/2007(BSR) New Delhi India for providing research fellowships under the scheme for the meritorious students in sciences (RFSMS) and Dr. J.P. Bridhkoti for his valuable discussions. Author MSM thanks Science and Engineering Research Board (SERB Department of Science & Technology) (EMR/2016/001110) (DST), Govt. of India for financial support.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.jlumin.2018.09.062

References

[1] J.R. Lakowicz, Principles of Fluorescence Spectroscopy, third ed., Springer, 2006, https://doi.org/10.1007/978-0-387-46312-4.
[2] H. Mishra, S. Pant, H.B. Tripathi, Temperature-dependent time-resolved fluorescence study of cinchonine alkaloid dication, J. Fluoresc. 18 (2008) 17–27, https://doi.org/10.1007/s10895-007-0229-6.
[3] M.S. Mehata, H.C. Joshi, H.B. Tripathi, Fluorescence characteristics of protonated form of 6-hydroxyquinoline in Nafion® film, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 59 (2003) 559–567, https://doi.org/10.1016/S1386-1425(02)00206-8.
[4] Y.H. Kim, J.S. Youk, S.H. Kim, S.K. Chang, Solvatochromic fluorescence behavior of 8-aminoquinoline-benzothiazole: a sensitive probe for water composition in binary aqueous solutions, Bull. Korean Chem. Soc. 26 (2005) 47–50, https://doi.org/10.5012/bkcs.2005.26.1.047.
[5] D. Panda, D. Ghosh, A. Datta, Acid-base behavior of 3-aminoquinoline in its ground and excited states, J. Photochem. Photobiol. A Chem. 207 (2009) 254–259, https://doi.org/10.1016/j.jphotochem.2009.07.017.
[6] R. Rautela, N.K. Joshi, H.C. Joshi, N. Tewari, S. Pant, Solvatochromic study of 2-hydroxy-4-methylquinoline for the determination of dipole moments and solute-solvent interactions, J. Mol. Liq. 154 (2010) 47–51, https://doi.org/10.1016/j.molliq.2010.04.004.
[7] N.K. Joshi, H.C. Joshi, R. Gahlaut, N. Tewari, R. Rautela, S. Pant, Steady State and time-resolved fluorescence study of isoquinoline: reinvestigation of excited state proton transfer, J. Phys. Chem. A 116 (2012) 7272–7278, https://doi.org/10.1021/jp210516t.
[8] A.K. Singh, S. Ghosh, R. Kancherla, A. Datta, Engineering the excited-state dynamics of 3-aminoquinoline by chemical modification and temperature variation, J. Phys. Chem. B 120 (2016) 12920–12927, https://doi.org/10.1021/acs.jpcb.6b09939.
[9] S.G. Schulman, L.B. Sanders, Fluorescence and phosphorescence of 5- and 8-aminoquinoline, Anal. Chim. Acta 56 (1971) 83–89, https://doi.org/10.1016/S0003-2670(01)80111-2.
[10] J.P. Bridhkoti, H.C. Joshi, S. Pant, Fluorescence characteristics of 5-aminoquinoline in acetonitrile: water, J. Mol. Liq. 164 (2011) 197–200, https://doi.org/10.1016/j.molliq.2011.09.014.
[11] J.P. Bridhkoti, H. Mishra, H.C. Joshi, S. Pant, Photochemistry of 5-aminoquinoline in protic and aprotic solvents, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 79 (2011) 412–417, https://doi.org/10.1016/j.saa.2011.02.044.
[12] D. Panda, A. Datta, The role of the ring nitrogen and the amino group in the solvent dependence of the excited-state, dynamics of 3-aminoquinoline, J. Chem. Phys. 125 (2006) 1–10, https://doi.org/10.1063/1.2232199.
[13] J.P. Bridhkoti, R. Gahlaut, H.C. Joshi, S. Pant, Effect of positional substitution of amino group on excited state dipole moments of quinoline, J. Lumin. 131 (2011) 1869–1873, https://doi.org/10.1016/j.jlumin.2011.04.038.
[14] S. Kumar, S. Bawa, H. Gupta, Biological activities of quinoline derivatives, Mini-Rev. Med. Chem. 9 (2009) 1648–1654, https://doi.org/10.2174/138955709791012247.
[15] K. Tiwari, P. Arora, N. Pandey, P. Pandey, H.C. Joshi, S. Pant, Experimental and computational approaches on dipole moment of 5-aminoisoquinoline, J. Mol. Liq. 200 (2014) 460–464, https://doi.org/10.1016/j.molliq.2014.10.027.
[16] M.S. Mehata, H.B. Tripathi, Fluorescence quenching of 6-methoxyquinoline: an indicator for sensing chloride ion in aqueous media, J. Lumin. 99 (2002) 47–52, https://doi.org/10.1016/S0022-2313(02)00294-6.
[17] R. Rautela, P. Arora, N.K. Joshi, S. Pant, H.C. Joshi, Fluorescence quenching of 8-methyl quinolinium: an efficient halide indicator mechanism, J. Mol. Liq. 218 (2016) 632–636, https://doi.org/10.1016/j.molliq.2016.03.002.
[18] N.K. Joshi, N. Tewari, P. Arora, R. Rautela, S. Pant, H.C. Joshi, Photophysical behavior and fluorescence quenching by halides of quinidine dication: steady state and time resolved study, J. Lumin. 158 (2015) 412–416, https://doi.org/10.1016/j.jlumin.2014.10.041.
[19] Z. Liu, W. He, Z. Guo, Metal coordination in photoluminescent sensing, Chem. Soc. Rev. 42 (2013) 1568, https://doi.org/10.1039/c2cs35363f.
[20] B. Zhang, H. Liu, F. Wu, G.F. Hao, Y. Chen, C. Tan, Y. Tan, Y. Jiang, A dual-response quinoline-based fluorescent sensor for the detection of Copper (II) and Iron(III) ions in aqueous medium, Sens. Actuators B Chem. 243 (2017) 765–774, https://doi.org/10.1016/j.snb.2016.12.067.
[21] J. Nie, N. Li, Z. Ni, Y. Zhao, L. Zhang, A sensitive tetraphenylethene-based fluorescent probe for Zn^{2+} ion involving ESIPT and CHEF processes, Tetrahedron Lett. 58 (2017) 1980–1984, https://doi.org/10.1016/j.tetlet.2017.04.027.
[22] M. Formica, V. Fusi, L. Giorgi, M. Micheloni, New fluorescent chemosensors for metal ions insolution, Coord. Chem. Rev. 256 (2012) 170–192, https://doi.org/10.1016/j.ccr.2011.09.010.
[23] Xiangming Meng, ShuXin Wang, Manzhou Zhu. Quinoline-Based Fluorescence Sensors, Molecular Photochemistry Satyen Saha, IntechOpen, doi:10.5772/31771, 2012. Available from: ⟨https://www.intechopen.com/books/molecular-photochemistry-various-aspects/quinoline-based-fluorescence-sensors⟩.
[24] Y. Hu, F. Zhao, S. Hu, Y. Dong, D. Li, Z. Su, A novel turn-on colorimetric and fluorescent sensor for Fe^{3+} and its application in living cells, J. Photochem.

Photobiol. A Chem. 332 (2017) 351–356, https://doi.org/10.1016/j.jphotochem.2016.09.006.
[25] J. Wu, W. Liu, J. Ge, H. Zhang, P. Wang, New sensing mechanisms for design of fluorescent chemosensors emerging in recent years, Chem. Soc. Rev. 40 (2011) 3483, https://doi.org/10.1039/c0cs00224k.
[26] P. Arora, K. Suyal, N.K. Joshi, H.C. Joshi, S. Pant, Fluorescence characteristics of 5-amino salicylic acid: an iodide recognition study, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 94 (2012) 119–125, https://doi.org/10.1016/j.saa.2012.03.056.
[27] Y. Li, H. Yu, G. Shao, F. Gan, A tetraphenylethylene-based “turn on” fluorescent sensor for the rapid detection of Ag^+ ions with high selectivity, J. Photochem. Photobiol. A Chem. 301 (2015) 14–19, https://doi.org/10.1016/j.jphotochem.2014.12.013.
[28] C. Liu, S. Huang, H. Yao, S. He, Y. Lu, L. Zhao, X. Zeng, Preparation of fluorescein-based chemosensors and their sensing behaviors toward silver ions, RSC Adv. 4 (2014) 16109, https://doi.org/10.1039/C3RA47392A.
[29] N.J. Wang, C.M. Sun, W.S. Chung, A highly selective fluorescent chemosensor for Ag^+ based on calix − 4-arene with lower-rim proximal triazolylpyrenes, Sens. Actuators B Chem. 171 (2012) 984–993, https://doi.org/10.1016/j.snb.2012.06.014.
[30] A. Affrose, S. Dawood, S. Parveen, B.S. Kumar, Selective sensing of silver ion using berberine, a naturally occurring plant alkaloid, Sens. Actuators B. Chem. 206 (2015) 170–175, https://doi.org/10.1016/j.snb.2014.09.042.
[31] S.N. Sathyanarayana, Electronic Absorption Spectroscopy and Related Techniques, Universities Press, 2001 (ISBN 8173713715,9788173713712).
[32] A.D. Arulraj, R. Devasenathipathy, S.M. Chen, V.S. Vasantha, S.F. Wang, Highly selective and sensitive fluorescent chemosensor for femtomolar detection of silver ion in aqueous medium, Sens. Bio-Sens. Res. 6 (2015) 19–24, https://doi.org/10.1016/j.sbsr.2015.10.004.
[33] L. Chen, Q. Zhao, X.Y. Zhang, G.H. Tao, Determination of silver ion based on the redshift of emission wavelength of quantum dots functionalized with rhodanine, Chin. Chem. Lett. 25 (2014) 261–264, https://doi.org/10.1016/j.cclet.2013.10.029.
[34] J. Prabhu, K. Velmurugan, Q. Zhang, S. Radhakrishnan, L. Tang, R. Nandhakumar, Symmetric fluorescent probes for the selective recognition of Ag-ion via restricted C=N isomerization and on-site visual sensing applications, J. Photochem. Photobiol. A Chem. 337 (2017) 6–18, https://doi.org/10.1016/j.jphotochem.2017.01.006.
[35] S.K. Sheet, B. Sen, R. Thounaojam, K. Aguan, S. Khatua, Ruthenium(II) complex-based luminescent bifunctional probe for Ag^+ and phosphate ions: Ag^+-assisted detection and imaging of rRNA, Inorg. Chem. 56 (2017) 1249–1263, https://doi.org/10.1021/acs.inorgchem.6b02343.
[36] L.Q. Li, L.J. Gao, A novel rosamine based fluorescent sensor for Ag^+ recognition, Spectrochim. Acta - Part A Mol. Biomol. Spectrosc. 152 (2016) 426–430, https://doi.org/10.1016/j.saa.2015.07.097.
[37] A.S. Al-Kady, M. Gaber, M.M. Hussein, E.Z.M. Ebeid, Fluorescence enhancement of coumarin thiourea derivatives by Hg^{2+}, Ag^+ and silver nanoparticles, J. Phys. Chem. A 113 (2009) 9474–9484, https://doi.org/10.1021/jp905566z.
[38] C.J. Chen, A.J. Zhang, J.B. Hu, Y.H. Miao, X. Liu, A simple quinoline-based fluorescent “turn on” probe with high selectivity for Ag^+ in aqueous media, Inorg. Chim. Acta 435 (2015) 137–141, https://doi.org/10.1016/j.ica.2015.06.014.
[39] G. Ambrosi, E. Borgogelli, M. Formica, V. Fusi, L. Giorgi, M. Micheloni, E. Rampazzo, M. Sgarzi, N. Zaccheroni, L. Prodi, PluS Nanoparticles as a tool to control the metal complex stoichiometry of a new thio-aza macrocyclic chemosensor for Ag(I) and Hg(II) in water, Sens. Actuators B Chem. 207 (2015) 1035–1044, https://doi.org/10.1016/j.snb.2014.07.107.
[40] Y. bao Jiang, C. Gao, X. Zhang, J. shui Yao, Q. ze Liu, X. xia Cai, A highly selective and sensitive fluorescence probe with A-π-D-π-A structure for detection of Ag^+, J. Mol. Struct. 1163 (2018) 33–40, https://doi.org/10.1016/j.molstruc.2018.01.058.
[41] T. Anand, G. Sivaraman, A. Mahesh, D. Chellappa, Analytica Chimica Acta Aminoquinoline based highly sensitive fluorescent sensor for lead (II) and aluminum (III) and its application in live cell imaging, Anal. Chim. Acta 853 (2015) 596–601, https://doi.org/10.1016/j.aca.2014.11.011.
[42] Y.W. Choi, J.J. Lee, C. Kim, A highly selective fluorescent chemosensor based on a quinoline derivative for zinc ions in pure water, RSC Adv. 5 (2015) 60796–60803, https://doi.org/10.1039/C5RA09954D.

Electrical And Thermal Characterization of Lithium-Ion Batteries with the help of Hardware and Software Simulation

Supervised by:

Prof. M. M. Tripathi
Electrical Engg. Department
Delhi Technological University,
New Delhi, India

&

Dr. Amrish K. Panwar
Department of Applied Physics
Delhi Technological University
New Delhi, India

Prepared by:

Snigdha Sharma
Electrical Engineering
Delhi Technological University
New Delhi, India

The Electrochemical Society
Advancing solid state & electrochemical science & technology

236th ECS Meeting
ATLANTA, GA
October 13-17, 2019
Hilton Atlanta

Outline of the Presentation

1. Introduction
2. Simulation using software
3. Interfacing of hardware with software
4. Result and discussion
5. Conclusion
6. References

Introduction

- To fulfil the growing demand of clean energy rechargeable batteries are used as energy storage and conversion devices.
- The most commercialized rechargeable batteries are Li-ion batteries.
- In this work, evaluation of Li-ion batteries has been done to determine the battery characteristics[5] which helps in realizing more reliable and safe batteries.
- Two types of battery characterization techniques i.e. Thermal and electrical characterizations has been done.
- Thermal characterization has been obtained using Multiphysics Software.
- Electrical characterization is done using data acquisition software 'LabVIEW' and by interfacing hardware with software system[9].

Simulation using software

Snigdha Sharma

Model geometry and formulation

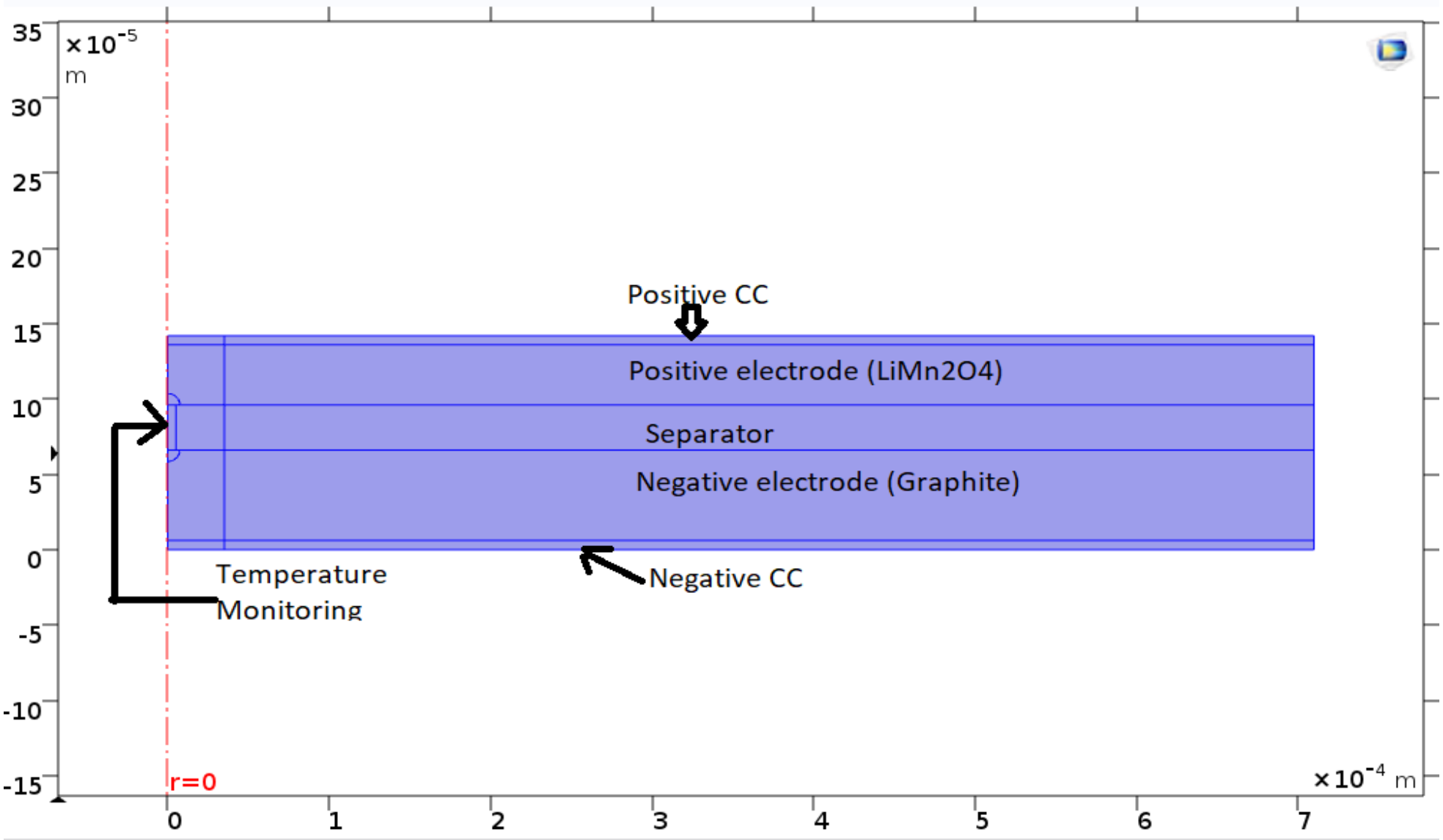


Fig. 1 Geometry of Lithium-ion Battery model

Model parameters

- Model contains 2D axisymmetric domain surrounded by infinite elements, that comprises the pseudo two- dimensional electrochemical structure.
- It allows heat transportation only in one direction.
- Characterization techniques provides thermal performances which contain various parameters for battery model as shown below.

Thermal properties for Battery model

Parameters	$LiMn_2O_4$	Graphite	$LiPF_6$	Aluminum CC	Copper CC
Thermal Conductivity(W(m.k))	3.4	1	0.15	400	238
Young Modulus (Pa)	$4.9*e^9$	$4.1*e^9$	-	$110*e^9$	$70*e^9$
Density (Kg/m^3)	3600	2260	940	8960	2700
Heat Capacity(J/Kg.k)	2500	881	1046	385	900

Measured and Identified Parameters for the Lithium-Ion Battery Model Simulation

Parameter	Units	Value	Description
E_cell	V	3.4	Cell Potential
L_tot	m	$1.42*e^{-4}$	Cell total Thickness
epsl_sep	-	0.4	Porosity of separator
$T_{initial}$	K	328	Initial Temperature
T_{final}	K	395	Final Temperature
Maximum electrodeSOC	SOCmax	0.79	Operational SOC
Minimum electrode SOC	SOCmin	0.02	Operational SOC
Q_B	A*h	12	Battery Capacity

Measured and Identified Parameters for the Lithium-Ion Battery Model Simulation

Parameter	Units	Value	Description
N_layers	-	25	No. of cell layers
kT_batt	L_batt/(L_pos/kT_pos+L_neg/kT_neg+L_pos_cc/kT_pos_cc+L_neg_cc/kT_neg_cc+L_sep/kT_sep)	0.87451 W/(m·K)	Battery thermal conductivity
Rho_batt	(rho_pos*L_pos+rho_neg*L_neg+rho_pos_cc*L_pos_cc+rho_neg_cc*L_neg_cc+rho_sep*L_sep)/L_batt	2189.3 kg/m³	Battery density
Cp_batt	(Cp_pos*L_pos+Cp_neg*L_neg+Cp_pos_cc*L_pos_cc+Cp_neg_cc*L_neg_cc+Cp_sep*L_sep)/L_batt	1391.8 J/(kg·K)	Battery heat capacity

Interfacing hardware with software

Simulation through Hardware

- **OVERVIEW OF EXPERIMENT**

1. Simple electronic circuit is connected with the Li-ion battery bank.
2. Discharge curves are optimized with the help of the software, LabVIEW.

- **MATERIALS AND EQUIPMENT USED**

1. A battery bank of six coin-cell of maximum 3.5V.
2. The variable resistance of 0-500 ohm.
3. Motor as a load, capacitors, three switches, wires,
4. A protoboard, a digital multimeter,
5. A personal computer connected with National Instrument 6001 DAQ card and commercial software LabVIEW with updated version were used.[11]

Experiment Setup

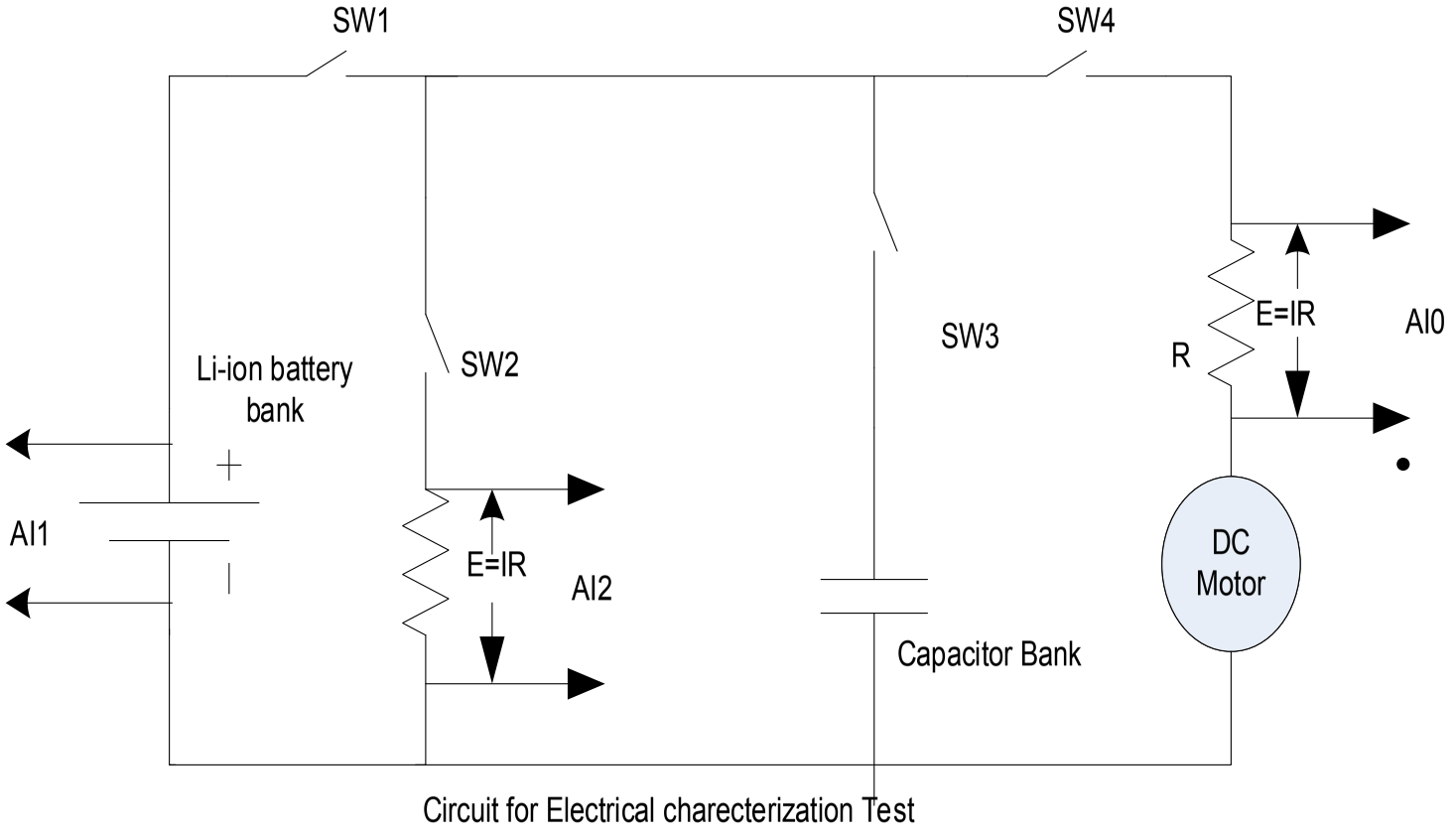
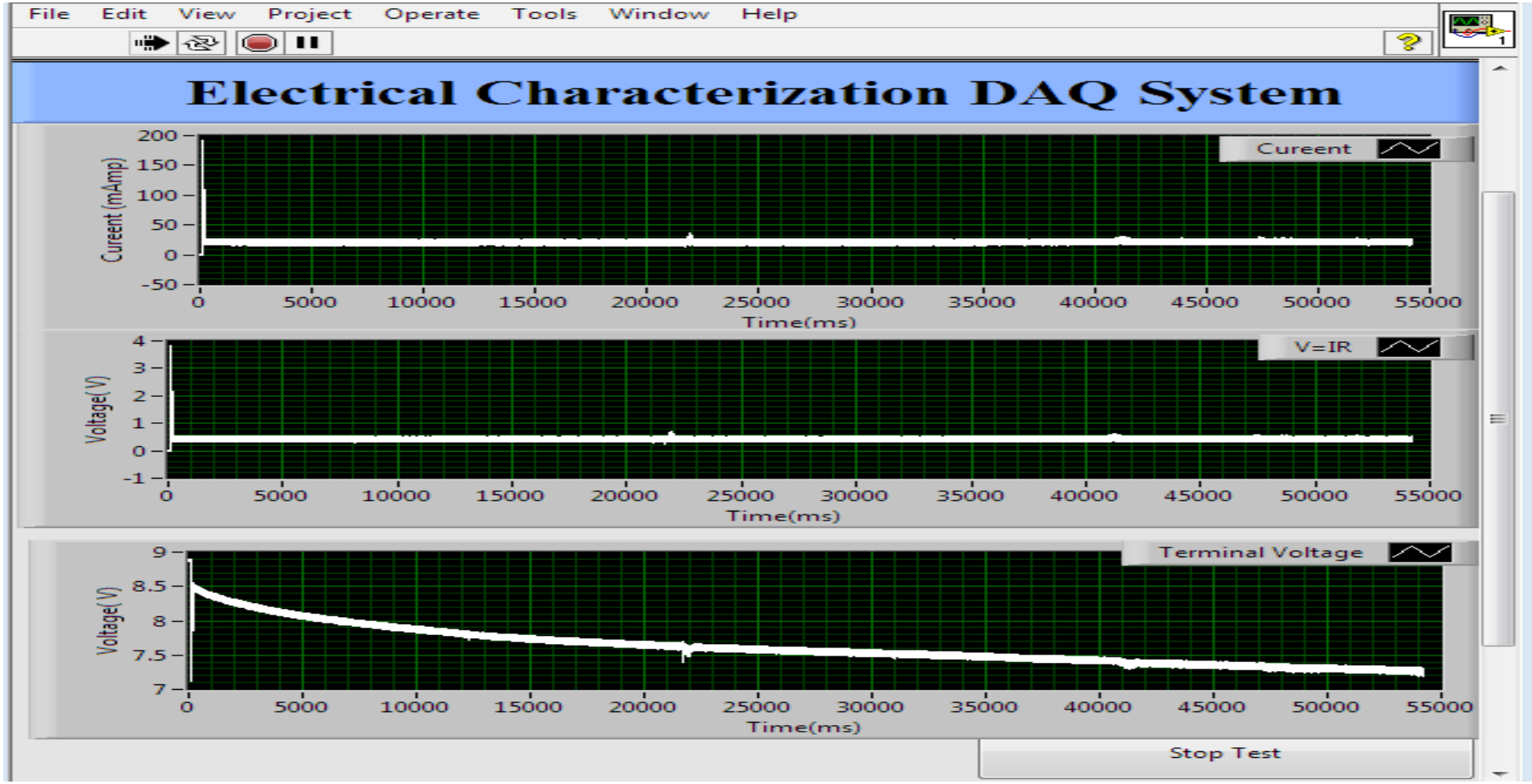


Fig 2. Circuit for Hardware design.

- Switch SW1, SW2 is closed for analyzing a short circuit current across the resistor through an analog input channel 2.

- Switch SW1, SW3, SW4 are closed for monitoring the battery characteristics, when the capacitor bank is connected parallel across the circuit.
- Three electrical tests were performed for optimization of different types of discharge curves.
- Switch SW1, SW4 is closed for monitoring the potential drop across the battery terminal

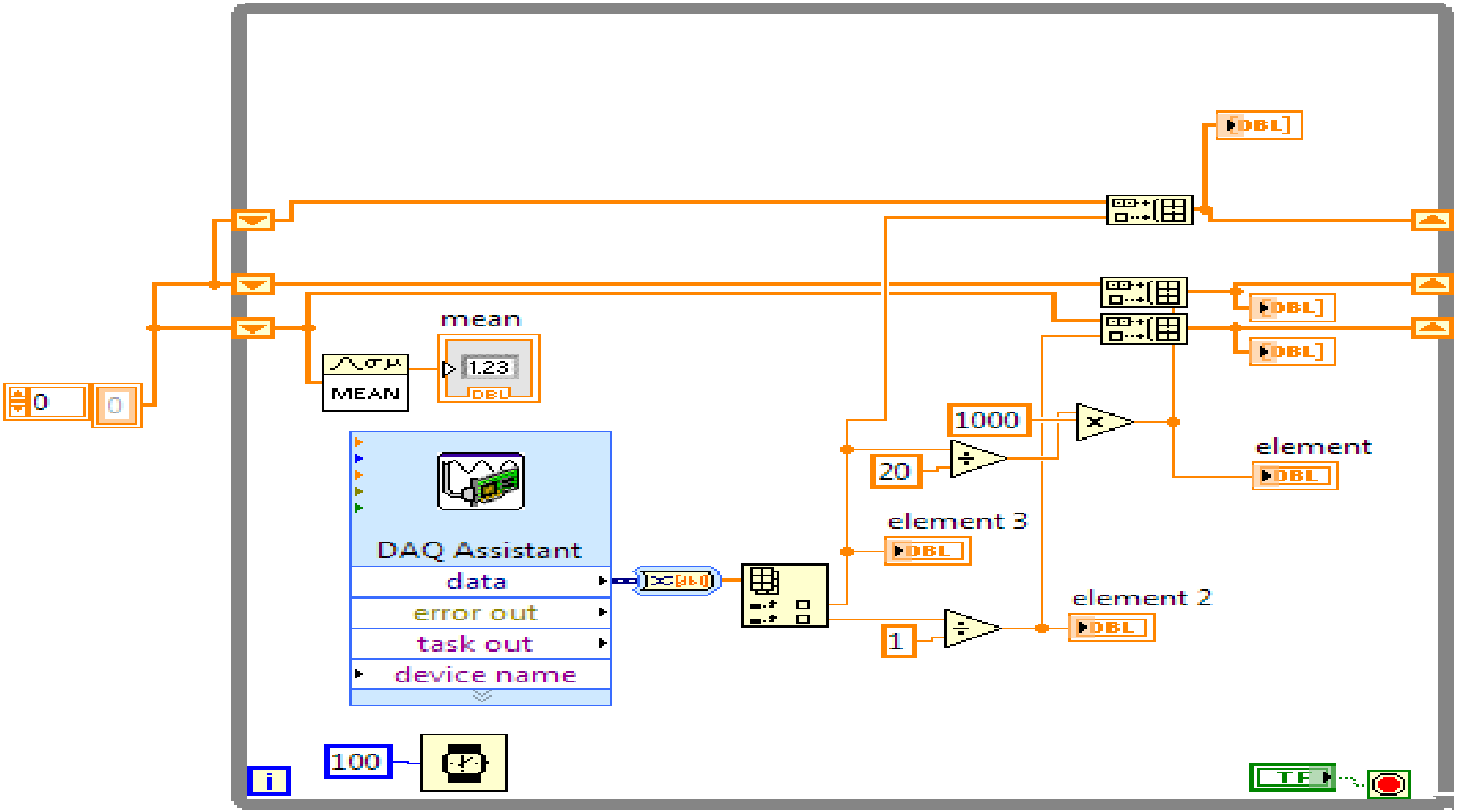
Battery discharge curve on the LabVIEW Front Panel



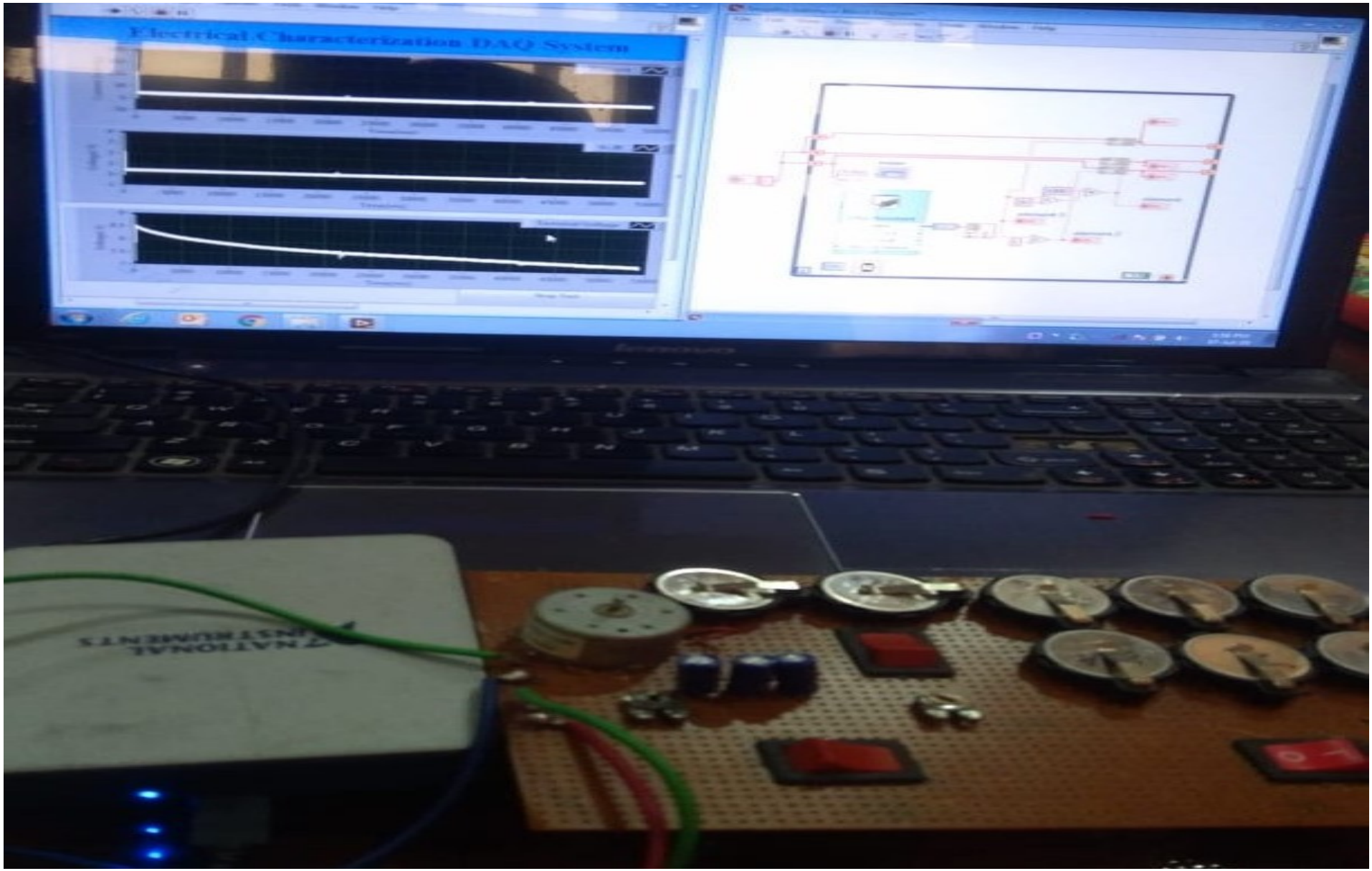
Front Panel works as an interface between the hardware and the PC[16].
It is also used for modifying the experimental condition and introduce virtual information across the panel and it is called a virtual instrument (VI).

Block Diagram of VI front panel

- The second panel is used for designing a block diagram and also provides a proper platform for writing a program by using various icons and connectors.



Hardware Setup for Electrical Characterization



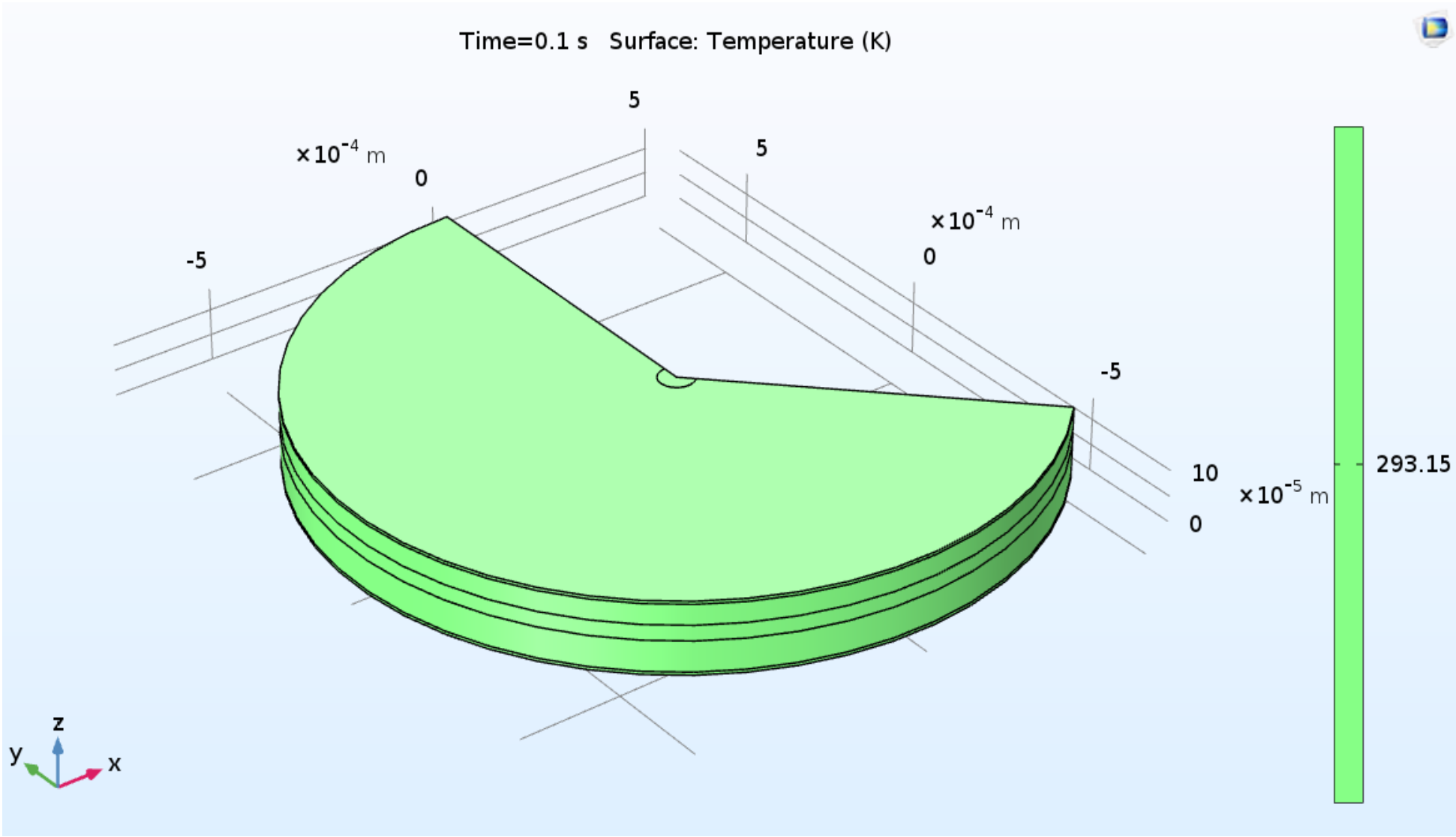
Results And Discussion

During battery characterization, the following two tests have been performed to optimize the behavior of the Li-ion battery.

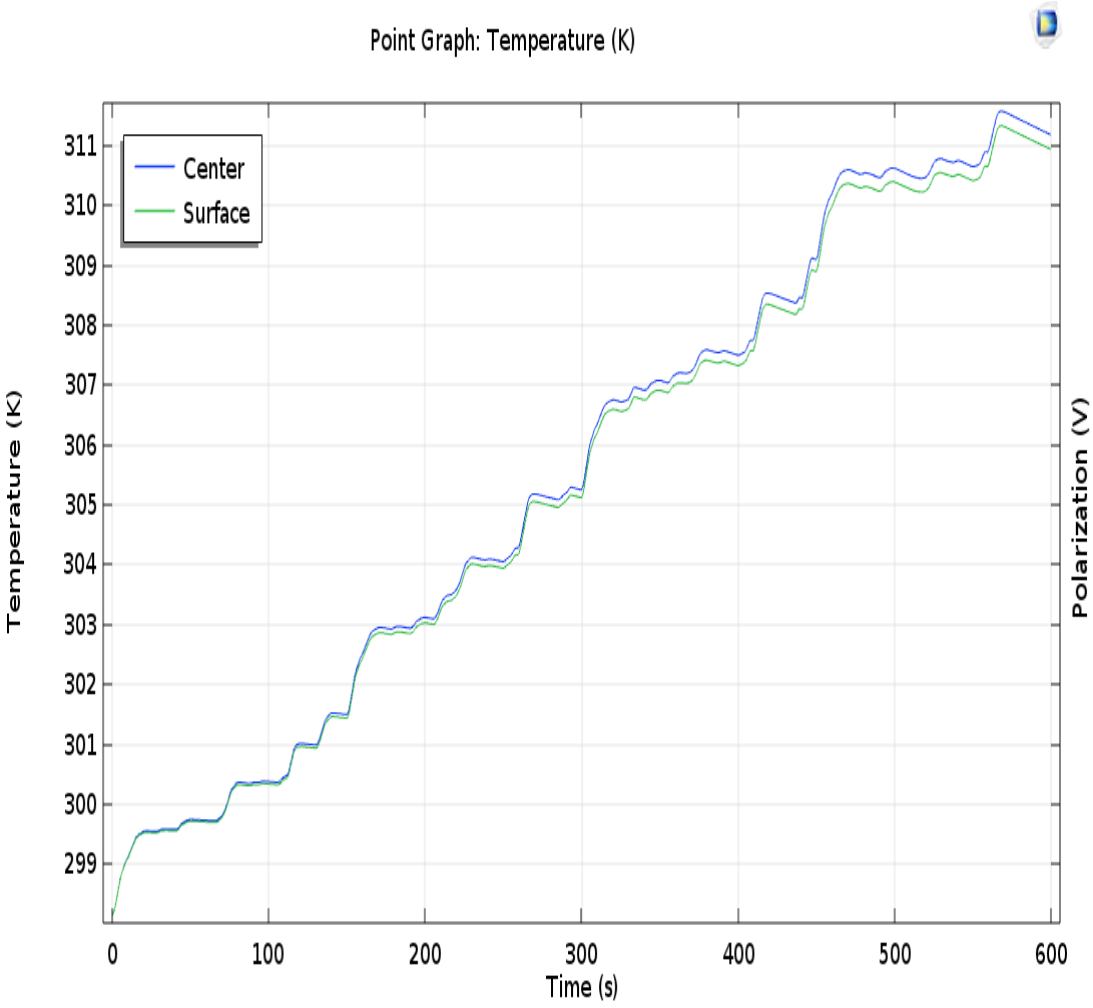
- Thermal Characterization
- Electrical Characterization

- All simulation results were analyzed at different boundary conditions.

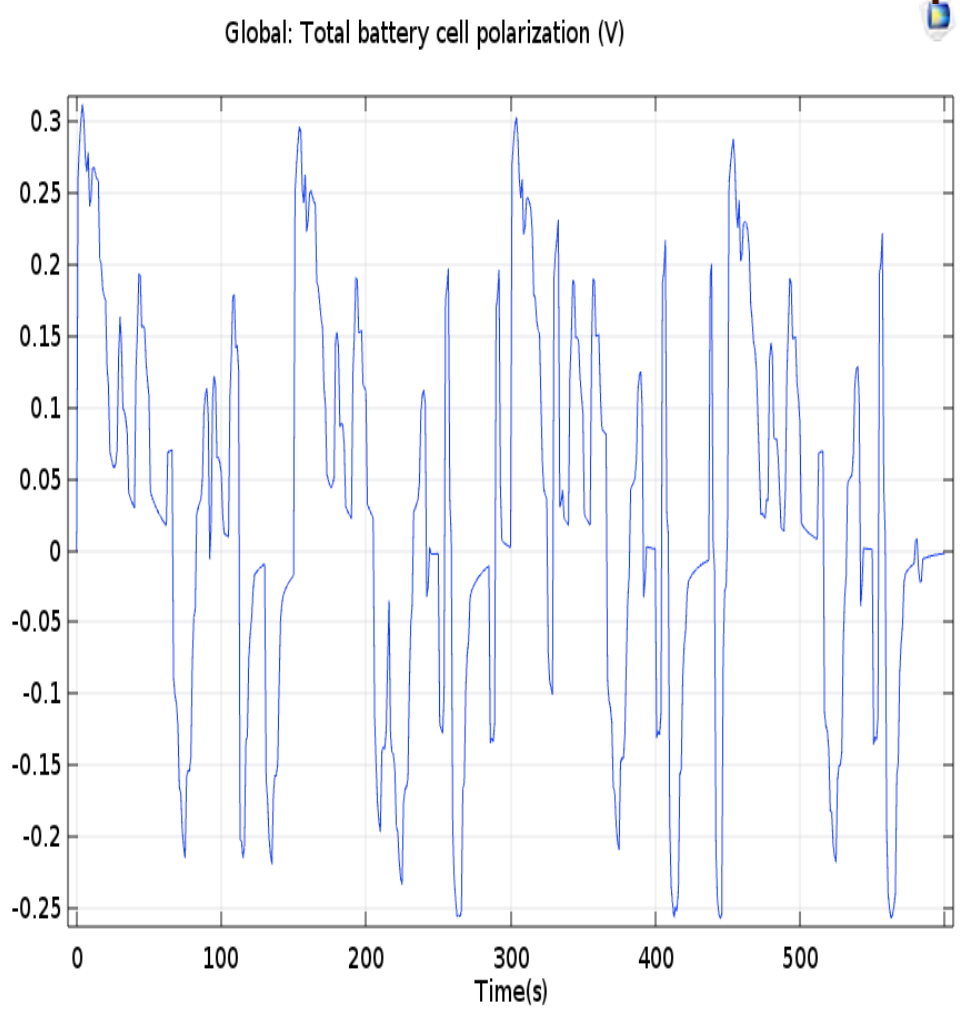
Thermal Characterization - Temperature distribution over the surface



Thermal Characterization

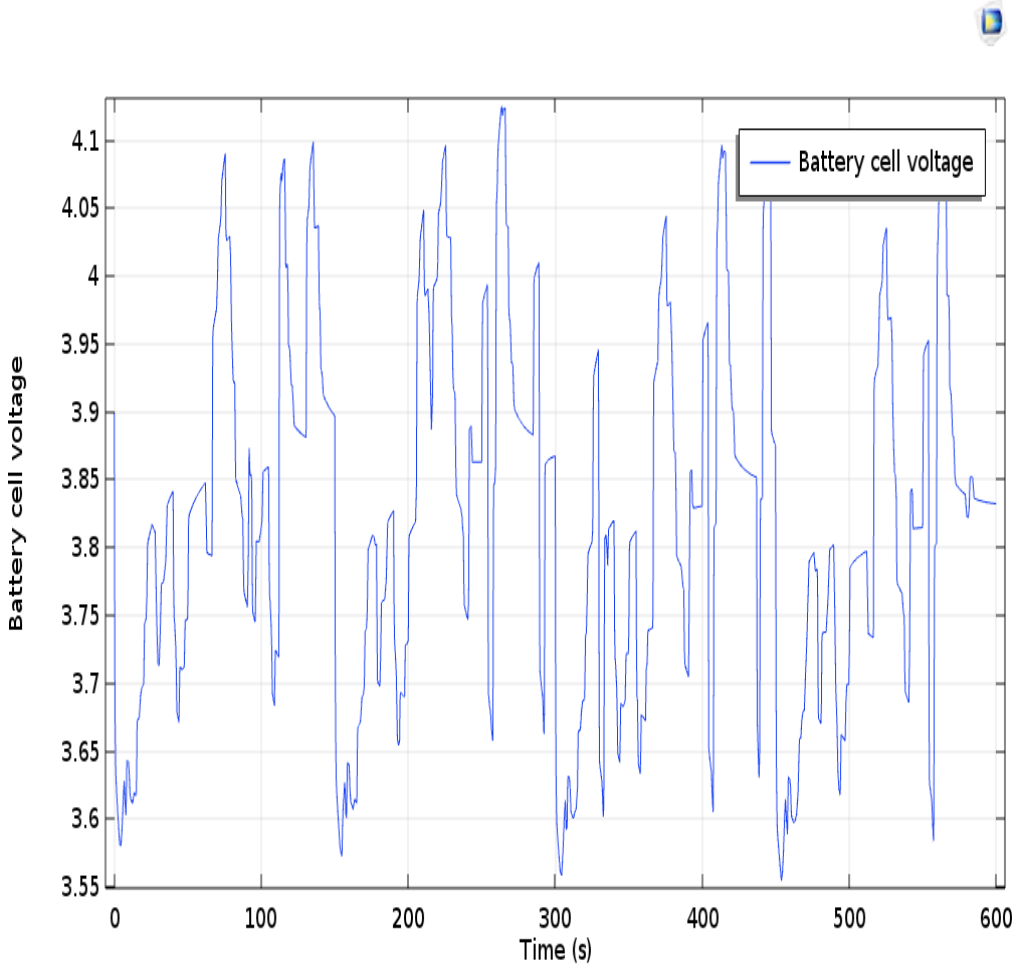


Temperature of the Battery During Drive Cycle

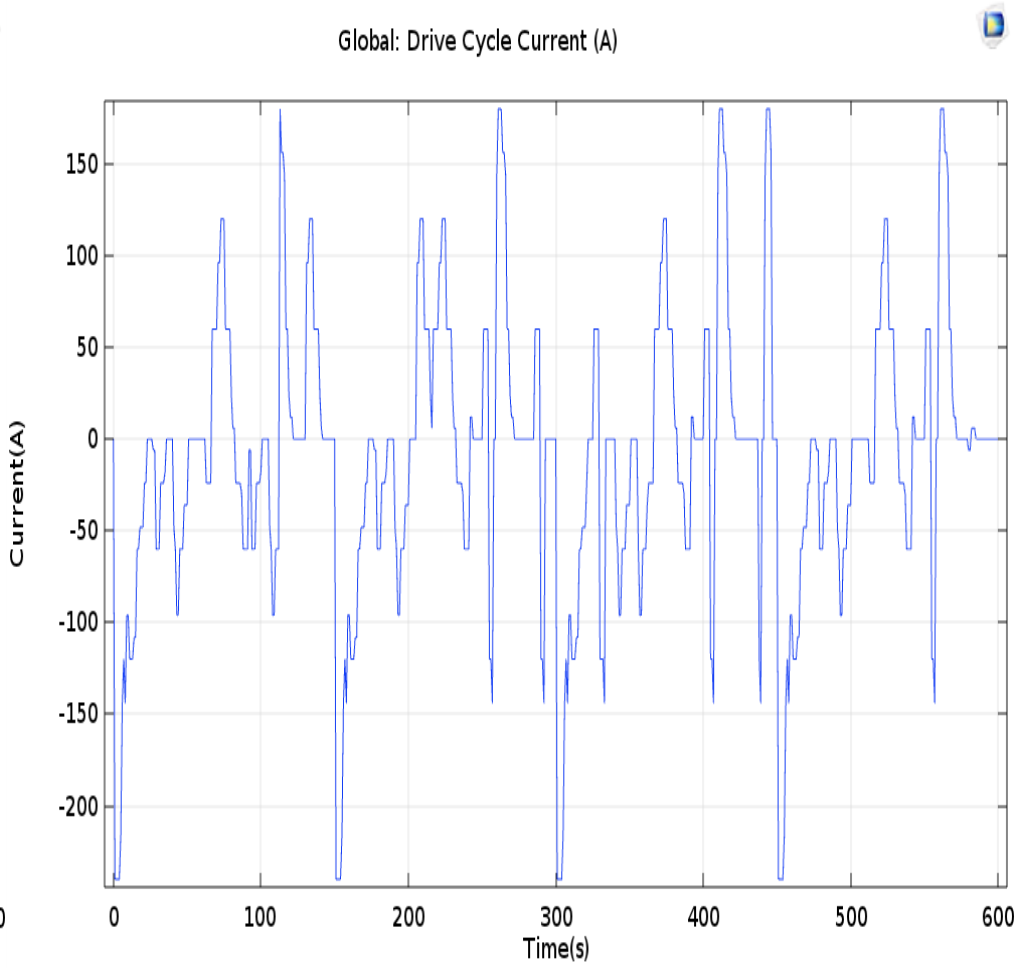


Total Polarization During Drive Cycle.

Thermal Characterization

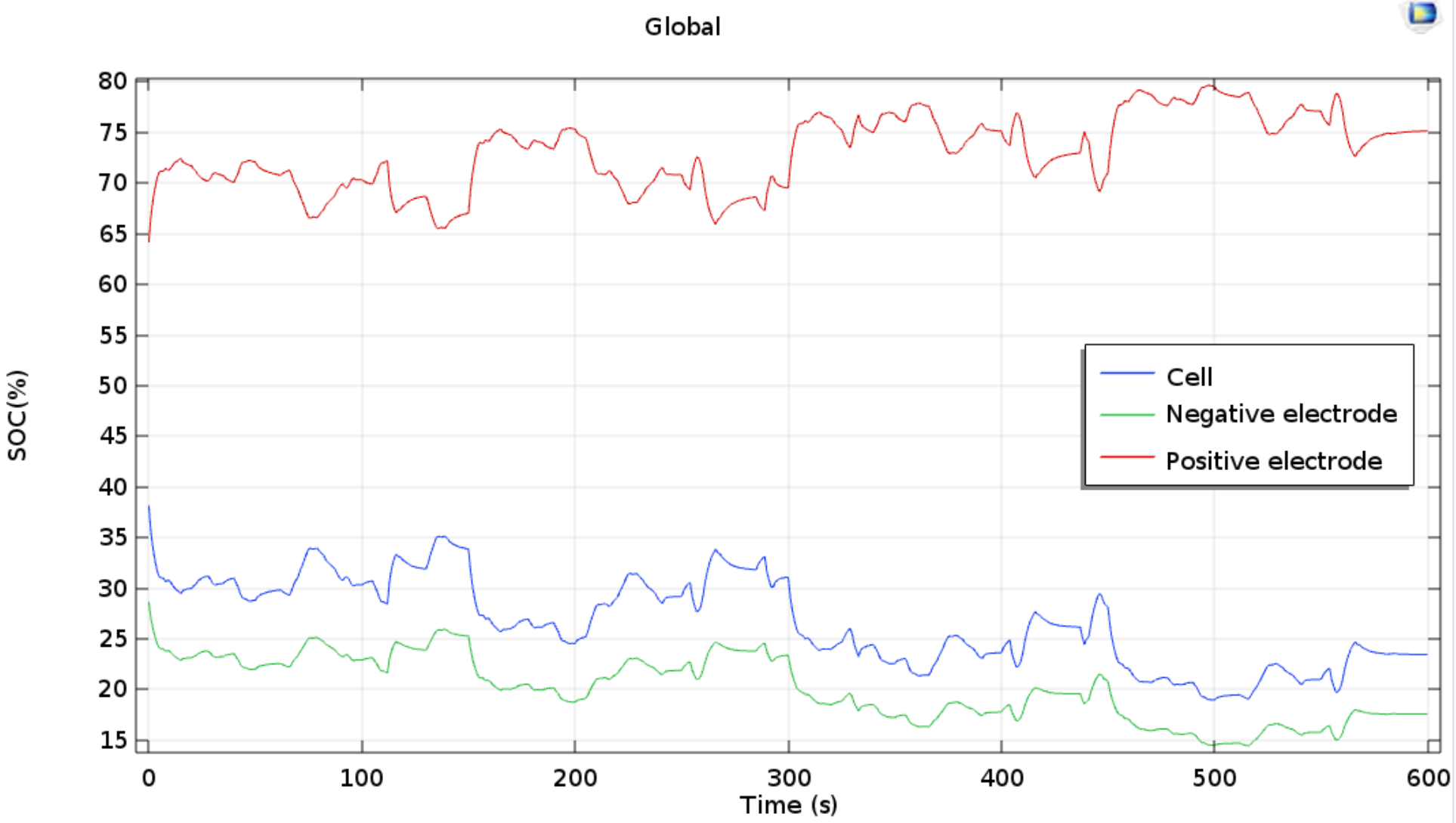


Cell Voltage during Drive Cycle.

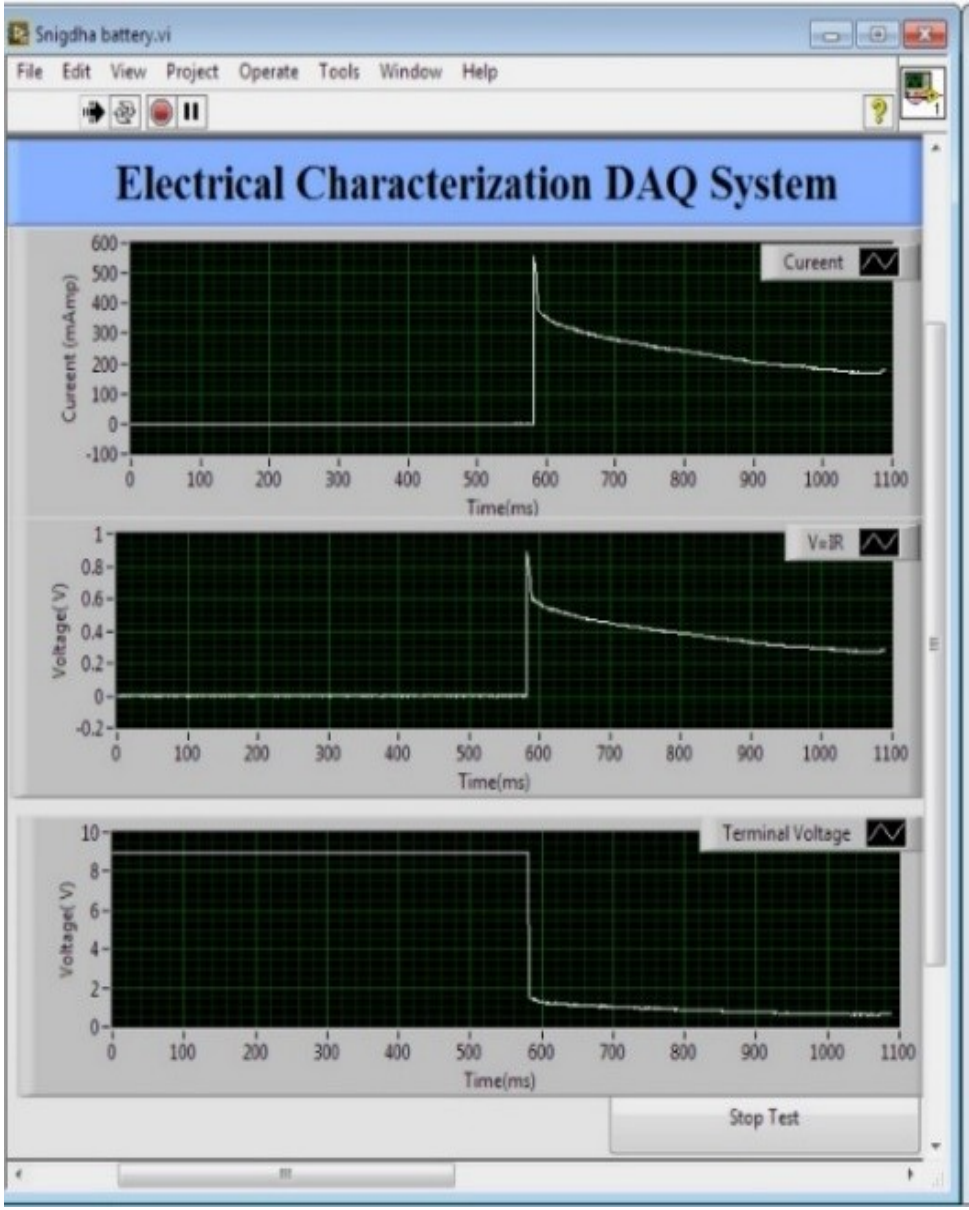


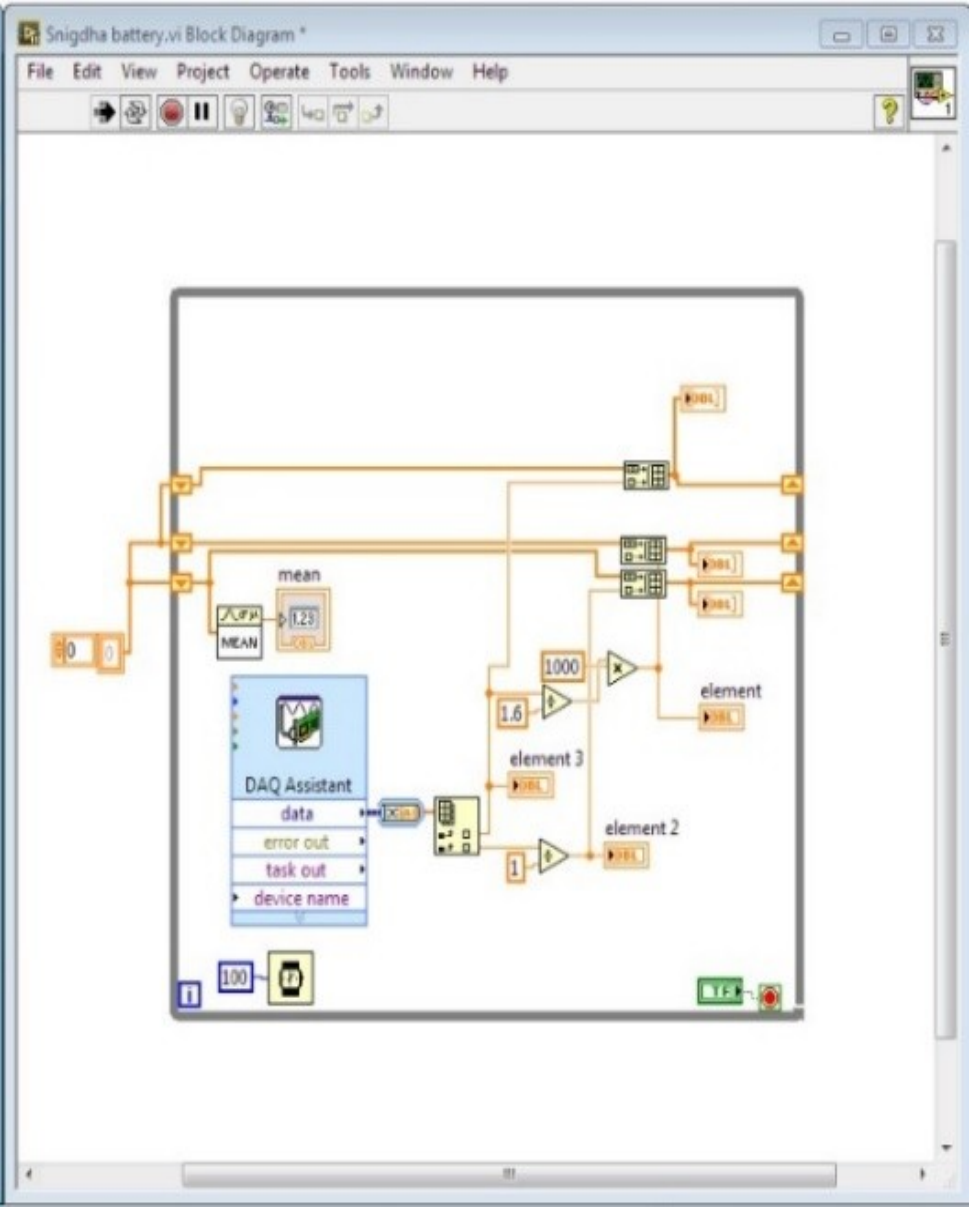
Drive Cycle Current on load.

Thermal Characterization- State of charge of battery during the drive cycle

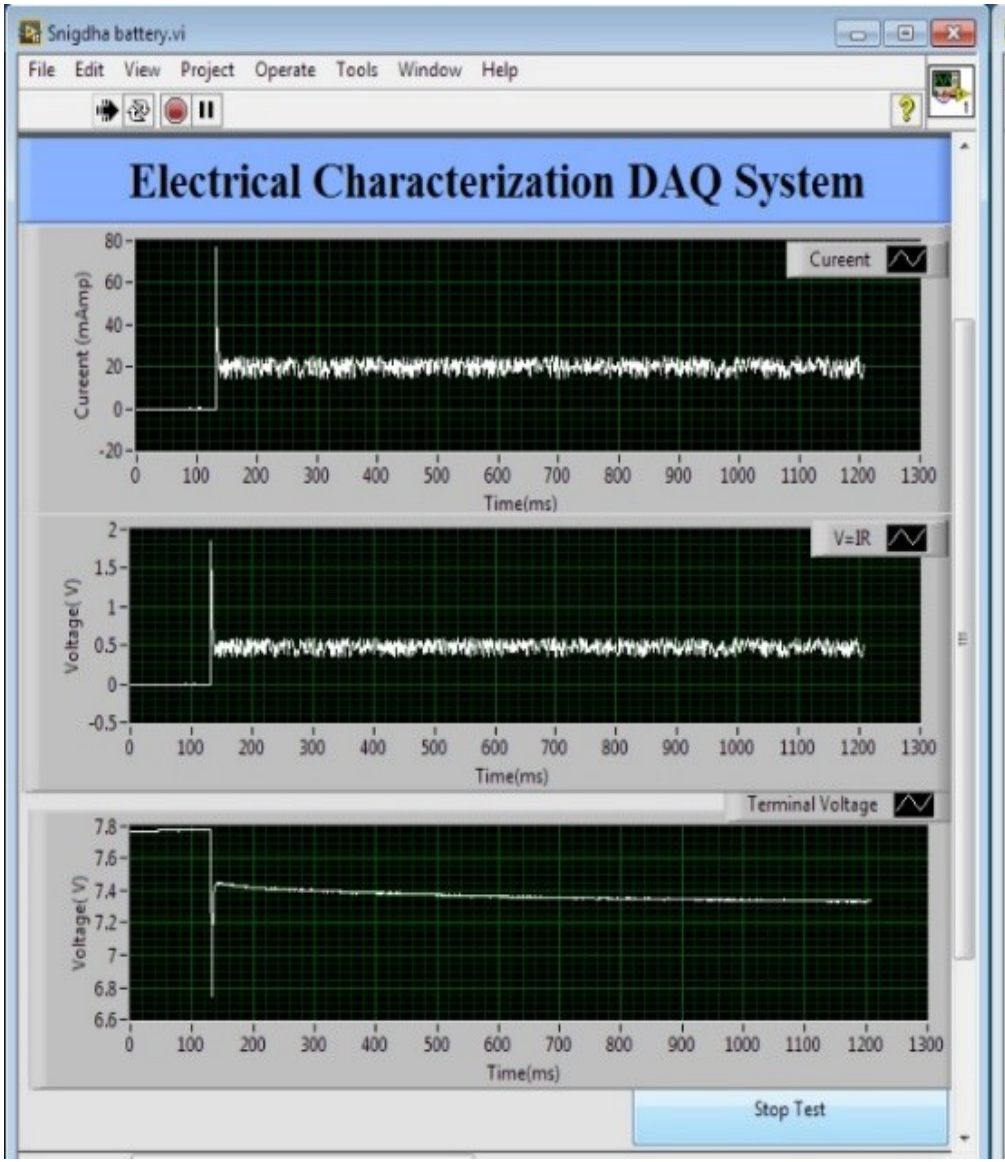


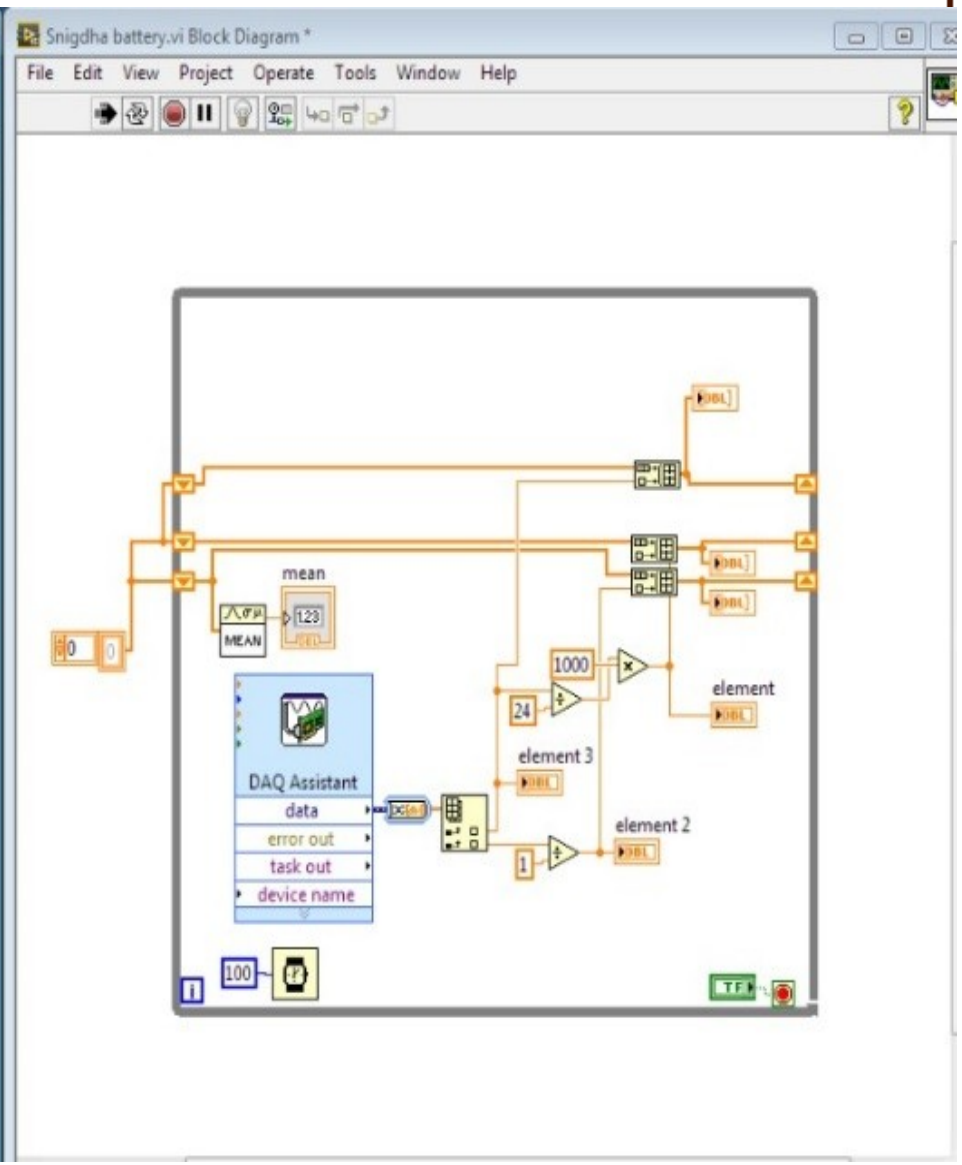
Electrical Characterization- Short Circuit Test Waveform



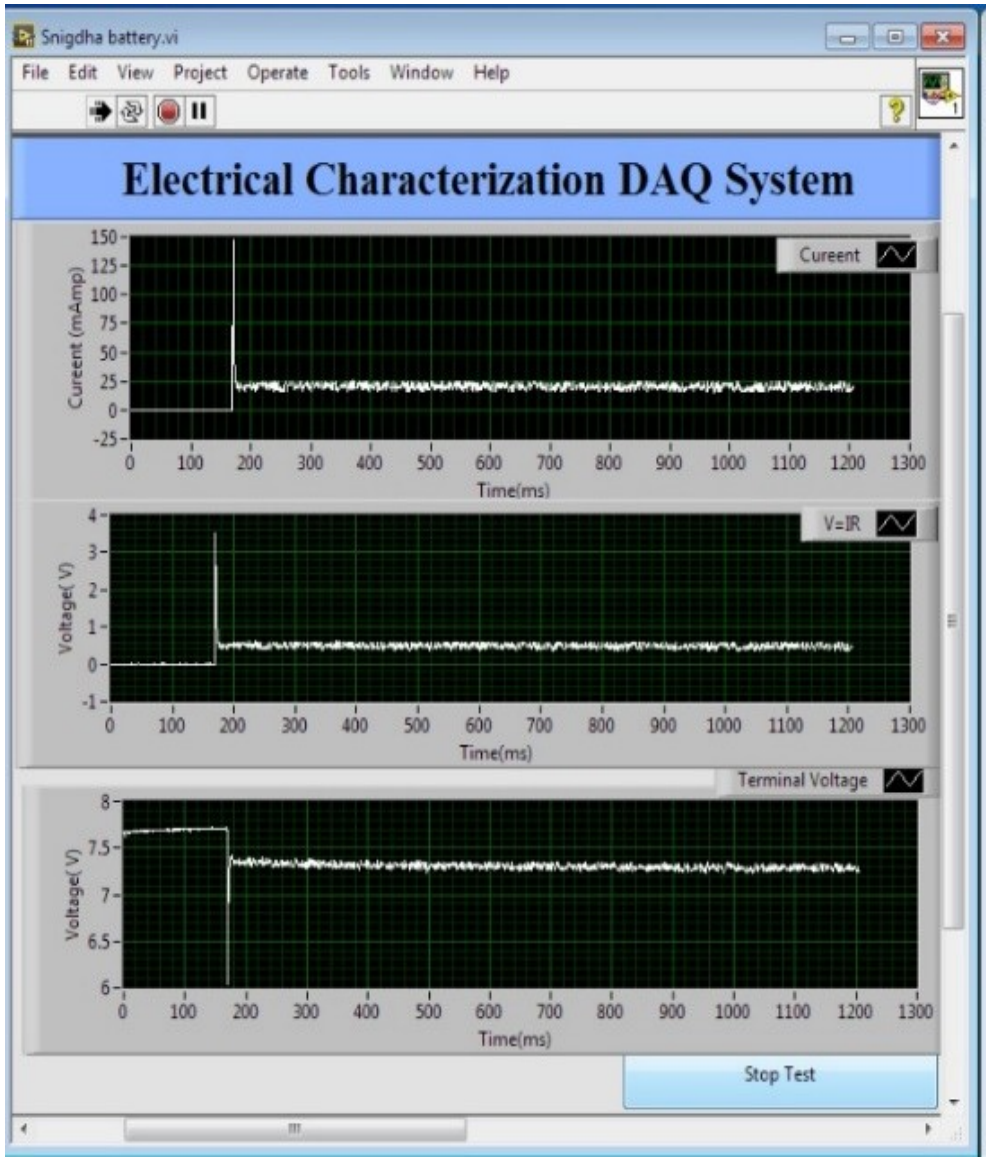


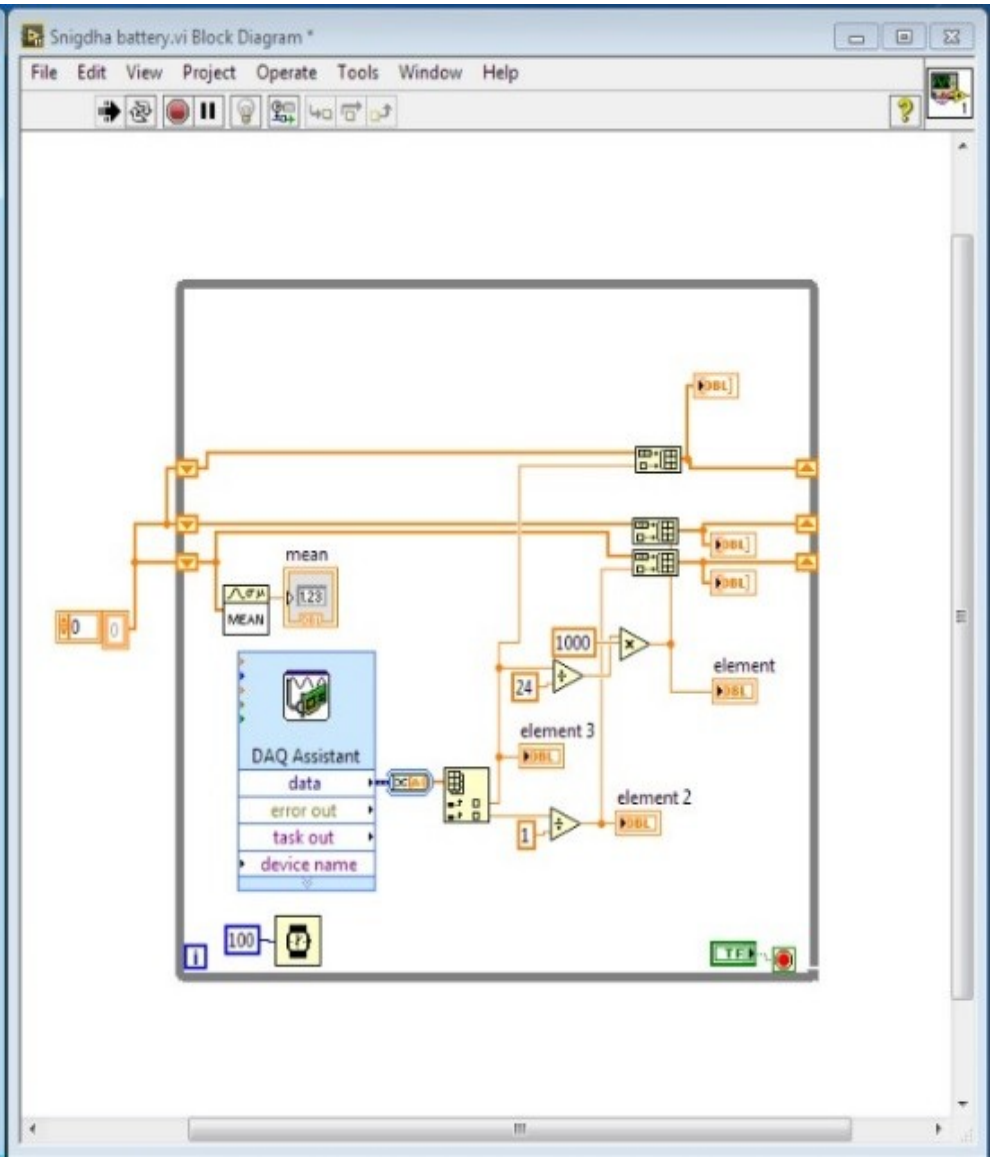
Electrical Characterization- Current And Voltage Waveform Without Capacitor Bank





Electrical Characterization-Current And Voltage Waveform With Capacitor Bank





Conclusion

- In this study, theoretical and hardware simulations were successfully performed using Multiphysics COMSOL and LabVIEW software for thermal and electrical characterization of Li-ion battery.
- Thermal characterization of Li-ion battery shows better heat management and good polarization capability with longer drive cycles for Li-ion batteries.
- During electrical characterization, a real-time hardware circuit was successfully designed to integrate the data acquisition unit with software.
- These electrical characterization results show good electrical behavior of the Li-ion battery bank as well as high accelerating current across the circuit in presence of a capacitor bank.

References

1. Thierry Cassagneau and Janos H Fendler, Advanced Materials **10** (11), 877 (1998).
2. Xiaosong Hu, Fengchun Sun, and Yuan Zou, Energies **3** (9), 1586 (2010).
3. Rakesh Saroha, Amrish K Panwar, and Yogesh Sharma, Ceramics International **43** (7), 5734 (2017).
4. Linda Ager-Wick Ellingsen, Guillaume Majeau-Bettez, Bhawna Singh, Akhilesh Kumar Srivastava, Lars Ole Valøen, and Anders Hammer Strømman, Journal of Industrial Ecology **18** (1), 113 (2014).
5. Rakesh Saroha, Amrish K Panwar, Yogesh Sharma, Pawan K Tyagi, and Sudipto Ghosh, Applied Surface Science **394**, 25 (2017).
6. Mingzhi Dai, Changhe Huo, Qiang Zhang, Karim Khan, Xingye Zhang, and Cai Shen, Advanced Theory and Simulations **1** (10), 1800023 (2018).
7. Farid Bahiraei, Maryam Ghalkhani, Amir Fartaj, and Gholam-Abbas Nazri, Applied Thermal Engineering **125**, 904 (2017).
8. Christian Hellwig, Seniz Sörgel, and Wolfgang G Bessler, ECS Transactions **35** (32), 215 (2011).
9. Xiao-song Hu, Feng-Chun Sun, and Yuan Zou, Journal of Central South University of Technology **18** (5), 1525 (2011).
10. Asan Gani and MJE Salami, presented at the Student conference on research and development, 2002 (unpublished).
11. Qiu Tang, Zhaosheng Teng, Siyu Guo, and Yaonan Wang, presented at the 2009 International Conference on Measuring Technology and Mechatronics Automation, 2009 (unpublished).
12. Long Cai and Ralph E White, Journal of Power Sources **196** (14), 5985 (2011).
13. Chih-Chiang Hua and Meng-Yu Lin, presented at the ISIE'2000. Proceedings of the 2000 IEEE International Symposium on Industrial Electronics (Cat. No. 00TH8543), 2000 (unpublished).
14. Stephan Buller, Eckhard Karden, Andreas Lohner, and Rik W De Doncker, presented at the INTELEC-Twentieth International Telecommunications Energy Conference (Cat. No. 98CH36263), 1998 (unpublished).
15. Anif Jamaluddin, Louis Sihombing, Agus Supriyanto, Agus Purwanto, and M Nizam, presented at the 2013 Joint International Conference on Rural Information & Communication Technology and Electric-Vehicle Technology (ICT & ICeV-T), 2013 (unpublished).

दिल्ली प्रौद्योगिकी विश्वविद्यालय
DELTECH
UNIVERSITY
DELHI TECHNOLOGICAL UNIVERSITY

THANK YOU

Advances in Intelligent Systems and Computing 1042

Neha Sharma
Amlan Chakrabarti
Valentina Emilia Balas *Editors*

Data Management, Analytics and Innovation

Proceedings of ICDMAI 2019, Volume 1

Springer

Advances in Intelligent Systems and Computing

Volume 1042

Series Editor

Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Advisory Editors

Nikhil R. Pal, Indian Statistical Institute, Kolkata, India
Rafael Bello Perez, Faculty of Mathematics, Physics and Computing, Universidad Central de Las Villas, Santa Clara, Cuba
Emilio S. Corchado, University of Salamanca, Salamanca, Spain
Hani Hagras, School of Computer Science and Electronic Engineering, University of Essex, Colchester, UK
László T. Kóczy, Department of Automation, Széchenyi István University, Gyor, Hungary
Vladik Kreinovich, Department of Computer Science, University of Texas at El Paso, El Paso, TX, USA
Chin-Teng Lin, Department of Electrical Engineering, National Chiao Tung University, Hsinchu, Taiwan
Jie Lu, Faculty of Engineering and Information Technology, University of Technology Sydney, Sydney, NSW, Australia
Patricia Melin, Graduate Program of Computer Science, Tijuana Institute of Technology, Tijuana, Mexico
Nadia Nedjah, Department of Electronics Engineering, University of Rio de Janeiro, Rio de Janeiro, Brazil
Ngoc Thanh Nguyen, Faculty of Computer Science and Management, Wrocław University of Technology, Wrocław, Poland
Jun Wang, Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing such as: computational intelligence, soft computing including neural networks, fuzzy systems, evolutionary computing and the fusion of these paradigms, social intelligence, ambient intelligence, computational neuroscience, artificial life, virtual worlds and society, cognitive science and systems, Perception and Vision, DNA and immune based systems, self-organizing and adaptive systems, e-Learning and teaching, human-centered and human-centric computing, recommender systems, intelligent control, robotics and mechatronics including human-machine teaming, knowledge-based paradigms, learning paradigms, machine ethics, intelligent data analysis, knowledge management, intelligent agents, intelligent decision making and support, intelligent network security, trust management, interactive entertainment, Web intelligence and multimedia.

The publications within "Advances in Intelligent Systems and Computing" are primarily proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

**** Indexing: The books of this series are submitted to ISI Proceedings, EI-Compendex, DBLP, SCOPUS, Google Scholar and Springerlink ****

More information about this series at http://www.springer.com/series/11156

Neha Sharma · Amlan Chakrabarti ·
Valentina Emilia Balas
Editors

Data Management, Analytics and Innovation

Proceedings of ICDMAI 2019, Volume 1

Springer

Editors
Neha Sharma
Society for Data Science
Pune, Maharashtra, India

Valentina Emilia Balas
Department of Automatics
and Applied Software
Aurel Vlaicu University of Arad
Arad, Romania

Amlan Chakrabarti
A.K. Choudhury School of Information
Technology
University of Calcutta
Kolkata, West Bengal, India

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-981-32-9948-1 ISBN 978-981-32-9949-8 (eBook)
https://doi.org/10.1007/978-981-32-9949-8

© Springer Nature Singapore Pte Ltd. 2020
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

These two volumes constitute the proceedings of the International Conference on *Data Management, Analytics and Innovation* (ICDMAI 2019) held from 18 to 20 January 2019. ICDMAI is a flagship conference of Society for Data Science, which is a non-profit professional association established to create a collaborative platform for bringing together technical experts across industry, academia, government laboratories and professional bodies to promote innovation around data science. ICDMAI 2019 envisions its role towards data science and its enhancement through collaboration, innovative methodologies and connections throughout the globe. The conference was hosted by Lincoln University College, Kuala Lumpur, Malaysia, and was supported by the industry leaders like IBM. Other partners for the conference were Wizer and DSMS College of Tourism & Management, West Bengal, India. The conference witnessed participants from 20 countries, 12 industries, 31 international universities and 94 premier Indian universities. Utmost care was taken in each and every facet of the conference, especially regarding the quality of the paper submissions. Out of 418 papers submitted to ICDMAI 2019, only 20% (87 papers) are selected for an oral presentation after a rigorous review process. Besides paper presentation, the conference also showcased workshops, tutorial talks, keynote sessions and plenary talks by the experts of the respective field.

The volumes cover a broad spectrum of computer science, information technology, computational engineering, electronics and telecommunication, electrical, computer application and all the relevant disciplines. The conference papers included in these proceedings are published post-conference and are grouped into the four areas of research such as data management and smart informatics; big data management; artificial intelligence and data analytics; and advances in network technologies. All the four tracks of the conference were very relevant to the current technological advancements and had Best Paper Award in each track. Very stringent selection process was adopted for paper selection, and from plagiarism check to technical chairs' review to double-blind review, every step was religiously followed. We compliment all the authors for submitting high-quality research papers to ICDMAI 2019. We would like to acknowledge all the authors for their contributions and also the efforts taken by reviewers and session chairs of the conference,

without whom it would have been difficult to select these papers. We appreciate the unconditional support from the members of the National and International Program Committee. It was really interesting to hear the participants of the conference highlight the new areas and the resulting challenges as well as opportunities. This conference has served as a vehicle for a spirited debate and discussion on many challenges that the world faces today.

We especially thank our General Chair, Dr. P. K. Sinha; Vice Chancellor and Director, Dr. S. P. Mukherjee, International Institute of Information Technology, Naya Raipur (IIIT-NR), Chhattisgarh; other eminent personalities like Mr. Eddy Liew, Cloud and Solutions Country Technical Leader at IBM, Malaysia; Kranti Athalye, Sr. Manager in IBM India University Relations; Dr. Juergen Seitz, Head of Business Information Systems Department, Baden-Wuerttemberg Cooperative State University, Heidenheim, Germany; Mr. Aninda Bose, Senior Publishing Editor, Springer India Pvt. Ltd.; Dr. Vincenzo Piuri, IEEE Fellow, University of Milano, Italy; Hanaa Hachimi, National School of Applied Sciences ENSA in Kenitra, Morocco; Amol Dhondse, IBM Senior Solution Architect; Mohd Helmy Abd Wahab, Senior Lecturer and Former Head of Multimedia Engineering Lab at the Department of Computer Engineering, Faculty of Electrical and Electronic Engineering, Universiti Tun Hussein Onn Malaysia (UTHM); Anand Nayyar, Duy Tan University, Vietnam; and many more who were associated with ICDMAI 2019. Besides, there was CSI-Startup and Entrepreneurship Award to felicitate budding job creators.

Our special thanks go to Janus Kacprzyk (Editor in Chief, Springer, Advances in Intelligent Systems and Computing Series) for the opportunity to organize this guest-edited volume. We are grateful to Springer, especially to Mr. Aninda Bose (Senior Publishing Editor, Springer India Pvt. Ltd.), for the excellent collaboration, patience and help during the evolvement of this volume.

We are confident that the volumes will provide the state-of-the-art information to professors, researchers, practitioners and graduate students in the area of data management, analytics and innovation, and all will find this collection of papers inspiring and useful.

Pune, India — Neha Sharma
Kolkata, India — Amlan Chakrabarti
Arad, Romania — Valentina Emilia Balas

Contents

Data Management and Smart Informatics

Empirical Study of Soft Clustering Technique for Determining Click Through Rate in Online Advertising 3
Akshi Kumar, Anand Nayyar, Shubhangi Upasani and Arushi Arora

ASK Approach: A Pre-migration Approach for Legacy Application Migration to Cloud 15
Sanjeev Kumar Yadav, Akhil Khare and Choudhary Kavita

A Fuzzy Logic Based Cardiovascular Disease Risk Level Prediction System in Correlation to Diabetes and Smoking 29
Kanak Saxena and Umesh Banodha

An Integrated Fault Classification Approach for Microgrid System 41
Ruchita Nale, Ruchi Chandrakar and Monalisa Biswal

Role of Data Analytics in Human Resource Management for Prediction of Attrition Using Job Satisfaction 57
Neerja Aswale and Kavya Mukul

A Study of Business Performance Management in Special Reference to Automobile Industry 69
Gurinder Singh, Smiti Kashyap, Kanika Singh Tomar and Vikas Garg

Secure Online Voting System Using Biometric and Blockchain 93
Dipti Pawade, Avani Sakhapara, Aishwarya Badgujar, Divya Adepu and Melvita Andrade

An Approach: Applicability of Existing Heterogeneous Multicore Real-Time Task Scheduling in Commercially Available Heterogeneous Multicore Systems 111
Kalyan Baital and Amlan Chakrabarti

Analyzing the Detectability of Harmful Postures for Patient with Hip Prosthesis Based on a Single Accelerometer in Mobile Phone 125
Kitti Naonueng, Opas Chutatape and Rong Phoophuangpairoj

Software Development Process Evolution in Malaysian Companies 139
Rehan Akbar, Asif Riaz Khan and Kiran Adnan

Automated Scheduling of Hostel Room Allocation Using Genetic Algorithm 151
Rayner Alfred and Hin Fuk Yu

Evaluation of ASTER TIR Data-Based Lithological Indices in Parts of Madhya Pradesh and Chhattisgarh State, India 161
Himanshu Govil, Subhanil Guha, Prabhat Diwan, Neetu Gill and Anindita Dey

Analyzing Linear Relationships of LST with NDVI and MNDISI Using Various Resolution Levels of Landsat 8 OLI and TIRS Data 171
Himanshu Govil, Subhanil Guha, Prabhat Diwan, Neetu Gill and Anindita Dey

Automatic Robot Processing Using Speech Recognition System 185
S. Elavarasi and G. Suseendran

Banking and FinTech (Financial Technology) Embraced with IoT Device 197
G. Suseendran, E. Chandrasekaran, D. Akila and A. Sasi Kumar

Big Data Management

GRNN++: A Parallel and Distributed Version of GRNN Under Apache Spark for Big Data Regression 215
Sk. Kamaruddin and Vadlamani Ravi

An Entropy-Based Technique for Conferences Ranking 229
Fiaz Majeed and Rana Azhar Ul Haq

MapReduce mRMR: Random Forests-Based Email Spam Classification in Distributed Environment 241
V. Sri Vinitha and D. Karthika Renuka

The Impact of Sustainable Development Report Disclosure on Tax Planning in Thailand 255
Sathaya Thanjunpong and Thatphong Awirothananon

Clustering and Labeling Auction Fraud Data 269
Ahmad Alzahrani and Samira Sadaoui

Big Data Security Challenges and Preventive Solutions 285
Nirmal Kumar Gupta and Mukesh Kumar Rohil

Role and Challenges of Unstructured Big Data in Healthcare 301
Kiran Adnan, Rehan Akbar, Siak Wang Khor and Adnan Bin Amanat Ali

Zip Zap Data—A Framework for 'Personal Data Preservation' 325
K. Arunkumar and A. Devendran

A Systematic Mapping Study of Cloud Large-Scale Foundation—Big Data, IoT, and Real-Time Analytics 339
Isaac Odun-Ayo, Rowland Goddy-Worlu, Temidayo Abayomi-Zannu and Emanuel Grant

Studies on Radar Imageries of Thundercloud by Image Processing Technique 365
Sonia Bhattacharya and Himadri Bhattacharyya Chakrabarty

Artificial Intelligence and Data Analysis

PURAN: Word Prediction System for Punjabi Language News 383
Gurjot Singh Mahi and Amandeep Verma

Implementation of hDE-HTS Optimized T2FPID Controller in Solar-Thermal System 401
Binod Shaw, Jyoti Ranjan Nayak and Rajkumar Sahu

Design of Sigma-Delta Converter Using 65 nm CMOS Technology for Nerves Organization in Brain Machine Interface 413
Anil Kumar Sahu, G. R. Sinha and Sapna Soni

Performance Comparison of Machine Learning Techniques for Epilepsy Classification and Detection in EEG Signal 425
Rekh Ram Janghel, Archana Verma and Yogesh Kumar Rathore

Novel Approach for Plant Disease Detection Based on Textural Feature Analysis 439
Varinderjit Kaur and Ashish Oberoi

Novel Approach for Brain Tumor Detection Based on Naïve Bayes Classification 451
Gurkarandesh Kaur and Ashish Oberoi

Automatic Classification of Carnatic Music Instruments Using MFCC and LPC 463
Surendra Shetty and Sarika Hegde

Semiautomated Ontology Learning to Provide Domain-Specific Knowledge Search in Marathi Language 475
Neelam Chandolikar, Pushkar Joglekar, Shivjeet Bhosale, Dipali Peddawad, Rajesh Jalnekar and Swati Shilaskar

Identifying Influential Users on Social Network: An Insight 489
Ragini Krishna and C. M. Prashanth

Factex: A Practical Approach to Crime Detection 503
Rachna Jain, Anand Nayyar and Shivam Bachhety

Analysis of Classification Algorithms for Breast Cancer Prediction 517
S. P. Rajamohana, K. Umamaheswari, K. Karunya and R. Deepika

Real-Time Footfall Prediction Using Weather Data: A Case on Retail Analytics 529
Garima Makkar

Normal Pressure Hydrocephalus Detection Using Active Contour Coupled Ensemble Based Classifier 543
Pallavi Saha, Sankhadeep Chatterjee, Santanu Roy and Soumya Sen

Question–Answer System on Episodic Data Using Recurrent Neural Networks (RNN) 555
Vineet Yadav, Vishnu Bharadwaj, Alok Bhatt and Ayush Rawal

Convoluted Cosmos: Classifying Galaxy Images Using Deep Learning 569
Diganta Misra, Sachi Nandan Mohanty, Mohit Agarwal and Suneet K. Gupta

Advances in Network Technologies

Energy-Based Improved MPR Selection in OLSR Routing Protocol 583
Rachna Jain and Indu Kashyap

A Novel Approach for Better QoS in Cognitive Radio Ad Hoc Networks Using Cat Optimization 601
Lolita Singh and Nitul Dutta

(T-ToCODE): A Framework for Trendy Topic Detection and Community Detection for Information Diffusion in Social Network 613
Reena Pagare, Akhil Khare and Shankar Chaudhary

ns-3 Implementation of Network Mobility Basic Support (NEMO-BS) Protocol for Intelligent Transportation Systems 633
Prasanta Mandal, Manoj Kumar Rana, Punyasha Chatterjee and Arpita Debnath

Modified DFA Minimization with Artificial Bee Colony Optimization in Vehicular Routing Problem with Time Windows 643
G. Niranjani and K. Umamaheswari

Coverage-Aware Recharge Scheduling Scheme for Wireless Charging Vehicles in the Wireless Rechargeable Sensor Networks 663
Govind P. Gupta and Vrajesh Kumar Chawra

A Transition Model from Web of Things to Speech of Intelligent Things in a Smart Education System 673
Ambrose A. Azeta, Victor I. Azeta, Sanjay Misra and M. Ananya

Intrusion Detection and Prevention Systems: An Updated Review 685
Nureni Ayofe Azeez, Taiwo Mayowa Bada, Sanjay Misra, Adewole Adewumi, Charles Van der Vyver and Ravin Ahuja

Simulation-Based Performance Analysis of Location-Based Opportunistic Routing Protocols in Underwater Sensor Networks Having Communication Voids 697
Sonali John, Varun G. Menon and Anand Nayyar

A Hybrid Optimization Algorithm for Pathfinding in Grid Environment 713
B. Booba, A. Prema and R. Renugadevi

Dynamic Hashtag Interactions and Recommendations: An Implementation Using Apache Spark Streaming and GraphX 723
Sonam Sharma

Author Index 739

About the Editors

Neha Sharma is the Founder Secretary of the Society for Data Science, India. She was the Director of the Zeal Institute of Business Administration, Computer Application & Research, Pune, Maharashtra, India, and and Deputy Director, of the Padmashree Dr. D.Y.Patil Institute of Master of Computer Applications, Akurdi, Pune. She completed her PhD at the prestigious Indian Institute of Technology (IIT-ISM), Dhanbad, and she is a Senior IEEE member as well as Execom member of IEEE Pune Section. She has published numerous research papers in respected international journals. She received the "Best PhD Thesis Award" and "Best Paper Presenter at International Conference Award" from the Computer Society of India. Her areas of interest include data mining, database design, analysis and design, artificial intelligence, big data, cloud computing, blockchain and data science.

Amlan Chakrabarti currently the Dean of the Faculty of Engineering and Technology, Professor and Director of the A.K.Choudhury School of Information Technology, University of Calcutta, India. He was a postdoctoral fellow at the School of Engineering, Princeton University, USA from 2011 to 2012. He has published around 130 research papers in refereed journals and conferences, and has been involved in research projects supported by various national funding agencies and international collaborations. He is a senior member of the IEEE and ACM, ACM Distinguished Speaker, Vice President of the Society for Data Science, and Secretary of the

IEEE CEDA India Chapter. He is also the Guest Editor of the Springer Nature Journal on Applied Sciences. His research interests include quantum computing, VLSI design, embedded system design, computer vision and analytics.

Valentina Emilia Balas is currently a Full Professor at the Department of Automatics and Applied Software at the Faculty of Engineering, "Aurel Vlaicu" University of Arad, Romania. She holds a Ph.D. in Applied Electronics and Telecommunications from the Polytechnic University of Timisoara. Dr. Balas is the author of more than 300 research papers in refereed journals and international conferences. Her research interests include intelligent systems, fuzzy control, soft computing, smart sensors, information fusion, modeling and simulation. She is the editor-in-chief of the International Journal of Advanced Intelligence Paradigms (IJAIP) and the International Journal of Computational Systems Engineering (IJCSysE), and is an editorial board member of several national and international journals.

Data Management and Smart Informatics

Empirical Study of Soft Clustering Technique for Determining Click Through Rate in Online Advertising

Akshi Kumar, Anand Nayyar, Shubhangi Upasani and Arushi Arora

Abstract Online advertising is an industry with the potential for maximum revenue extraction. Displaying the ad which is more likely to be clicked plays a crucial role in generating maximum revenue. A high click through rate (CTR) is an indication that the user finds the ad useful and relevant. For suitable placement of ads online and rich user experience, determining CTR has become imperative. Accurate estimation of CTR helps in placement of advertisements in relevant locations which would result in more profits and return of investment for the advertisers and publishers. This paper presents the application of a soft clustering method namely fuzzy c-means (FCM) clustering for determining if a particular ad would be clicked by the user or not. This is done by classifying the ads in the dataset into broad clusters depending on whether they were actually clicked or not. This way the kind of advertisements that the user is interested in can be found out and subsequently more advertisements of the same kind can be recommended to him, thereby increasing the CTR of the displayed ads. Experimental results show that FCM outperforms k-means clustering (KMC) in determining CTR.

Keywords Advertising · Click Through Rate · Clustering

A. Kumar
Department of Computer Science and Engineering, Delhi Technological University, Delhi, India
e-mail: akshi.kumar@gmail.com

A. Nayyar
Graduate School, Duy Tan University, Da Nang, Viet Nam
e-mail: anandnayyar@duytan.edu.vn

S. Upasani
Department of Electronics and Communication Engineering, Delhi Technological University, Delhi, India
e-mail: shubhangi.upasani@gmail.com

A. Arora (✉)
Department of Electrical Engineering, Delhi Technological University, Delhi, India
e-mail: aroraarushi1997@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_1

1 Introduction

Online advertising has become an important source of revenue for a wide range of businesses. With the tremendous growth in online advertising each year, it has also taken over a major area of research. Most of the revenues generated by widely used search engines as well as prevalent websites come from advertisements. It is therefore important to display relevant advertisements (ads) to the users and avoid the advertisements that are often disliked by them.

Online advertising is more economical than traditional ways of advertising like mass markets and niche media. Internet ads have a wider audience and can be viewed for days and nights altogether, in contrast to ads on television and radios that last for shorter durations and are displayed with limited frequency. Market segmentation is much more effective over the Internet than in any other medium. Thorough study of markets, customer preferences and habits and segmenting consumers into cohesive groups can be done efficiently through online advertising. Online advertising also offers small businesses numerous benefits like robust targeting, consumer insights and more effective return on investment.

Advertisements fall into two broad categories—sponsored search advertisements and contextual advertisements. Sponsored search ads are displayed on the same web pages that show results of search queries entered by users. The core purpose behind sponsored advertising is to enhance the advertiser's brand image as the ads displayed have the same form and qualities as the advertiser's original content. Contextual advertising, on the other hand, uses automated systems that display ads relevant to the user's identity and website's content. Google AdSense is one of the many well-known examples of contextual advertising. Google robots display only those ads that the users find relevant and useful. When a user visits a website, features like ad size, ad placement, etc. are extracted from the search query and sent to a server. Relevant ads are selected based on user's past history, CTR and other data. Increasing the number of ads is not a good idea as it will shoot up the earnings only for a short time before the user switches to other search engines due to poor user experience. To maximize revenue, precise placements of ads are therefore required.

CTR refers to the number of times the advertiser's ad has been clicked (clicks) divided by number of times the ad appears on the screen (impression). Relevant placement of ad is a precondition for increasing the CTR for it. Ad performance can also be measured using CTR. CTR determination has several issues associated with it. The advertisers need to pay every time whenever an ad is clicked. It is on the basis of CTR that the search engine decides what ads are to be displayed and in what order of appearance. This is ensured by combining together the likelihood of an ad being clicked and the cost of the ad per click to create a display format that will yield maximum return. Many algorithms based on the supervised methodology have been designed to predict the CTR of advertisements like support vector machines, decision trees, Naïve Bayes, etc.

Motivated by this, the goal is to analyse and assess the application of unsupervised approach namely FCM for determining the CTR of an advertisement. This

clustering algorithm divides the total ads in the dataset into broad clusters based on whether they have been clicked or not by the user. This classification of ads helps in assessing the kind of advertisements that the user is really interested in and hence predict more of these kinds of ads. This would, in turn, result in a rise in the CTR of the displayed ads because of a greater number of clicks. The results obtained from FCM have been contrasted with KMC which has been used as a baseline model. The two techniques are assessed based on metrics like precision, recall and accuracy.

The following content is compiled as follows: Sect. 2 reviews work done by various researchers along with the brief idea of the algorithm used by them and their results. Section 3 elaborates the dataset taken, methods employed for preparing the data before implementing the algorithm and the algorithms used. The scrutiny of the experimental outcomes is done in the Sect. 4. Section 5 culminates the results along with the suggestions for future research prospects.

2 Related Work

After a thorough assessment of various studies in the past, it was found that a lot of algorithms have been proposed for the prediction of relevant ad to be displayed to the user. The author Avila Clemenshia et al. [1] in 2016, proposed a CTR prediction model using Poisson's regression, linear regression and support vector regression algorithms and displayed the ads accordingly. The dataset was provided by a digital marketing agency. Their results stated that support vector regression performed best among the three. Evaluation of the results was done on the basis of root mean squared error (RMSE) and correlation coefficient.

Authors Graepel et al. [2] in 2010 devised a new Bayesian CTR algorithm. The ad predictor presented showed better outcomes when compared to the baseline Naïve Bayes in terms of relative information gain (RIG) and areas under the curve (AUC). Hillard et al. [3] in 2010, implemented a model for estimating ad relevance. They refined it by including indirect feedback after consolidating the basic features of text overlap. In case of presence of adequate observations, click history was used. In case of no or few observations, a model was learned that could also be used for unpredicted ads. The precision, recall and f-score values were noted and improvement was observed with the new model.

A logistic regression approach was suggested by author Kondakindi et al. [4] in 2014, to predict whether an ad will be clicked or not. The dataset used for this purpose was from Avazu provided as a part of Kaggle competition. They started off with simple Naive Bayes followed by Vowpal Wabbit and finally got the best scores with logistic regression together with proper data preprocessing. Shi et al. [5] in 2016, designed a framework for prediction of CTR and average cost per click (CPC) of a keyword using some machine learning algorithms. The performance data for the advertiser's keywords was gathered from Google AdWords. The author had applied different machine learning algorithms such as regression, random forest

and gradient boosting to evaluate the prediction performance on both CTR and CPC. Results concluded that random forest transpires to be the best for both the metrics while gradient descent results are least reliable.

Wang et al. [6] in 2011, implemented a model for learning of user's click behaviours from advertisement search and click logs. Decision tree (DT), CRF, SVM and backpropagation neural networks (BPN) are the algorithms which were employed to carry out the imposition. The experimentation finally led to proving that CRF model outperformed the two baselines and SVM remarkably. Chakrabarti et al. [7] in 2008, developed a model for contextual advertising in which the revenue accrued by the site publisher and the advertising network depends upon the suitability of the ads displayed. This was followed by mapping the model to standard cosine similarity matching.

Cheng et al. [8] in 2010, developed the model for customization of click models in sponsored search. The results demonstrated that the accuracy of CTR could significantly be improved by personalized models in sponsored advertising. Edizel* et al. [9] in 2017, proposed the use of that deep convolutional neural networks for CTR prediction of an advertisement. One approach involved query-ad depiction being learned at character while the second entail word-level model by pretrained word vectors. The conclusion signified better outcome than the standard machine learning algorithms trained with well-defined features.

The analysis of the background work clearly connotes that most of the algorithm used in the past are supervised classification algorithms that involves two classes namely clicked and not clicked-demonstrating whether advertisement was clicked or not. The study of unsupervised clustering is still not much discovered in this domain to the extent of our understanding. This research paper is an endeavour to compare FCM and KMC algorithm using the dataset.

3 Data Characteristics

The following Fig. 1 demonstrates the system architecture of the research.

The following subsection explains the details.

3.1 Data Collection

The dataset is acquired from Avazu (Kaggle) for the purpose of writing this paper. It contains 11 days' worth of data in order to build and test prediction models using various machine learning algorithms. As the given data is approximately 6 GB, the data taken was 10 h of data for training and 2 h of data for testing. The data is ordered chronologically and the clicks and non-clicks are sampled according to different strategies.

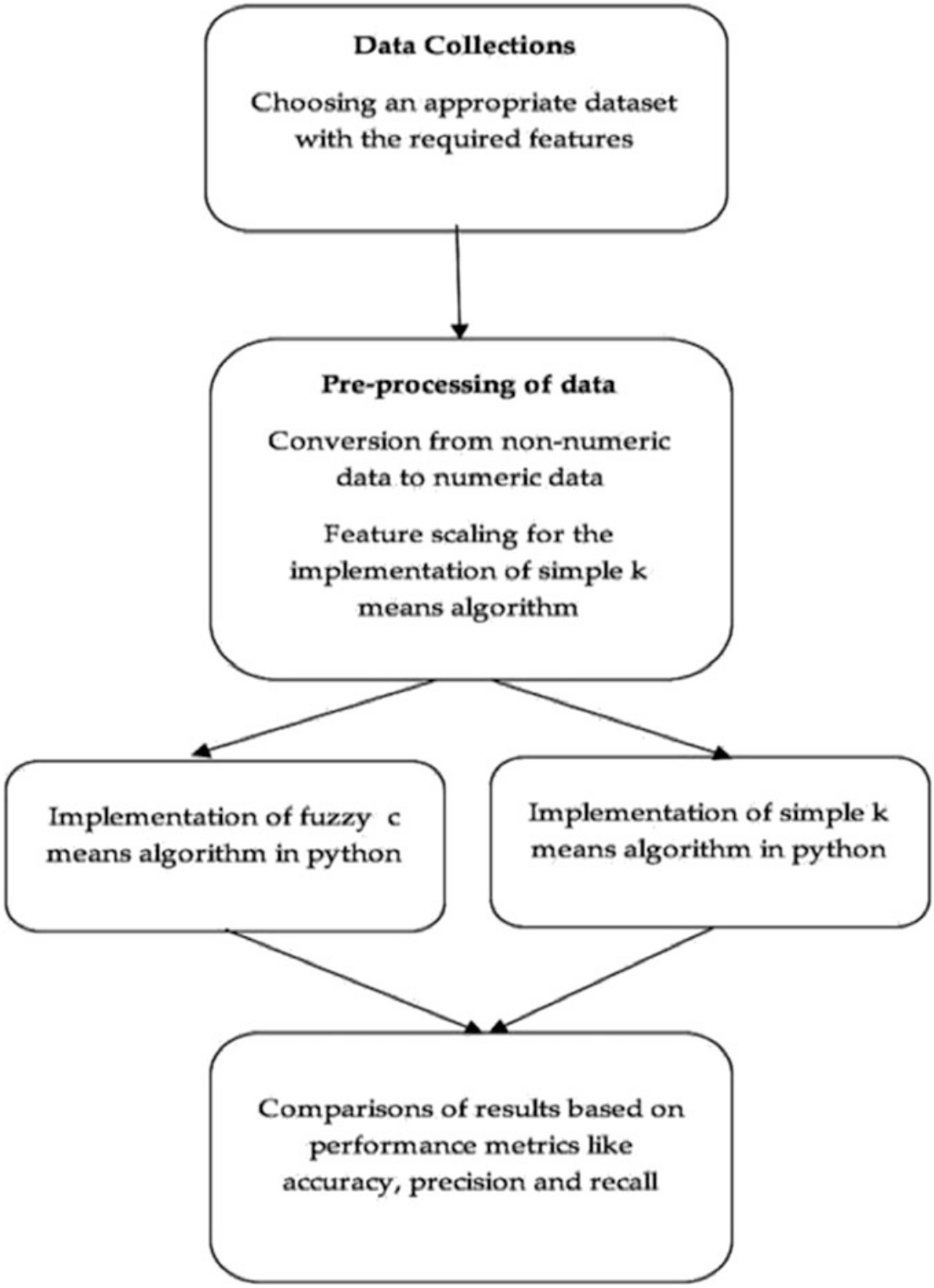


Fig. 1 System architecture

The following features are included in the dataset:

- id (unique id to identify advertisement)
- click (0 for not clicked and 1 for clicked)

- hour (in format of YYMMDDHH; it refers to the date and time the data is recorded)
- banner_pos (0 for top and 1 for bottom)
- site_id (refers to the website id)
- site_domain (website domain id)
- site_category (category to which website belongs)
- app_id (app id)
- app_domain (app domain)
- app_category (class to which app belongs)
- device_id (id of device from which ad was clicked
- device_ip (ip address of the device)
- device_model (model number of device)
- device_type (mobile/laptop/desktop)
- device_conn_type(internet connection type-wifi/mobile data)
- C1 and C14–C21 (anonymized categorical variables).

3.2 Preprocessing of Data

The class attribute is removed owing to the fact that unsupervised clustering is the technique applied which will learn from the data and classify it into two clusters. The dataset contains features in both integer and string format. In order to carry out the implementation, conversion of strings into integer is done. Pandas package in Python is used to load the csv file, i.e. the raw dataset into memory, identify the columns with string values and convert them into integer values using python standard hash function.

Feature scaling is done with the help of scikit-learn library in python in the case of KMC. It is done in order to ensure fast convergence to the optimal solution and normalize the range of variables.

3.3 Implementation of Machine Learning Algorithms

The following section explains the clustering algorithm used.

K-means Clustering: It is one of the most straightforward learning algorithms available for unsupervised learning [10]. The procedure basically follows two main steps to classify the dataset into k clusters. The number of clusters are fixed apriori. The algorithm starts by defining k cluster centers. Each value in the dataset is then associated to its nearest cluster. This is followed by a recalculation of the k cluster centers. A new binding is done between discrete units of information in the dataset and the new cluster centers and these steps are repeated till all units have been assigned to their respective clusters.

The pseudocode is given as follows (Fig. 2):

X = {x1, x2 …xn} be the collection of observations and u = {u1, u2 …uc} are the cluster centroids. We arbitrarily select 'c' cluster centers (1). Then, the distance

Given: a training set x(1),...,x(m) and feature vectors for each data point $x(i) \in \mathbb{R}^n$ with no labels $y^{(i)}$

(unsupervised learning problem).

Objective: Predict k centroids and a label $c^{(i)}$ for each datapoint

1. Initialize *k* points at random as cluster centers.
2. Assign data point to their closest cluster center according to the Euclidean distance function.

   For every i, set

   $$c(i) := arg\ min(j)||x^{(i)} - \mu j||^2$$

3. Find the new cluster centroid by calculating mean of all data points that belong to the cluster.

   For each j, set $\mu_j := \frac{\sum_{i=1}^{m} 1\{c^{(i)}=j\}x^{(i)}}{\sum_{i=1}^{m} 1\{c^{(i)}=j\}}$

4. Repeat steps 2 and 3 until the same points are assigned to each cluster in consecutive rounds.

Fig. 2 Pseudocode for k-means clustering

between each observation in dataset and center is calculated. Following this, each observation is assigned to that cluster for which the distance between the observation and the center is least from all the available centers (2). This assignment step is followed by a recalculation of the cluster centers (3) by executing step 1 again. If all observations belong to the clusters calculated in the previous iteration, then the algorithm is terminated.

1. Initialize $U=[u_{ij}]$ membership matrix, $U^{(0)}$
2. At k-step: calculate the centers vectors $C^{(k)}=[c_j]$ with $U^{(k)}$

$$c(j) = \frac{\sum_{i=1}^{N} u_{ij}^{m} . x_i}{\sum_{i=1}^{N} u_{ij}^{m}}$$

3. Update $U^{(k)}$, $U^{(k+1)}$

$$u_{ij} = \frac{1}{\sum_{k=1}^{c} \left(\frac{\| x_i - c_j \|}{\| x_i - c_k \|} \right)^{\frac{2}{m-1}}}$$

4. If $|| U^{(k+1)} - U^{(k)} || < \varepsilon$ then STOP; otherwise return to step 2.

Where u_{ij} is the degree of membership of x_i in the cluster j,

x_i is the ith of d-dimensional measured data,

c_j is the d-dimension center of the cluster

ε is a termination criterion between 0 and 1.

Fig. 3 Pseudocode for fuzzy c-means clustering

Fuzzy C-Means Clustering: FCM is also referred to as soft clustering [11]. Each observation can be a part of more than one cluster at the same time. The distance between each cluster center and value in the dataset is evaluated and based on this distance, the degree of membership is assigned to each observation. A higher degree of membership towards a cluster means that the observation is closer to that cluster center than compared to the rest of the clusters. The summation of the degree of membership of each point in dataset equals to one. Data points belong to distinct clusters in hard or non-fuzzy clustering. The degree of membership assigned to each observation plays a pivotal role here. This denotes the extent to which an observation belongs to each cluster. As we move from center to the boundary of the cluster, the degree of belongingness of the observation decreases.

The pseudocode for FCM is as follows (Fig. 3).

Let X = {x1, x2, x3 …, xn} be the collection of discrete units in dataset and V = {c1, c2, c3 …, ck} be the centers. Select 'k' cluster centers and initialize the membership matrix in step (1). Compute the centers 'cj' in step (2). Calculate the membership 'μij' in the membership matrix using step (3). Repeat step (2) and (3) until the smallest value of J is obtained.

4 Results and Analysis

The results were analysed using the accuracy, precision and recall [12] as an efficacy criterion. Accuracy is the measure of the closeness of the predicted observations to the actual value. It is the ratio of rightly predicted inspections to the total inspections made. Higher value of accuracy indicates more true positives and true negatives. Precision refers to the correctness of a model. It is simply the ratio of all the precisely predicted positives to the total number of positives predicted. More the number of true positives implies more precision. Recall is a measure of responsiveness or sensitivity of a machine model. It is the ratio of the correctly predicted positives to the count of values that belong to the class 'yes'. Recall and precision are inversely related. The following Table 1 depicts the performance analysis of FCM and KMC.

As Table 1 shows the accuracy, recall and precision values obtained for the two clusters, i.e. clicks/non-clicks predictions for the users is higher for FCM than for

Table 1 Comparison of results for k means and fuzzy c-means clustering

	k-means		Fuzzy c-means	
	Cluster 0	Cluster 1	Cluster 0	Cluster 1
Accuracy	43.85438544	43.85438544	76.61767177	76.61767177
Precision	15.62280084	81.0428737	17.12049388	83.51293103
Recall	52.05182649	42.16809357	91.03165299	9.345230918

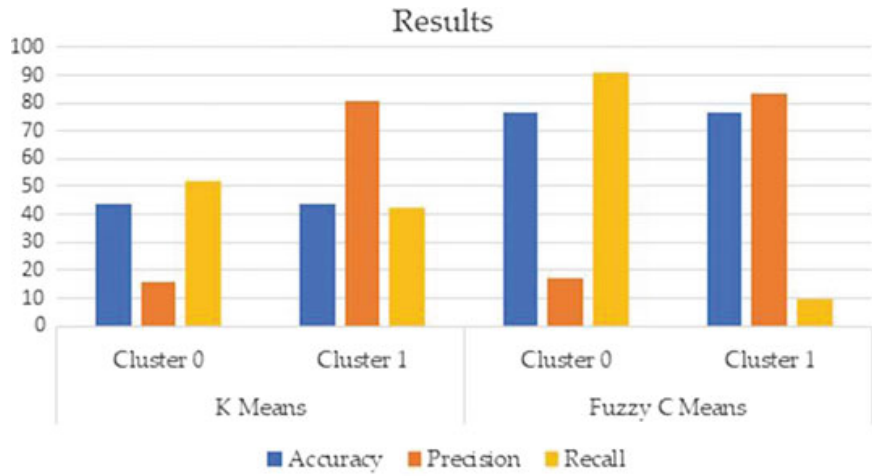


Fig. 4 Results

KMC. This means that the rightly predicted true positives and true negatives (accuracy), the degree of exactness (precision) and degree of completeness (recall) are relatively higher for FCM than for KMC.

Both the algorithms divide the data into two clusters, Cluster 0 and Cluster 1, which in turn represent a division of the dataset into two categories (Fig. 4). These categories are based on click/non-click classification for the ads, i.e. one cluster has all the ads clicked by the user. The other cluster has all ads which were not clicked by him. This gives a clear idea of the type of ads the user is inquisitive about and willing to click on.

After the classification of the ads in the dataset into categories—'Click' and 'Non-Click', we get a fair estimate of the kind of advertisements the user really wants to see. We can thereby predict more of such ads, hence increasing the CTR for the same.

The number of 'clicks' divided by the total number of times the ad is displayed, i.e. 'impressions' gives the CTR.

5 Conclusion

This paper compares the CTR as assessed by KMC and FCM algorithms. The outcome shows that FCM achieves better values of performance metrics than KMC. The results are better because it considers the fact that each data point can lie in more than one cluster and involves complex calculation of membership matrix. Conventional KMC, on the other hand, relies on hard clustering and definitively assigns each data point to the clusters.

The accuracy of the model can further be improved by assessing the shape of the clusters, including other more refined attribute selection and extraction methods that could assist in better modelling of the present system. Attribute selection is

concerned with selecting a subgroup of valid features for model selection whereas attribute extraction means deriving attributes from the already existing ones for subsequent learning. Superior results could be obtained by enhancing fuzziness coefficient. The degree of overlap between clusters is determined by the fuzziness coefficient. Higher value of m means larger overlapping between clusters. There exists a vast scope of application of other soft computing methodologies also, like swarm optimization, etc. that can be applied for determining CTR for other datasets as well.

References

1. Avila Clemenshia, P., Vijaya, M.S.: Click through rate prediction for display advertisement. Int. J. Comput. Appl. **975**, 8887 (2016)
2. Graepel, T., Candela, J., Borchert, T., Herbrich, R.: Web-scale bayesian click-through rate prediction for sponsored search advertising in microsoft's bing search engine. In: IJCA (2010)
3. Hillard, D., Schroedl, S., Manavoglu, E., Raghavan, H., Leggetter, C.: Im-proving ad relevance in sponsored search. In: ACM (2010)
4. Kondakindi, G., Rana, S., Rajkumar, A., Ponnekanti, S.K., Parakh, V.: A logistic regression approach to ad click prediction (2014)
5. Shi, L., Li, B.: Predict the click through rate and average cost per click for key-words using machine learning methodologies. In: IEOM (2016)
6. Wang., C.J., Chen, H.H.: Learning user behaviors for advertisements click predictions. In: ACM (2011)
7. Chakrabarti, D., Agarwal, D., Josifovski, V.: Contextual advertising by combining relevance with click feedback. In: WWW (2008)
8. Cheng, H., Cantú-Paz, E.: Personalized click prediction in sponsored search. In: ACM (2010)
9. Edizel, B., Mantrach, A., Bai, X.: Deep character-level click-through rate prediction for sponsored search. In: Stat.ml (2017)
10. Yadav, J., Sharma, M.: A review of K means algorithm. Int. J. Eng. Trends Technol. **4**(7), 2972–2976 (2013)
11. Yanyun, C., Jianlin, Q., Xiang, G., Jianping, C., Dan, J., Li, C.: Advances in research of fuzzy c means algorithm. In: International Conference on Network Computing and Information Security (2011)
12. Bhatia, M.P.S., Kumar, A.: Information retrieval and machine learning: supporting technologies for web mining research and practice. Webology **5**(2), 5 (2008)

ASK Approach: A Pre-migration Approach for Legacy Application Migration to Cloud

Sanjeev Kumar Yadav, Akhil Khare and Choudhary Kavita

Abstract Legacy application migration is a mammoth task, if migration approach is not well thought at the very start, i.e. pre-migration, and supported by robust planning especially at pre-migration process area. This paper proposes a mathematical pre-migration approach, which will help the enterprise to analyse existing/legacy application based on the application's available information and parameters an enterprise would like to consider for analysis. Proposed pre-migration assessment will help in understanding the legacy application's current state and will help in un-earthing the information with respect candidate application. Proposed pre-migration approach will help to take appropriate well-informed decision, whether to migrate or not to migrate the legacy application. As it is said that application migration is a journey, if kick-started once, needs to reach its destination else it can result into a disaster hence pre-migration is one of the important areas of migration journey.

Keywords Application migration · ASK approach · ASK confidence level (ACL) · LAMP2C · Legacy application · Migration process · Pre-migration

1 Introduction

Technology paradigm is changing at fast pace, so is the customer requirement. In order to serve the fast-changing customer requirement is to keep abreast with ever-changing technology. Important is, if enterprise does not embrace or adopt

S. K. Yadav (✉) · C. Kavita
Jayoti Vidyapeeth Women's University, Jaipur, India
e-mail: sanjeevyadav@yahoo.com

C. Kavita
e-mail: kavita.yogen@gmail.com

A. Khare
CSED MVSR Engineering College, Hyderabad, India
e-mail: khareakhil@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_2

new or up-coming technology then enterprise may face challenges such as unavailability or shortage of skilled manpower as they would have moved to other latest technology, which will impact software and hardware maintenance cost (complex patching and modifications/enhancement) because scarcity of people in market with the required technology will increase the manpower cost and at the same time even the hardware will also become costly. All this may put application scalability, business continuity, customers' experience and enterprise image [1] in jeopardize and may result in inability to meet current business needs.

Each of the challenges identified may have direct or indirect impact on one or more factors. At times these identified factors seem to be interrelated, refer Fig. 2. There may be other factors, which may be enterprise or legacy application specific.

Enterprises may be ardent and ready to adopt technology such as cloud computing, which delivers resource as virtualize service [2] and offers so many other benefits. However, the biggest challenge for an organization is that enterprise can't discard the investment made in the past in terms of existing business critical and in-use applications including its infrastructure. These business-critical applications/systems are referred as an asset. These assets are built or might have evolved over a span of 10–20 years; and have become core to business process, might have become complex and large enough to be understood by a person.

Cloud computing offers multiple benefits which any enterprise will be interested in such as lower operating cost, no up-front investment and ease of maintenance. Fact is that cloud reduces capital expenditure by leveraging cloud business model of pay-as-you use, which turns capital expenditure to operation expenditure.

As exiting applications cannot be discarded, so to embrace the new technology enterprise has below-mentioned three options for lasting benefits/goal, refer Fig. 1:

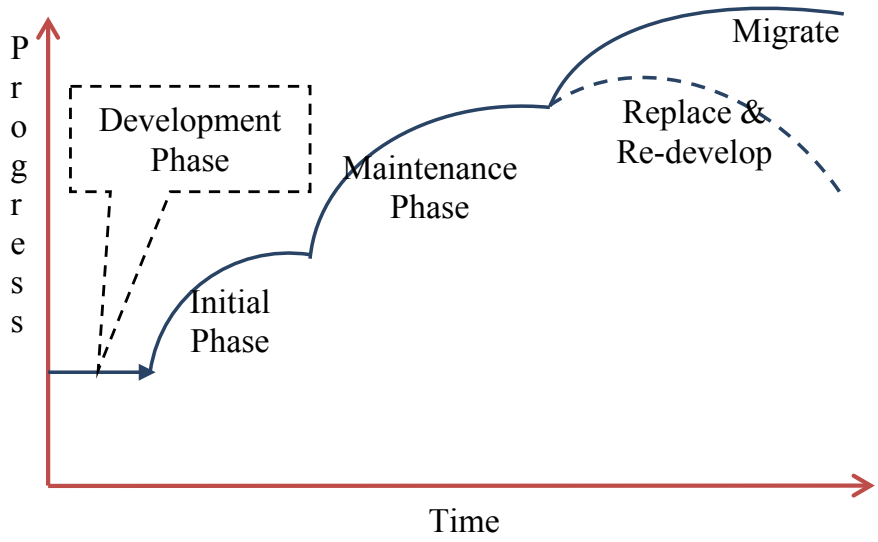


Fig. 1 Software life cycle phases and modernisation options©

i. Application Migration—Migrate legacy application migration to new platform.
ii. Application Replacement—Replace legacy application with off-the-shelf, available in market, application.
iii. Application Redevelopment—Redevelop the existing application all over again.

Two options, Option ii, i.e. Application replacement or option iii, i.e. Application redevelopment, does not seem to be feasible options because the existing business-critical application might have got developed with lot of effort, time and cost. Putting similar effort, cost and time all over again may not be feasible, so feasible option available with enterprise is option i, i.e. Application migration, which will help to migrate current application to new technology/ platform and embrace new technology, i.e. Cloud at the same time, will keep the business-critical application intact.

Paper is organized in six sections. Section 2 provides the literature review, which provides background of related work. Section 3 defines the problem. Section 4 provides the detail about the proposed framework. Section 5 details the ASK approach. Finally, the paper concludes with Sect. 6, which presents the concluding comments and future scope.

2 Literature Review

Cloud computing acceptance has increased over a period of time because of the benefits it brings in the area of operational and efficiency. Migration frameworks are proposed and are compared in the selected literature but none provides an end-to-end migration framework, especially in the pre-migration stage of the migration. The focus of literature review is to bring out any available pre-migration method/approach.

Summary of literature studied and identified research gap in context of chosen topic of the paper are detailed below:

Shoaib et al. [3] highlighted that due to technology advancement enterprises need to migrate from one platform to other, so existed framework was examined and reviewed. Based on this, the author is of view that multiple methods exists for migration and at the same time there are multiple risks involved during migration all these risks can be taken care if migration is performed properly. In this paper, the author describes different methods and frameworks which provides guideline for developers to enhance software migration process. As per the author, the focus area for migration is framework along with process and activities. The author concludes that there is research gap and need to be explored further.

Musale et al. [4] provided basic knowledge about the migration concept and initial factors required to migrate to cloud. The paper also highlighted that could computing is an emerging technology and there is substantial scope of operational

and efficiency improvement. If IT application is migration to cloud or some of the processes are migration to cloud environment, off premises. The author puts forward the cloud migration concept and discusses the essential points for migration and further discusses the after migration essential. The author is of view that migration to cloud is an art than it is a science.

Gholami et al. [5] highlighted that research around migrating legacy application to cloud has increased because of the adoption of cloud computing as an outsourcing strategy has grown in recent past. The author is of view that there is no integrated overarching cloud migration exists in current literature. In order to fulfil this gap, the author tries to propose cloud platform-independent conceptual view and reusable migration process, which has phases, activities and task. Validity of the proposed metamodel is demonstrated by analysing three existing cloud migration process models. Model can be extended by adding new constructs and can be customized for other migration paradigm.

Sabiri et al. [6] , the authors said that enterprises adopting cloud is increasing because of the significant advantage cloud computing brings in; however, the migration (move) to cloud has challenges. The authors studied and compared the existing cloud migration methods to understand the strengths and weaknesses of each model based on Model-Driven Engineering (MDE) approach. The author proposed and explained a cloud migration method based on Architecture Driven Modernization (ADM), which focuses on architectural modes instead of code artefacts.

Gouda et al. [7] highlighted that cloud migration can be of two types; either from on-site premises to the cloud or moving them from one cloud environment to another, where the information can be accessed over the Internet. The authors also highlighted that cloud migration is broadly classified into two categories, i.e. Big Bang Migration and Trickle Migration. The authors have suggested multiple cloud migration patterns or categories as their requirement. The authors also suggested migration to cloud-step wise approach or procedure which will help the user to successful and fruitful migration of its IT resources to cloud.

Shawky [2] highlighted that in order to utilize cloud's capability, the system should be migrated to cloud. The author also highlights that application migration to cloud leads multiple technical challenges and these challenges cannot be ignored. According to the author, the main challenge is mapping application to cloud services as multiple aspects need to be considered, which includes other system component dependencies.

Rashmi and Sahoo [8], 'A five-phased approach for cloud migration'; highlighted that cloud computing is the rapidly growing segment of IT. Organization is moving to cloud due to its long-term benefits but they are unsure about the migration process. The authors proposed a five-phased waterfall-based migration model to cloud having feedback path to earlier phase of the process but have not dwelled into the details of each phase of the proposed model.

Pahl et al. [9] mentioned in this paper that though cloud computing has gained the attention, adoption of the technology is yet to gain acceptance as how to migrate to cloud is unanswered for many. Point in case is highlighted with the help of three provider-driven case studies and tries to extract common layer-specific migration

process. Highlights that no common procedures and tools are exiting, however, establish migration core elements such as activities and steps.

Stavru et al. [10] acknowledged the fact that enterprises are unable to take full strategic and operational advantage as they are not able to migrate legacy application to cloud due to non-existence of mature process. The authors identified and extracted challenges with the help of systematic literature review. Using expert judgment, evaluates how different agile techniques, taken from Scrum and Extreme Programming (XP), could address or overcome the identified challenges in software modernization projects in specific context.

3 Problem Specification

Challenge can be in terms of skilled resource's scarcity, software maintenance cost, lack of software scalability, business continuousness, customers experience, reduced efficiencies, knowledge retention, enterprise image [1] and inability to meet current business needs. Each of the challenges identified may have direct or indirect impact on one or more factors, so identified impacted factors seem to be interrelated. Tried categorizing these challenges into the below category; refer Fig. 2. There may be other factors, which may at times be enterprise specific or legacy application specific.

- Customer
  - (a) Application inefficient in catering to business requirement will impact the enterprise interest.
  - (b) Applications may become inefficient in catering to any area of customer requirement (customer can be internal and/or external) over time; in terms of functionality/feature(s) and better service. Inefficient application will provide an edge to competition.
- Knowledge
  - (a) Lack of in-depth knowledge in terms of understanding the existing system and its related process will or may have direct impact on the system maintenance, i.e. developing new functionality and enhancements.

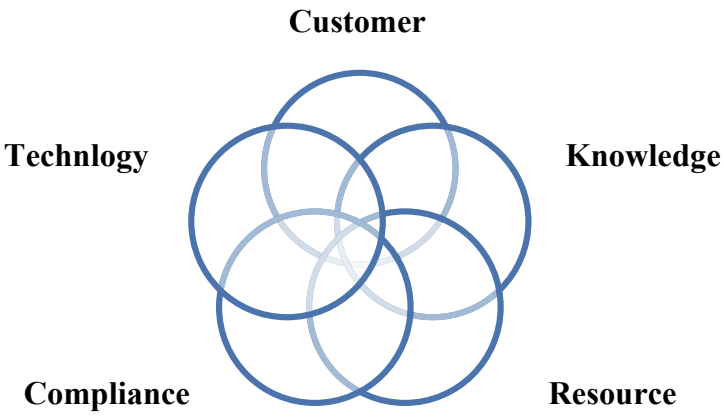


Fig. 2 System maintenance and enhancement challenges

(b) Due to non-availability of knowledgeable resources will have difficulty in getting best practices shared with respect to these legacy applications, i.e. knowledge sharing. It will increase the risk of knowledge concentration amount small set of people.

- Resources

(a) Reducing resource pool (attrition) or resource pool desire to move onto market relevant technology will have direct impact on resource availability for the maintenance/enhancement of legacy application.
(b) As legacy application knowledge gets concentrated into few resources, it will lead to increase in people retaining cost; related aspect will be increase cost of gaining resource in desired technology.
(c) Contracted resource pool will have risk of high dependency, in terms of understanding the existing system and its related process on to a very small set of people in an organization.

- Technology

(a) Enterprise may have multiple technologies in use for existing business-critical applications, so maintaining them will have high cost because the legacy application might have become unsupported due to non-existence of support from the product vendor as they might have stopped supporting the in-use version or vendor might release new version, i.e. vendor moved ahead in technology.

- Compliance

(a) Legacy applications may become non-regulatory compliance over a period of time if no enhancement is made due to multiple reasons mentioned above.
(b) Non-compliance to regulatory norms may have bigger impact than adopting new technology and moving legacy application to the technology.

Challenges highlighted above are few of the challenge, many other challenges will come which will be application or enterprise-specific challenges. These challenges motivate to come up with a framework, which will help the enterprise to migrate legacy application to cloud in an efficient manner.

4 Proposed Framework

Legacy application migration is a journey which requires a careful detailed planning irrespective of the size of the application identified for migration because no one-size-fit-approach will work in application migration. One of the biggest challenges an enterprise faces is the uncertainty of where to begin the migration process. If enterprises do not carefully plan, execute and monitor the transformation using

established processes [11] then migration can be fraught with pitfalls. Before we proceed further, it is essential to understand the reasons or trigger points for migration. At high level there are only two factors which trigger migration and these factors are either: Business Driven, Technology Driven or both. Whichever may the reason for migration, i.e. business and/or technology; one thing which will happen first is current portfolio analysis.

Cloud computing concepts and approach itself is very different from the traditional software concepts and approach. In fact, moving to cloud is a serious step to take in a holistic manner [1], which required a careful elaborate planning for successful migration because cloud computing brings changes such as economic, legal, privacy/security, [1] etc. especially the pre-migration area of the migration process.

Cloud covers all application layers, i.e. application (SaaS), platform (PaaS) and infrastructure (IaaS), which makes legacy application migration not a straightforward task as it looks like. Rashmi et al. [8], especially when migrating a legacy application, which is tightly coupled [12] as compared to cloud application.

In order to address this issue, a framework is proposed called '**L**egacy **A**pplication **M**igration **P**rocess Framework to **C**loud'; referred as LAMP2C, Fig. 3. LAMP2C migration framework acts as a guiding principle for migrating application to cloud and process to be followed while migrating legacy application. Scope of the proposed LAMP2C framework covers entire legacy application as shown in Fig. 3 and pre-migration is highlighted as ASK Approach is part and limited to the pre-migration area of the LAMP2C framework.

Framework helps to plan the migration in an effective and efficient manner and it also helps in speeding up migration execution process with clear and hazel free migration path.

LAMP2C, address pre-migration, migration and post-migration area of migration with an umbrella of governance area across three cloud layers. Framework is divided into area and area is further sub-divided into phases. The phases will have activities under it. LAMP2C—high-level framework is shown in Fig. 4.

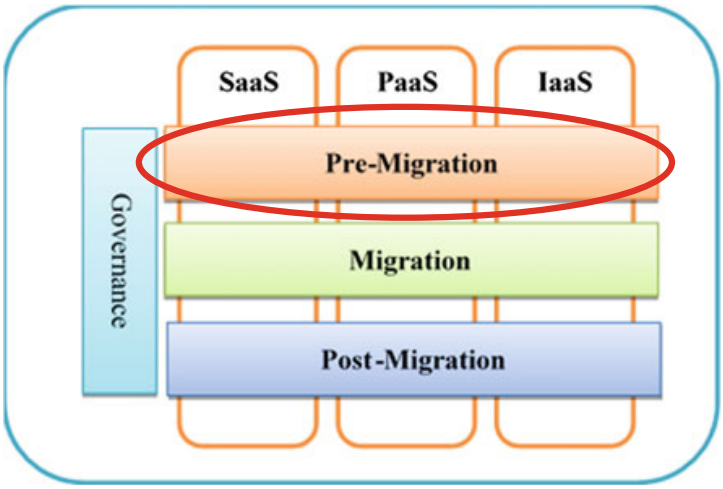


Fig. 3 High-level framework

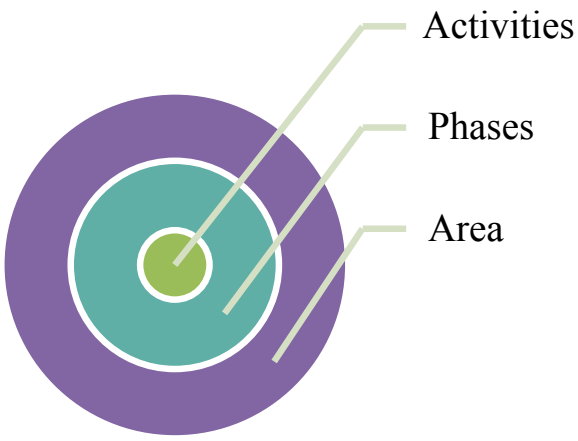


Fig. 4 LAMP2C constituents

LAMP2C is a flexible framework, which can be tweaked as per the enterprise's requirement as proposed framework does not force to follow the process end-to-end rather it suggests to go phased manner.

One of the important parts of LAMP2C is the pre-migration area, which helps in evaluating existing portfolio or application which in turn helps an enterprise in understanding the complexity, challenges and efforts that will be required for application migration. It will also address the uncertainty of 'where to begin' and other than whether to migrate or not to migrate to the cloud.

LAMP2C, Figs. 1 and 3 is copyright of the authors: Registration Number L-60550/2014.

5 Ask Approach

It is assumed that at high level there are only two factors which can trigger migration and they are either technical or non-technical (business) parameter, so assessment and its parameters are also clubbed into these two categories only. Another assumption is that a) technical assessment should consider all the three cloud's service model, i.e. IaaS, SaaS and PaaS; however, enterprise is free to choose one or more for assessment and their respective parameters. Parameters which cut across three models should be grouped into one as 'Common Parameters', as shown in Fig. 5.

Common parameters are to avoid the consideration of same parameter(s) across three layers of service model, however it will depend upon individual, how they would like to perform the pre-migration analysis. b) Non-technical, i.e. business parameter are those parameters, which are not technical in nature should be considered as one group or as suitable.

ASK Approach—a pre-migration assessment methodology is proposed to perform portfolio or application analysis under pre-migration are of LAMP2C framework. ASK Approach is an important constituent of the LAMP2C framework and its scope is limited to the pre-migration area of the LAMP2C framework as it

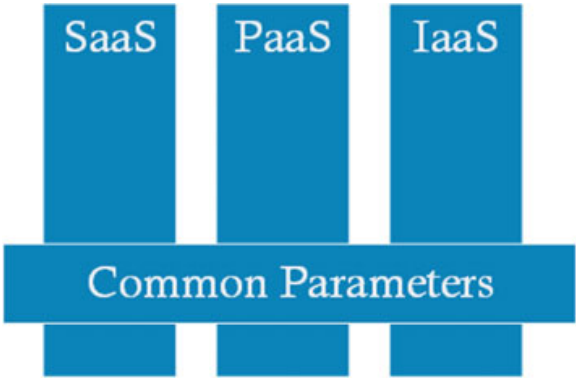


Fig. 5 Parameter identification area

Fig. 6 Pre-migration approach

provides inputs and helps in taking GO/No-Go decision with respect to legacy application migration to cloud. Proposed pre-migration assessment methodology is parameter based mathematical approach, which will be appropriate for any application identified for migration and will help enterprise to view and analyse the enterprises IT application portfolio in details based on parameters which are important for organization and application so assessment is carried out on any set of identified parameters under the mentioned category or both, i.e. technical and non-technical (business) aspect.

Pre-migration approach can be divided into five steps, as shown in Fig. 6 and detailed in subsequent paras:

As approach is parameter based, as a first step, all parameters and its sub-parameters needs to be identified. Each identified and considered parameter needs to be assigned the importance. Importance is rated between the range of 0.1, being the lowest to 5.0, being the highest, based on the parameter's impact, as tabled below (Table 1):

Table 1 Impact rating range

Importance	Impact rating range
Very high	4.1–5.0
High	3.1–4.0
Neutral	2.1–3.0
Low	1.1–2.0
Very low	0.1–1.0

Assuming that there are n number of parameters namely C_1, C_2, C_3 and so on till C_n. Based on parameters' importance an impact rating between the range of 0.1–5.0 will be assigned for each parameter. Adding all these integrals and dividing them by the total number of parameters (in this case n) considered for analysis will give us the mean value of the data, which is being referred as '**ASK Point**' (ASKP), i.e. μ_c.

$$\text{ASKP} = \mu_c = \frac{\int C_1 + \int C_2 + \int C_3 + \int C_4 \ldots + \int C_n}{\text{n}}$$

$\text{ASKP} = \mu_c = \textit{Parameters'Mean}$
$n = \textit{Number of parameters}$
$i = \textit{Importance of parameter}$
In reduced form the above formula can be written as follows:

$$ASKP = \mu_c = \sum_{1}^{n} \frac{\int C_n}{n}$$

Considering the same data calculate the standard, which will be referred as **ASK Confidence Level (ACL)**. ACL will be calculated by adding the squares of the values found from subtracting the ASKP from each parametric value. This summed value is divided by the total number of parameters. This will give us the standard deviation σ_c, i.e. ACL.

$$\text{ACL} = \sigma_c = \frac{\left(\int C_1 - \mu_c\right)^2 + \left(\int C_2 - \mu_c\right)^2 + \left(\int C_3 - \mu_c\right)^2 + \cdots \left(\int C_n - \mu_c\right)^2}{n}$$

$\text{ACL} = \sigma_c = \textit{Parameters' Standard Deviation}$
$\text{ASKP} = \mu_c = \textit{Parameters'Mean}$
$n = \textit{Number of parameters}$

In reduced form the above formula can be rewritten as follows:

$$ACL = \sigma_c = \sum_{1}^{n} \frac{\left(\int C_n - \mu_c\right)^2}{n}$$

In order to generate boundary for the identified parameter, which is going to be referred as **ASK Belt** (ASKB). ASKP and ACL will help to build ASK Belt. ASKB will have two components a) ASKB Right, i.e. ASKB_R and b) ASKB Left, i.e. ASKB_L. ASKB_R and ASKB_L can be calculated as shown below:

$$\text{ASKB}_\text{R} = \text{ASKP} + \text{ACL}$$
$$\text{ASKB}_\text{L} = \text{ASKP} - \text{ACL}$$

Once, ASKP, ACL and ASKB ($ASKB_R$ and $ASKB_L$); bell curve can be calculated using the below formula (Gaussian equation)

$$f(x) = \frac{1}{\sqrt{2\pi\sigma_c^2}} e^{\frac{-(x-\mu_c)^2}{2\sigma_c}}$$

The following are the value for the above formula:

π = 3.14
e = 2.71 and
x = Parameter Impact
μ_c = ASKP
σ_c = ACL

As fifth and the last step; calculate various other values required in the Gaussian function. The parameters have been arranged in the ascending order of the impact rating for facilitating the plotting process. Plotting the above $f(x)$ values versus x values, we get the following bell curve (as defined by the Gaussian function), refer Fig. 7.

Points which falls between $ASKB_R$ and $ASKB_L$ can be safely considered for legacy application migration, however, points which fall outside the belt need to be re-looked and considered with caution.

These inputs will provide enough data points to conclude as to Go ahead with the legacy application migration or still there are points/parameters, which need to be re-looked at. And will provide inputs to come up with migration plan.

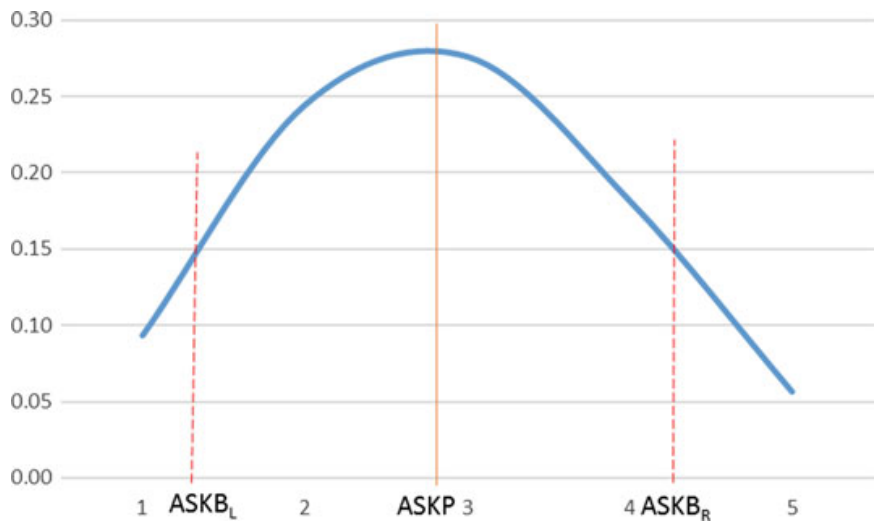


Fig. 7 Bell curve

6 Conclusion

Proposed mathematical pre-migration methodology called 'ASK Approach' under the LAMP2C migration framework. Important is, impact rating between the range of 0.1–5.0 is assigned to each identified and considered parameters. 0.1 being the lowest and 5.0 being the highest. Assigning an impact rating to each parameter will eliminate any guesswork among any of the person associated with respective application migration. Impact rating will help to drive ASKP, i.e. μ_c, which in turn will help to drive ACL, i.e. σ_c. ASKP and ACL will help to drive ASKB_R and ASKB_L which in turn will help to calculate $f(x)$ for each considered parameter. Bell curve can be plotted to identify the parameters, which helps to provide a view about, which legacy application parameters should be considered with confidence and which parameter needs more attention. It requires ultimate help to take informed decision of GO or NO-GO with respect to application migration. With the help of this pre-migration approach enterprise can also make phase or staggered legacy migration approach. Phased manner approach provides the required feel confident factor and return on their investment can be seen as migration project progress. Big bang migration approach may not provide the required feel good factor and always have a project failure risk.

Proposed ASK Approach is focused on parameter which falls with the ASK Belt, but does not provide any inputs as to what needs to be done with parameters, which are outside the ASK Belt, so ASK Approach can be further enhanced to come up with mathematical formula or otherwise, as to how to deal with the parameters, which falls outside the ASK Belt, i.e. ASKB_R and ASKB_L.

References

1. Khajeh-Hosseini, A., Sommerville, I., Sriram, I.: Research challenges for enterprise cloud computing (2010)
2. Shawky, D.M.: A cost-effective approach for hybrid migration to the cloud. Int. J. Comput. Inf. Technol. **2**(1), 57 (2013)
3. Shoaib, M., Ishaq, A., Ahmad, M.A., Talib, S., Mustafa, G., Ahmed, A.: Software migration frameworks for software system solutions: a systematic literature review. Int. J. Adv. Comput. Sci. Appl. (IJACSA) **8**(11), 192–204 (2017)
4. Musale, A.D., Khot, P.G.: A roadmap towards cloud migration. IOSR J. Comput. Eng. (IOSR-JCE). e-ISSN: 2278-0661, p-ISSN: 2278–8727
5. Gholami, M.F., Low, G., Beydoun, G.: Conceptualising cloud migration process. In: Twenty-Fourth European Conference on Information Systems (ECIS), İstanbul, Turkey (2016)
6. Sabiri, K., Benabbou, F., Hain, M., Akodadi, K.: A survey of cloud migration methods: a comparison and proposition. Int. J. Adv. Comput. Sci. Appl. (IJACSA) **7**(5), 598–604 (2016)
7. Gouda, K.C., Dwivedi, D., Patro, A., Bhat, N.: Migration management in cloud computing. Int. J. Eng. Trends Technol. (IJETT) **12**(9), 466–472 (2014)
8. Rashmi, M.S., Sahoo, G.: A five-phased approach for cloud migration. Int. J. Emerg. Technol. Adv. Eng. (IJETAE) **2**(4), 286–291 (2012)

9. Pahl, C., Xiong, H., Walshe, R.: A comparison of on-premise to cloud migration approaches —a tale of four cloud migration processes. In: European Conference on Service-Oriented and Cloud Computing (2013)
10. Stavru, S., Krasteva, I., Ilieva, S.: Challenges for migrating to the service cloud paradigm: an agile perspective (2012)
11. PMI White Paper: Cloud computing: the new strategic weapon (2012)
12. Bisbal, J., Lawless, D., Wu, B., Grimson, J., Wade, V., Richardson, R., O'Sullivan, D.: An overview of legacy information system migration. In: Asia-Pacific Software Engineering Conference (1997)
13. Yadav, S.K., Khare, A.: Legacy applications migration process to cloud—an approach framework. Int. J. Comput. Technol. Electron. Eng. (IJCTEE) **4**(2), 11–14 (2014)

A Fuzzy Logic Based Cardiovascular Disease Risk Level Prediction System in Correlation to Diabetes and Smoking

Kanak Saxena and Umesh Banodha

Abstract The cardiovascular disease (CVD) is one of the major causes of death among the people having diabetes in addition to smoking habits. It will create tribulations for every organ of the human body. Smoking becomes fashion among the youth from their childhood which results in premature death. The intention of this paper is to explain the impact of diabetes and smoking along with high BP, high pulse rate, angina affect, and family history on the CVD risk level. The concept used is based on the knowledge-based system. We have proposed a fuzzy-logic-based prediction system to evaluate the CVD risk among the people having diabetes with smoking habits. The aim is to facilitate the experts to provide the medication as well as counsel the smokers well in advance. This will not merely save the individual but also an immense relief to concern. The data set is used from UCI (Machine Learning Repository). Most of the researchers worked on diabetes or smoking impact on CVD separately, but the proposed system demonstrates how drastically it will affect ones' health condition.

Keywords Approximation function · Cardiovascular disease · Decision-making · Diabetes · Fuzzy logic · Smoking

1 Introduction

Cardiovascular Diseases (CVD) are very widespread disease that becomes the ground of death and morbidity among the humans [3, 5, 19]. The data from the national heart association (2012) showed that 65% of people with diabetes die from some sort of heart disease or stroke. The Framingham study was the earliest sub-

K. Saxena (✉) · U. Banodha
Department of Computer Application, Samrat Ashok Technological Institute, Vidisha, India
e-mail: ks.pub.2011@gmail.com

U. Banodha
e-mail: ub.pub.2011@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_3

stantiation to demonstrate that people with diabetes are more (near about 2–4 times) susceptible to CVD disease than those who did not have diabetes [12]. It is very clear that diabetes if not controlled then it becomes the unending cause of not only the CVD but many more diseases [16]. As the WHO the percentage of the affected group is of 30 million in the year 1985, which increased to 135 million in the year 1995, and in the year 2005 it became 217 million. The prediction is near about 366 million people will suffer from diabetes in 2030, either they belong to developing or developed countries. On the other hand, smoking is the most imperative avertable cause of death which also accelerates the chances of occurrence of CVD among people, i.e., enhances the risk level [4]. It is also stated that smoking is the second leading cause after the high BP for CVD mortality. The WHO reported that more than one billion people smoke and the frequency rising surprisingly [14]. According to [8] the risk level of CVD estimated 15- year long period ahead for a smoker in comparison to a nonsmoker. Nowadays, the health care domain is one of the key domain where research is going on in the area of knowledge base to predict the impending risk in a patient and future of a healthy person with bad habits. The importance of the knowledge base is that it analyzes the data using various techniques. The experts took advantage of using the computerized diagnosis system where the information draws together in direct as well as indirect form. Earlier the work is done manually, though the decisions were taken by the help of the computer-based decision support system. With the improvement in the technology, the manually collected data problems have solved and now the system acquires the data which are directly or indirectly related to the physical condition. Today, there are many health care apps which assist the users as well as the experts for quick evaluation, thus avoiding uninvited events in ones' life. The evaluation varies on the input parameters and the techniques applied for the evaluation. Currently, the focus is on the flexibility of the conditions with intelligent evaluation. For the proposed system the highly suitable basis is the fuzzy set theory and the logic. The concept helps us in the interpretation of the facts with its findings and enhances the system with the cross diagnosis and estimation of the risk level.

In the proposed prediction system the main focus is on the CVD in relation to diabetes and smoking (urine nicotine). For testing purpose, the data set from UCI is used with major eight attributes as shown in Fig. 1. Among these, the major inputs are eight (inclusion of smoking) with adding together and one is used for the output purpose.

2 Literature Survey

Phuong et al. [13] illustrated the use of fuzzy logic and its application in the field of medicine. Oad and DeZhi [10] described the fuzzy-rule-based system for prediction of heart disease. Neshat et al. [11] proposed a fuzzy expert for liver disorder. Anooj [1] developed the weighted fuzzy clinical decision support system for the heart disease risk level prediction. Alessandra Saldanha deMattosMatheus et al. discussed

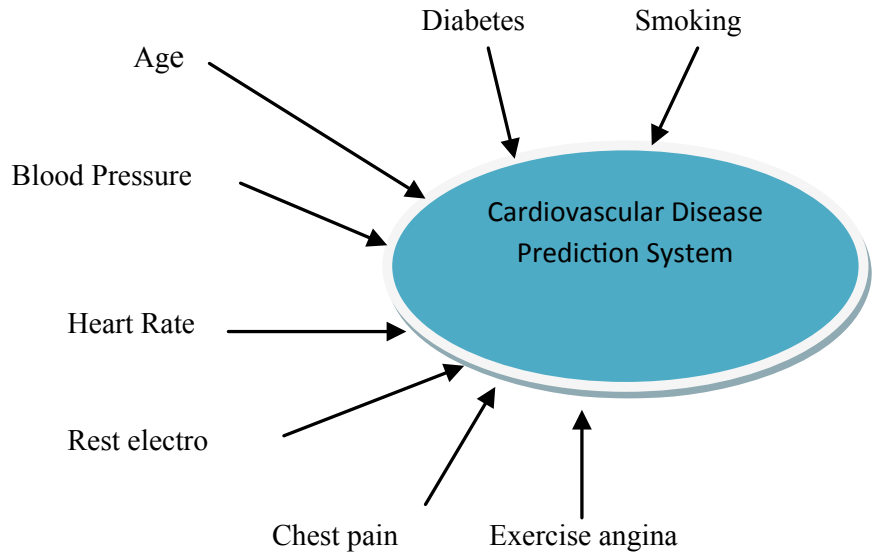


Fig. 1 Inputs of the cardiovascular disease risk level prediction system

the impact of the diabetes on the cardiovascular disease with obesity property in the patient. Rajeswari and Vaithiyanathan [15] developed heart disease diagnosis system using fuzzy logic and genetic algorithm. Anbe et al. [9] proposed the fuzzy set diagnosis system for the valvular heart disease. Kumar and Kaur [17] exemplified the fuzzy rules based on the six parameters to predict the heart disease. Phongsuphap and Pongsupap [18] evaluated the health care policies to assist the authorities on the basis of the survey data which consists of qualitative as well as qualitative. The authors surveyed on the epidemiology of the use of tobacco smoking in the United States. They also emphasized the study of the characteristics of the nicotine and the clinical applications of medication [6]. There is the correlation with the smoking and the cardiovascular disease as it may be one of the causes of altered lipid metabolism, increased demand for myocardial oxygen and blood and many more [20, 21]. In [2, 7], the group of the smokers was studied with the impact of the serum cotinine levels in smokers as well as in nonsmokers.

3 Cardiovascular Disease Prediction System

The proposed prediction system is based on the fuzzy logic technique which may feasibly be used to predict the risk level of the CVD in the patients with diabetes and smoking habits. The preference of the fuzzy techniques is through on the uniqueness of self-explanatory assessment making procedure. In fuzzy logic, its effectiveness depends on the formation of the set of acts. The biased predicament can be unconcerned if the set of acts be capable to form automatically without the intervention of the experts. According to Table 1 the stages and steps of fuzzy logic are illustrated, which commence with the preprocessing of the data followed by the

Table 1 Stages and Steps of Fuzzy Logic

Stages	Steps
Preprocessing	Mining of the parameters
Fuzzification	Define the linguistic variables with their ranges
	Construction of the membership functions for input as well as for output variables
	Conversion of crisp data into the fuzzy data sets with the help of membership functions
Fuzzy rule base	Construction of knowledge base rules
Defuzzification	Mapping to quantifiable risk
Decision-making	Selection of the choice on the basis of the relevant information available

fuzzification and the rules generation. Thus the prediction system instigates with the symptoms performance-based analysis. Further, it will also depict the risk level in the midst of investigation and lend a hand to the experts to take the suitable decision at suitable point in time.

Figure 2 demonstrates the steps used in the CVD prediction system. The preprocessing of the system that consists of patient clinical data pattern to facilitate the input into the system by converging the data into accessible form, by performing the data mining to find the missing values (if any), convert it into the permissible range and significant task is the selection of the attributes which directly or indirectly fatally affect the prediction system. As it's a nontrivial fact that the data

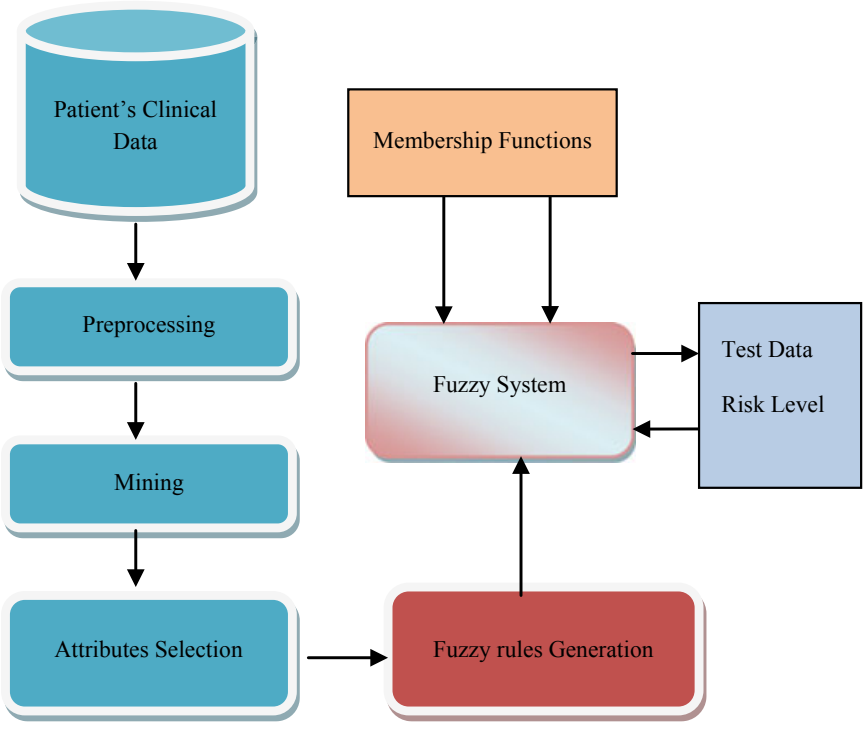


Fig. 2 Proposed fuzzy-based prediction system

which are not normalized cannot envisage the system procedure. The next step is to classify the data set in two categories, either suffering from the CVD or not.

The behavior of prediction system is uncertain. The fuzzy system is more appropriate to handle such type of formless medical events. The fuzzy system consists of the fuzzifier, fuzzy rules, fuzzy inference engine, and defuzzifier. In the fuzzification procedure, the crisp data set is to converted into a fuzzy value and the fuzzifier is the mapping from the data set to the fuzzy linguistic data value—a realistic approximation which is attained by the membership function. The fuzzy rule is to get fractioned into two parts as (1) condition and (2) conclusion. The inference engine interprets the rules logic with the appropriate reasoning, thus results in the form of the output—risk factor. Finally, the defuzzification is the reverse procedure of the fuzzification, where the mapping is to done from fuzzy value to the crisp data value. The system works with the identification of the input and output variables. Member function is used to characterize the fuzzy system and the suitable and succinct way is to characterize the member function mathematically. The significance of the mathematical formula is that it establishes contribution of the input variable to fuzzy set. The output of the prediction system has to be defuzzified, that is it has to be converted into a precise quantity, which can be the logical union of two or more fuzzy membership functions defined on the cardiovascular affect world.

The following process is to fuzzify all inputs and outputs and bring close to the contribution to which these inputs and outputs fit into each of the appropriate fuzzy sets. In the prediction system, the subsequent process is to discover the weighted fuzzy rules which have to engender from the various attributes resolution weighted rules. The method used for the automatic engenders the set of acts that are based on the structure of the state of fuzzy modals. The decision rules were attained from IF and THEN fractions of the specific class. These were acquired by converting into linguistic variables on the basis of the fuzzy membership functions. The first attribute used is age which is classified into four fuzzy sets (minor, young, adult, and old). The triangular as well as trapezoidal membership functions are used for the fuzzy sets. The same place into practice for blood pressure, chest pain, heart pain, blood sugar, rest electro, affected, and urine nicotine. In this, the fuzzy sets were separated into three or four levels that as follows.

Age (separated into 4 levels)			
Minor	Young	Adult	Old
$\mu_{minor}(x) = 1$, $x \in [1,15]$ $\mu_{minor}(x) = (18 - x)/(18 - 15)$, $x \in [15,18]$ $\mu_{minor}[x] = 0$, otherwise	$\mu_{young}(x) = (x - 15)/(24 - 15)$, $x \in [15,24]$ $\mu_{young}(x) = (42 - x)/(42 - 24)$, $x \in [24,42]$ $\mu_{young}(x) = 0$, otherwise	$\mu_{adult}(x) = (x - 33)/(40 - 33)$ $x \in [33,40]$ $\mu_{adult}(x) = (61 - x)/(61 - 40)$, $x \in [40,61]$ $\mu_{adult}(x) = 0$, otherwise	$\mu_{old}(x) = (x - 54)/(54 - 40)$, $x \in [40,54]$ $\mu_{old}(x) = (75 - x)/(75 - 54)$, $x \in [54,75]$ $\mu_{old}(x) = 0$, otherwise

Chest pain (separated into 3 levels)		
Alert	Risk	High risk
$\mu_{alert}(x) = (x - 1)/(2 - 1)$, $x \in [1, 2]$ $\mu_{alert}(x) = 0$, otherwise	$\mu_{Risk}(x) = (x - 2)/(3.5 - 2)$, $x \in [2,3.5]$ $\mu_{Risk}(x) = 0$, otherwise	$\mu_{HighRisk}(x) = (x - 3.5)/(4 - 3.5)$, $x \in [3.5,4]$ $\mu_{HighRisk}(x) = 0$, otherwise

Blood pressure (separated into 5 levels)				
Low	Normal	Alert	High	Very high
$\mu_{low}(x) = 1$, $x \in [1,100]$ $\mu_{low}(x) = (120 - x)/(120 - 99)$, $x \in [99,120]$ $\mu_{low}[x] = 0$, otherwise	$\mu_{normal}(x) = (x - 110)/(121 - 110)$, $x \in [110,121]$ $\mu_{normal}(x) = (132 - x)/(132 - 121)$, $x \in [121,132]$ $\mu_{young}(x) = 0$, otherwise	$\mu_{alert}(x) = (x - 125)/(143 - 125)$, $x \in [125,143]$ $\mu_{alert}(x) = (166 - x)/(166 - 143)$, $x \in [166\text{-}143]$ $\mu_{alert}(x) = 0$, otherwise	$\mu_{high}(x) = (x - 158)/(184 - 158)$, $x \in [125,143]$ $\mu_{high}(x) = (207 - x)/(207 - 184)$, $x \in [184,207]$ $\mu_{high}(x) = 0$, otherwise	$\mu_{very\ high}(x) = 0$ $x \in [1,157]$ $\mu_{very\ high}(x) = (x - 157)/(186 - 157)$, $x \in [157,186]$ $\mu_{very\ high}(x) = 1$, $x \in [207,350]$

Blood sugar (separated into 5 levels)				
Low	Normal	Alert	High	Very high
$\mu_{low}(x) = 1$, $x \in [1,65]$ $\mu_{low}(x) = (80 - x)/(80 - 65)$, $x \in [65,80]$ $\mu_{low}[x] = 0$, otherwise	$\mu_{normal}(x) = (x - 71)/(127 - 71)$, $x \in [71,127]$ $\mu_{normal}(x) = (139 - x)/(139 - 127)$, $x \in [127,139]$ $\mu_{normal}(x) = 0$, otherwise	$\mu_{alert}(x) = (x - 127)/(155 - 127)$, $x \in [127,155]$ $\mu_{alert}(x) = (167 - x)/(167 - 155)$, $x \in [167\text{-}155]$ $\mu_{alert}(x) = 0$, otherwise	$\mu_{high}(x) = (x - 148)/(155 - 148)$, $x \in [148,155]$ $\mu_{high}(x) = (250 - x)/(250 - 199)$, $x \in [199,250]$ $\mu_{high}(x) = 0$, otherwise	$\mu_{very\ high}(x) = 0$ $x \in [1,210]$ $\mu_{very\ high}(x) = (x - 210)/(250 - 210)$, $x \in [210,250]$ $\mu_{very\ high}(x) = 1$, $x \in [250,500]$

Heart pain (separated into 4 levels)			
Low	Normal	High	Very high
$\mu_{low}(x) = 1$, $x \in [1,55]$ $\mu_{low}(x) = (71 - x)/(71 - 55)$, $x \in [55,71]$ $\mu_{low}[x] = 0$, otherwise	$\mu_{normal}(x) = (x - 66)/(73 - 66)$, $x \in [66,73]$ $\mu_{normal}(x) = (98 - x)/(98 - 73)$, $x \in [73,98]$ $\mu_{normal}(x) = 0$, otherwise	$\mu_{high}(x) = (x - 87)/(123 - 87)$, $x \in [87,123]$ $\mu_{high}(x) = (123 - x)/(146 - 123)$, $x \in [123,146]$ $\mu_{high}(x) = 0$, otherwise	$\mu_{very\ high}(x) = 0$ $x \in [1,210]$ $\mu_{very\ high}(x) = (x - 172)/(198 - 172)$, $x \in [172,198]$ $\mu_{very\ high}(x) = 1$, $x \in [198,314]$

Rest electro (separated into 3 levels)		
Alert	Risk	High risk
$\mu_{alert}(x) = 1$, $x \in [1]$	$\mu_{Risk}(x) = 2$, $x \in [2]$	$\mu_{HighRisk}(x) = 3$, $x \in [3]$

Affected
$\mu_{affected}(x) = 1$, $x \in [1]$, $\mu_{affected}(x) = 0$, otherwise

Urine nicotine (separated into 3 levels)		
Alert	Risk	High risk
$\mu_{alert}(x) = (x - 225)/(243 - 225)$, $x \in [225,243]$ $\mu_{alert}(x) = (366 - x)/(366 - 243)$, $x \in [366\text{-}243]$ $\mu_{alert}(x) = 0$, otherwise	$\mu_{high}(x) = (x - 358)/(384 - 358)$, $x \in [358,384]$ $\mu_{high}(x) = (507 - x)/(507 - 384)$, $x \in [384,507]$ $\mu_{high}(x) = 0$, otherwise	$\mu_{very\ high}(x) = 0$, $x \in [1,557]$ $\mu_{very\ high}(x) = (x - 557)/(786 - 557)$, $x \in [557,786]$ $\mu_{very\ high}(x) = 1$, $x \in [631,1473]$

4 Proposed Algorithm

In this paper, we have used the conception of the fuzzy logic and establish the appropriate weights by normalizing the data set according to the derived fuzzy said rules. These normalized data sets are used to predict the risk level in any human being of cardiovascular disease with the impact of diabetics and/or smoking. These will also recognize the approximate acceleration of the disease in the case of smoking. The nicotine factors are included but a common value is considered instead of the type of smoking (cigarette, cigar, bidi, or huka). We have used the above rules and test it using the functions in SciLab.

Proposed Algorithm:

Step 1 [Preprocessing]
Normalized the given data set with respect to the input parameters fuzzy rules described in the Sect. 3. Let (F_1, D), (F_2, S), and (F_f, C) are the fuzzy sets, where (F_1, D) denotes the domination of the diabetics among the patients and (F_2, S) denotes the domination of smoking among the patients data sets and (F_f, D) is defined as (F_f, D X S), where $F_f(\alpha, \beta) = F_1(\alpha)$ ñ $F_2(\beta)$, $\forall\alpha \in D$ and $\forall\beta \in S$ and ñ is the fuzzy intersection operation of two fuzzy sets.

Step 2 [cardinal set]
Discover the cardinal set from the preprocessed data in Step 1.

Step 3 [Approximation Function]
Congregate the multiple aggregation functions of the symptoms to a single fuzzy set using the following

$$\frac{1}{|S|}\sum_{x\in S}\mu_c\Gamma_c(x)\mu_{\gamma c}(x)(p)$$

where S is the set of symptoms and p represents the patients with the membership and cardinal functions.

Step 4 [Decision-Making]
Finally, the maximum value has to be selected from the result obtained by

$$e_i = \sum_{j=1}^{n} p_{ij}$$

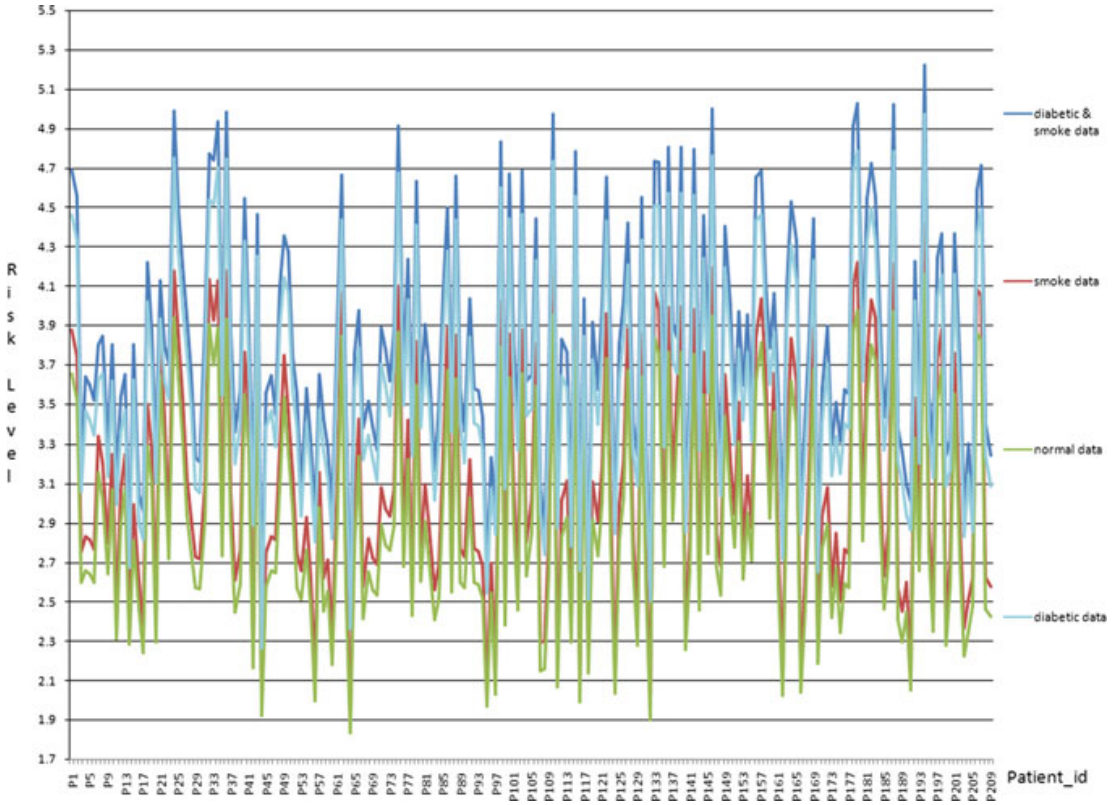
where i represents the signs and j represents the patients affected by that sign.

Thus, the sum of the ith row is represented by e_i and P_{ij} indicates the symptoms in which s_i dominates all the other patients. This will alarm the experts to take the appropriate action immediately.

The above algorithm is to apply on the data set with and without the diabetics as well as the urine nicotine (to find the consequence of smoking), in order to find out the impact of one of the cardiovascular disease. First, we applied the algorithm on the normal data set (data set of the patients with no diabetics and no smoking). Second, the algorithm is applied to the data where the patients are having sugar level (data set of the patients with the diabetics). Third, it is applied on the patients who used to smoke only (data set of the patients not having the diabetic but presence of the nicotine (smoking)) and lastly it is applied to the data of the patients with diabetic and nicotine (smoke). Thus, the algorithm determines the impact of the smoke on the cardiovascular disease with diabetics in the patients. This will not only predict the risk in the patient's life but also supports the experts to take the right and acceptable decision on time so that a person can save his/her precious life and can change the lifestyle accordingly. For this, the experts may treat them (if required) by means of counseling or proper medication.

5 Result Analysis

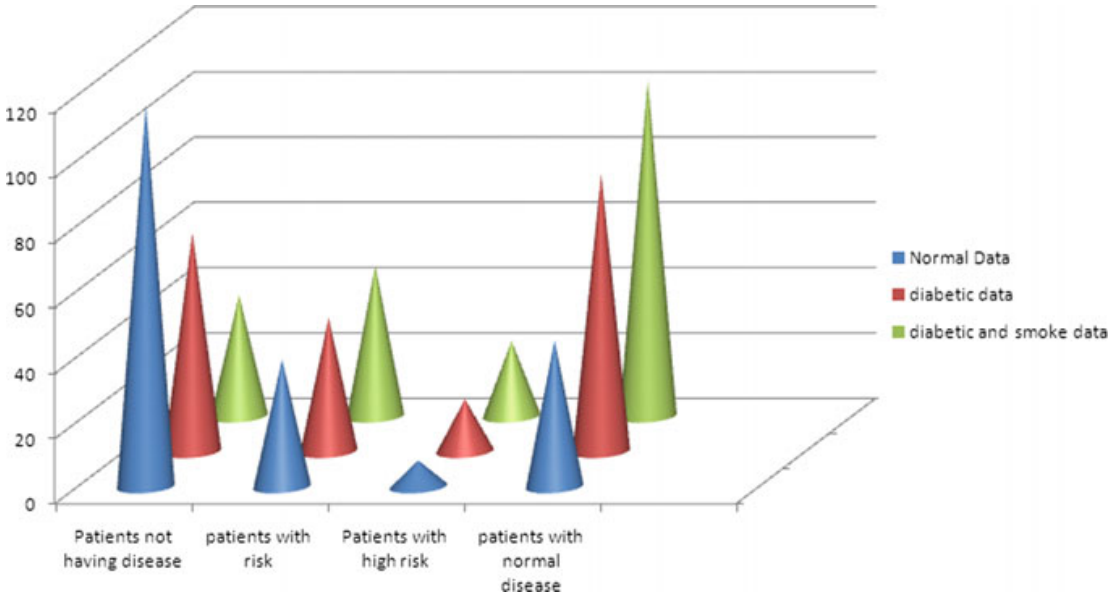
We applied the algorithm in SciLab on the normalized data set. Graph 1 represents the four responsiveness of the algorithm on the UCI data set of cardiovascular disease (in all 209 tuples in the data set, thus 209 patients are considered) which enhanced with the urine nicotine used to measure the smoking factor in the patients. The normal data stands for the input variables except the blood sugar (diabetic) and urine nicotine (smoke). Similarly, the diabetic data set contains the blood sugar, smoke data set consists of the urine nicotine, and diabetic and smoke data set comprise the blood sugar as well as the urine nicotine. Graph 1 represents the normal data (green color), smoke data (red color), diabetic data (light blue color), and smoke and diabetic data (dark blue color). By comparing the facts that diabetic data is more prominent and when it is enhanced with nicotine factor the risk level is at the highest level. The result is evidence for patients accelerating the chances of



Graph 1 Representation of the normal data, diabetic data, smoke data, and diabetic and smoke data

occurrence of the CVD who were not having the CVD initially. Thus, the diabetic and smoke factors augment the risk level of the cardiovascular disease in the patients who even do not have initially the CVD disease. For example, the patient P65 is not affected with the cardiovascular disease but when diabetic as well as smoke factors added the risk factor increased from 1.87 units to 4.26 units with the diabetics and enhanced to 4.63 units when smoking factor added with the diabetics. This illustrates the impact of the diabetic which enhanced it with the additional factor of smoking. This implies that smoking accelerates the risk factor of cardiovascular disease in any human being if he/she is not having the problem at present. On the other hand, the patient is suffering from CVD with the initial value approximately 2.7 augmented to approximately 5.3 with acceleration in the weighted values in presence of the blood sugar and nicotine.

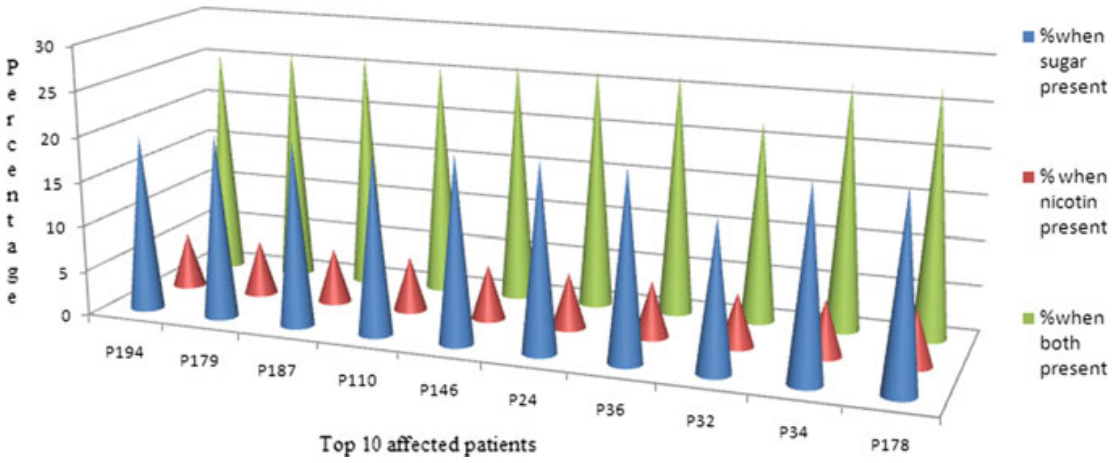
Graph 2 represents the result after the applicability of Step 3 of the algorithm. The approximation function demonstrates the enhancement in the number of patients from normal to risk and risk to high risk factor and more specific the number of the patient with CVD disease increases (earlier it was 117 which were not having the disease, but after the diabetic and with the impact of smoke it remains only 37 who were not yet affected with the cardiovascular disease). This implies that the high risk and risk patients were increased in the ratio of 17.9%, 65.21%, and more observing is 128.88% to the person who was not having the



Graph 2 Number of patients at various stages with respect to data

disease. In the above prediction, the weights of the other input factors were also considered and found that were more prominent with the inclusion of the blood sugar and urine nicotine. Hence, there is a leap in the increased percentage of patients with the risk and not of high risk. This demonstrates that the blood sugar (diabetic) and nicotine (smoke) comprise massive adverse impact on the other factors too. The result congregates the multiple aggregation functions of the symptoms to a single fuzzy set.

Graph 3 represents the patients' who were are at high risk and want an immediate medication. This is the result of Step 4 of the proposed algorithm. The detection of the enhancement in the patient status is based on the multiple-input sign data set. The Step 4 sums up the entire row, and then the difference is to calculate between the row sum and the column sum, in order to find the sign



Graph 3 The top 10 affected patients

overriding on the others. The results showed that the nicotine values were more prominent in the case of the patients with high sugar level. Thus, it is very clear that if a patient is having diabetic, the smoking accelerates the cardiovascular disease. The graph illustrates the top 10 affected patients' records only. It exhibits the amount of the percentage nicotine overrides the blood sugar in respect to the blood pressure, chest pain, angina, report of the electrocardiogram, and heart rate.

6 Conclusion

The concept is used of fuzzy theory which is very effective to handle the dynamic situations because of its characteristics. The proposed algorithm is designed to handle the multiple observations of the input variables which may or may not have the uncertain behavior. The algorithm dealt with the aggregation of the various observation functions based on the observed variables with and without the nicotine (smoke) factor inclusion and finally concludes with the recognition of critical patients who need immediate medication. This will not only facilitate the experts but also helps out in the prediction of the cardiovascular disease on the basis of the current lifestyle of the patient. At last, maybe our experiments results considered to be the starting point for the formation of the fuzzy prediction system for the evaluation of the cardiovascular risk where nicotine is one of the accelerating factors.

The limitation of the paper is of the data considered to be static. In reality, the data is of cumulative nature which augmented in due course time, if appropriately not assessed. Thus, the results will be more worst if the nicotine level amplifies in the body.

References

1. Anooj, P.K.: Clinical decision support system: risk level prediction of heart disease using weighted fuzzy rules. J. King Saud Univ. – Comput. Inf. Sci. **24**(1), 27–40 (2012)
2. Benowitz, N.L., Bernert, J.T., Caraballo, R.S., Holiday, D.B., Wang, J.: Optimal serum cotinine levels for distinguishing cigarette smokers and nonsmokers within different racial/ethnic groups in the United States between 1999 and 2004. Am. J. Epidemiol. **169**, 236–248 (2009)
3. D'Agostino, R.B., Vasan, R.S., Pencina, M.J., et al.: General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation **117**, 743–753 (2008)
4. Elliott, P., Andersson, B., Arbustini, E., et al.: Classification of the cardiomyopathies: a position statement from the European Society of Cardiology Working Group on Myocardial and Pericardial Diseases. Eur Heart J **29**, 270–276 (2008)
5. Nomesko, (2013).: Health Statistics for the Nordic Countries 100. http://norden.diva-portal.org/smash/get/diva2:941584/FULLTEXT01.pdf

6. Onor, I.O., Stirling, D.L., et al.: Clinical effects of cigarette smoking: epidemiologic impact and review of pharmacotherapy options. Int. J. Environ. Res. Public Health **14**, 1147 (2017). https://doi.org/10.3390/ijerph14101147
7. Iftikhar, A.R., Ahmad, M., Siddiqui, A.S.: A study on smoking problem using fuzzy matrix method. Int. J. Math. Arch. **7**(1), 147–152 (2016)
8. Keto, J., et al.: Cardiovascular disease risk factors in relation to smoking behaviour and history: a population-based cohort study. Open Heart **3**, e000358 (2016). https://doi.org/10.1136/openhrt-2015-000358
9. Anbe, J., Tobi, T.: Fuzzy set system application to medical diagnosis: a diagnostic system for valvular heart diseases. Fuzzy Theory Syst **2**, 937–956 (1999)
10. Oad, K. K., DeZhi, X.: A fuzzy rule based approach to predict risk level of heart disease. Glob. J. Comput. Sci. Technol: C Softw. Data Eng. **14**(3.1), 17–22 (2014)
11. Neshat, M., Yaghobi, M., Naghibi, M.B., Esmaelzadeh, A.: Fuzzy expert system design for diagnosis of liver disorders. In: International Symposium on Knowledge Acquisition and Modeling, pp. 252–256 (2008)
12. Matsue, Y., Suzuki, M., Nakamura, R., et al.: Prevalence and prognostic implications of pre-diabetic state in patients with heart failure. Circ. J. **75**, 2833–2839 (2011)
13. Phuong, N.H., et al.: Fuzzy logic and its applications in medicine. In: Proceedings of Asian Pacific medical informatics conference APAMI-MIC'2000, Hong Kong, Sep 27–30, pp 1–11 (2000)
14. Ponikowski, P., Voors, A.A., Anker, S.D., et al.: ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur. J. Heart Fail. **18**, 891–975 (2016)
15. Rajeswari, K., Vaithiyanathan, V.: Heart disease diagnosis: an efficient decision support system based on fuzzy logic and genetic algorithm. Int. J. Decis. Sci. Risk Manag. **3**(1–2), 81–97 (2011)
16. Wild, S., Roglic, G., Green, A., et al.: Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care **27**(5), 1047–1053 (2004)
17. Kumar, S., Kaur, G.: Detection of heart diseases using fuzzy logic. Int. J. Eng. Trends Technol. (IJETT) **4**(6), 2694–2699 (2013)
18. Phongsuphap, S., Pongsupap, Y.: Evaluation of responsiveness of health systems using fuzzy-based technique. In: IEEE International Conference on Fuzzy Systems, pp. 1618–1623 (2014)
19. Mahmood, S.S., et al.: The Framingham Heart Study and the epidemiology of cardiovascular diseases: a historical perspective. Lancet **383**(9921), 999–1008 (2014)
20. U.S. Department of Health and Human Services. The Health Consequences of Smoking—50 Years of Progress: A Report of the Surgeon General. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, Atlanta, GA, USA, pp. 1–36 (2014)
21. United States Department of Health and Human Services. How Tobacco Smoke Causes Disease: A Report of the Surgeon General. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, Atlanta, GA, USA (2010)

An Integrated Fault Classification Approach for Microgrid System

Ruchita Nale, Ruchi Chandrakar and Monalisa Biswal

Abstract In this paper, a moving windowing approach-based integrated fault classification algorithm is proposed for microgrid system. In a microgrid system, the nonlinear operation of control devices connected to distributed generation (DG) imposes problem for identifying the exact faulty class. In order to mitigate this issue, an integrated moving window averaging technique (IMWAT) is proposed. The method utilizes current signal at the line end. In this technique, first, the decision of the fault detection unit (FDU) is analyzed and based on that fault class is detected. The FDU uses the conventional moving window averaging technique. Different logics are framed to identify the symmetrical and unsymmetrical faults. The method is tested on a standard microgrid network and obtained results for different fault cases prove the efficacy of the proposed method.

Keywords Close-in fault · Fault classification · High resistance fault · Microgrid system · Moving window averaging technique · PV system

1 Introduction

Microgrid system is emanating as a primary part of distribution network owing to the technological advancement in distributed generation (DG) system [1–6]. It consists of various renewable energy sources such as wind power, fuel cells, and

R. Nale · R. Chandrakar · M. Biswal (✉)
Department of Electrical Engineering, National Institute of Technology, Raipur 492010, Chhattisgarh, India
e-mail: mbiswal.ele@nitrr.ac.in

R. Nale
e-mail: ruchita0119@gmail.com

R. Chandrakar
e-mail: ruchichandrakar14@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_4

photovoltaic cells providing a means to reduce greenhouse gas emissions and to supply clean and continuous power. These DG sources can be employed near loads for reliable power supply as it decreases the power transmission losses and enhances the energy efficiency of the power system [7–9]. Microgrids can be operated in two modes, i.e., when connected to grid and also when it is isolated from the main grid (islanded mode). The most significant feature of microgrid is its ability to operate in autonomous mode when there is fluctuation in voltage, deviation in frequency or when there is a fault in main grid. Hence, ensuring continuity of supply to critical loads. Enhanced service quality, improved reliability, and efficiency are the common advantages offered by the microgrid [10]. However, usage of different types of DG for generation of power poses challenges for operating, integrating, controlling, and protecting the microgrid [11–16].

Fault detection is an essential and critical issue in microgrid system because of the nonlinear operation of control devices connected with renewable sources and due to the wide variation in fault current magnitude in different operating modes such as grid-connected mode and islanded mode. Also, the presence of different types of DGs affect the magnitude of fault current. Owing to high penetration of inverter interfaced DG protection of microgrid is a tedious task. The contribution of fault current by inverter interfaced DGs is limited to two to three times the rated current [17].

Different solutions have been provided by different researchers to detect and classify the types of fault in microgrid system. The diagram of a general fault classification algorithm used in protective relay is provided in Fig. 1. In this diagram, the basic procedure for fault classification function performed by relay is presented. With the help of fault generated voltage and current signal, different sequence components are calculated and the relative angle between two quantities are used to estimate the exact faulty phase.

In [18], two new fault classification logics are proposed to identify the accurate faulty phase in microgrid system. The first method is based on voltage angle and magnitude and the second method used voltage angle to identify the fault class. A new approach-based on optimal wavelet functions is proposed to detect the exact faulty class and reported in [19].

In this paper, a new method is proposed. The method operates on the basis of two integrated logics. This IMWAT method overcomes the drawbacks of existing moving window averaging which fails to work correctly during high resistance fault, nonlinear operation of control devices connected to wind turbines, switching of nonlinear loads, and mode changing operation of microgrid. The proposed method is tested on a standard microgrid system, i.e., Aalborg distribution system model. The power system model is developed in EMTDC/PSCAD environment. The response of the method is verified for different symmetrical faults, unsymmetrical fault, high resistance fault, close-in fault, switching of nonlinear loads, and

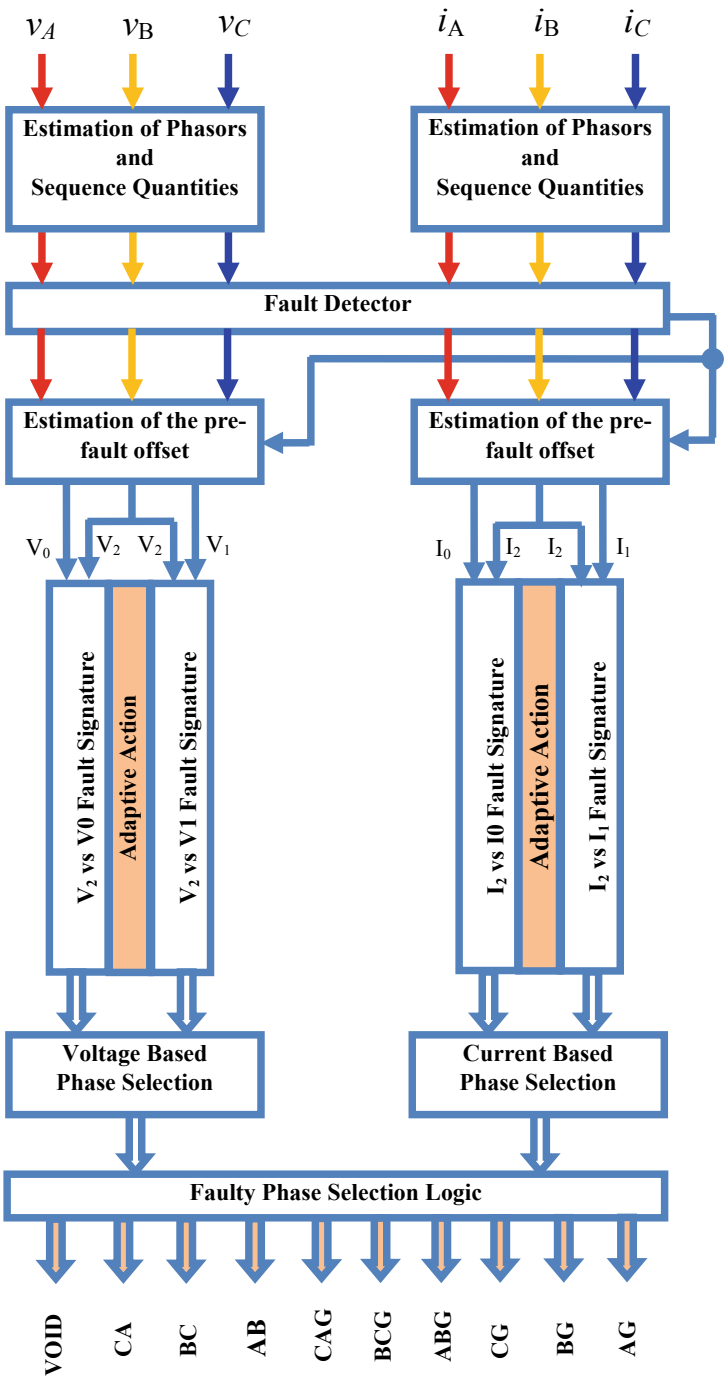


Fig. 1 Fault classification algorithm used in protective relay

different operating modes. From the analysis report, it is confronted that the proposed method is able to identify the exact faulty phase within half-cycle time period from the instant of fault occurrence and the decision time is fixed irrespective of the system and fault conditions.

The main body of the paper is organized as follows. In Sect. 2 fault classification based on IMWAT approach is described. Next, simulation results for different test conditions are provided in Sect. 3. The conclusion of the paper is provided in Sect. 4.

2 Fault Classification in Microgrid

2.1 Moving Window Averaging Technique

During transient phenomenon or fault scenarios, the average of one window current signal is a non-zero quantity. With suitable threshold value exact fault detection is possible. In this work, a moving window averaging technique based on fault detector unit is considered for processing the fault classification task using IMWAT technique.

Using moving window averaging technique, the magnitude of one cycle current samples can be modeled as

$$i_{\mathrm{MW}}(s) = \frac{1}{\mathrm{N}} \sum_{p=s-\mathrm{N}+1}^{s} i(p) \tag{1}$$

where s is the sampling instant, p is the instantaneous sample, and N is the number of samples per cycle.

The relation in recursive form can be represented as

$$i_{\mathrm{MW}}(s) = i_{\mathrm{MW}}(s-1) + i(s) - i(s-\mathrm{N}) \tag{2}$$

Using (2), a fault detector is developed in the proposed method. The decision of the FDU is employed in the classification algorithm to classify the exact faulty phase in microgrid system. From the several advantages, detection capability during high resistance fault is a unique feature of moving window averaging technique. Under far-end high resistance fault, the faulty phase current will be very less and can be in the same order as healthy phase. Even though the scenario is rare, but with the arrival of such an event, moving window averaging technique will fail to detect exact faulty phase within the stipulated time period. Such a case is undesirable and degrade the performance of protection function. From the study, it is observed that under faulty condition, the amplitude of affected phase moving window average value is more as compared to other phases if an observation is drawn on the basis of a fixed half-cycle calculated value just after fault inception. This unique feature can be incorporated for the detection of exact faulty phase in microgrid system. In order

to further enhance the reliability of relay, an integrated approach-based fault classification technique is devised that integrates the decision of FDU along with other computed paradigms which are described in the next section.

2.2 IMWAT Algorithm [19]

The proposed algorithm for classification of faults is as follows:

(1) The value of i_{MW} is computed using Eq. (1) by considering current sample for one cycle.
(2) Using recursive update formula mentioned in Eq. 2, value of i_{MW} for each phase is calculated.
(3) For detection of faults, threshold (β_t) is chosen to prevent false operation of relaying. For any further deviation above the threshold value, the threshold counter (j) counts 1. Each counter is provided for each phase.
(4) Check if $|i_{MW}| \geq \beta_t$

If yes, j = 1, otherwise 0. Then $j_{new} = j_{old} + 1$.

Then threshold counter continues to evaluate for consecutive three samples. If any time j = 0 comes in-between counter again reset to zero otherwise continue counting.

If $j_{new} = 3$, FDU = 1 otherwise 0.

The aforementioned process is performed for the fault detection task in three phases. FDU is reset to zero once the fault is detected. Next, the further steps of fault classification are evaluated and are explained below.

(5) Next the residual current through the ground can be calculated using formula,

$$i_g(s) = (|i_a(s) + i_b(s) + i_c(s)|) \tag{3}$$

In (3), i_a, i_b, and i_c represent the instantaneous current for phase a, b, and c, respectively. i_g aids in distinguishing ground fault from phase fault.

(6) Along with fault detection, the developed fault classification algorithm evaluates, $|i_{MW_a}|, |i_{MW_b}|, |i_{MW_c}|$, the summation and difference of two-phase moving sum current based on continuous half-cycle considering from the moment of occurrence of fault.

For example, the fault is detected at sth sample, so the proposed method initiates computing above data from (s − 2)th to (s + 7)th sample. The observation duration

is kept as half-cycle for better comparison and correct classification of different faults.

The computational procedure of consecutive phase summation and difference are provided below.

$$\mathrm{MS_{ab}}(s) = ||i_{\mathrm{MW_a}}(s)| + |i_{\mathrm{MW_b}}(s)||$$
$$\mathrm{MS_{bc}}(s) = ||i_{\mathrm{MW_b}}(s)| + |i_{\mathrm{MW_c}}(s)||$$
$$\mathrm{MS_{ca}}(s) = ||i_{\mathrm{MW_c}}(s)| + |i_{\mathrm{MW_a}}(s)||$$

Similarly,

$$\mathrm{MD_{ab}}(s) = ||i_{\mathrm{MW_a}}(s)| - |i_{\mathrm{MW_b}}(s)||$$
$$\mathrm{MD_{bc}}(s) = ||i_{\mathrm{MW_b}}(s)| - |i_{\mathrm{MW_c}}(s)||$$
$$\mathrm{MD_{ca}}(s) = ||i_{\mathrm{MW_c}}(s)| - |i_{\mathrm{MW_a}}(s)||$$

For each calculated value of i_g, $|i_{MW}|$, *MS*, and *MD* a data matrix of size 10 × 10 is devised. Suppose the data matrix of size n x n is denoted by D_{nn}, where n is equal to 10. D_{nn} can be expressed as,

$$\mathrm{D_{nn}} = \begin{bmatrix} \mathrm{D_{11}} & .. & .. & \mathrm{D_{n1}} \\ .. & .. & .. & .. \\ .. & .. & .. & .. \\ .. & .. & .. & .. \\ .. & .. & .. & .. \\ \mathrm{D_{1n}} & .. & .. & \mathrm{D_{nn}} \end{bmatrix}$$

where $D_{11} = i_{g1}$, D_{21} to D_{n1} are $|i_{\mathrm{MW_a}}|_1, |i_{\mathrm{MW_b}}|_1 |i_{\mathrm{MW_c}}|_1$,

$$MS_{\mathrm{ab_1}}, MS_{\mathrm{bc_1}}, MS_{\mathrm{ca_1}}, MD_{\mathrm{ab_1}}, MD_{\mathrm{bc_1}}, \text{and } MD_{\mathrm{ca_1}}.$$

(7) From each column indices are selected. For example, the maximum value in column 1 is *f*1 which denotes the maximum of i_g from a data matrix of size 10. As the calculated values are not constant so direct comparison is not possible, hence index values are selected.

Similarly, *f*2 = maxim $(|i_{\mathbf{MW_a}}|)$, *f*3 = maxim $(|i_{\mathbf{MW_b}}|)$, *f*4 = maxim $(|i_{\mathbf{MW_c}}|)$, *f*5 = maxim (MD_{ab}), *f*6 = maxim (MD_{bc}), *f*7 = maxim (MD_{ca}), *f*8 = maxim (MS_{ab}), *f*9 = maxim (MS_{bc}), *f*10 = maxim (MS_{ca}).The different indices are represented as *f*1 to *f*10 that will compare among each similar group with minim, maxim function. Rules for each fault type is developed and is given in Table 1.

Table 1 Fault classification logic

Fault Type	Logic-1	Logic-2	Logic-3	Logic	Output
	$f1$	minim ($f5$, $f6$, $f7$)	minim ($f8$, $f9$, $f10$)		
a–g	$\geq \beta_t$	$f6$	$f9$	AND	1
bg		$f7$	$f10$		2
c–g		$f5$	$f8$		3
		minim ($f5$, $f6$, $f7$)	maxim ($f8$, $f9$, $f10$)		
ab–g	$\geq \beta_t$	$f5$	$f8$	AND	4
bcg		$f6$	$f9$		5
cag		$f7$	$f10$		6
		minim ($f2$, $f3$, $f4$)	minim ($f5$, $f6$, $f7$)		
ab	$\leq \beta_t$	$f4$	$f5$	AND	7
bc		$f2$	$f6$		8
ca		$f3$	$f7$		9
abc		$f2 \geq \beta_t$ AND $f3 \geq \beta_t$ AND $f4 \geq \beta_t$			10

As mentioned in Table 1, if all the rules (rule 1,2 and 3) are satisfied at the same time for any fault in the line for which AND logic is applied then a particular fault case is recorded. For example, a fault occurred with $f1 \geq \beta_t$, AND minim ($f5$, $f6$, $f7$) = $f6$ AND minim ($f8$, $f9$, $f10$) = $f9$ then output is 1, i.e., a-g fault. To avoid ambiguity in decision of the relay during single-line-to-ground and double-line-to-ground fault two different functions are utilized. In double-line-to-ground fault both minim and maxim function is used, whereas in single-line-to ground fault only minim function is employed. Hence, Table 1 indicated that a secure fault classification algorithm can be devised using integrated moving sum technique.

3 Simulation Results

The performance of the proposed technique is evaluated by considering a standard microgrid system shown in Fig. 2. It is a part of a distribution network, owned by Himmerlands Elforsyning, in Aalborg, Denmark. The system consists of three wind turbine generators and one combined heat and power plant unit. Generator data, line data, and grid data are considered from [20]. To test the efficiency of the proposed fault classification approach in the presence of various types of distributed generation units, the system is modified by replacing DG3 with photovoltaic (PV) system at bus 14. The nominal power generation of DFIG wind turbine generator and PV is

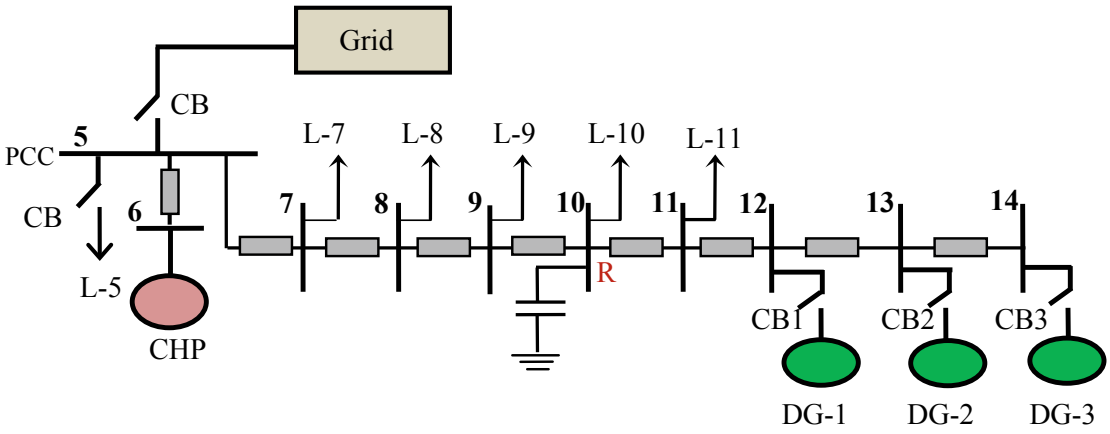


Fig. 2 Standard distribution network

2 MW and 630 kW, respectively. The simulation of the given network is performed in EMTDC/PSCAD. MATLAB software is used for logic development and the response of the technique is evaluated for different fault cases such as unsymmetrical and symmetrical faults, high resistance faults, close-in faults, nonlinear switching of load, and change in operating mode of microgrid system. The obtained results from each case are described below.

3.1 Single-Line-to-Ground Fault (LG)

In distribution system, LG faults are frequently occurring faults. To assess the potential of the proposed technique, a-g fault is simulated considering 5 Ω as the fault path resistance at 0.3 s in-between bus 10–11. The response of both fault detector unit and IMWAT are presented in Fig. 3. The computed absolute values of the average of one window current signal for each phase are shown in the figure. For phase-a, the measurement for three consecutive computed values is greater than the threshold, then the Counter (j) is set to 3. Hence, the output of FDU for all the three phases is [1, 0, 0] at 0.303 ms and is depicted in Fig. 3(b). So, the algorithm takes 3 ms to detect the fault using IMWAT. The proposed method further evaluates the value of i_g, MD, and MS as shown in Fig. 3. From the plots, it is noticed that $f1 \geq \beta_t$ AND minim ($f5$, $f6$, $f7$) = $f6$ AND minim ($f8$, $f9$, $f10$) = $f9$. Since all the conditions for a-g fault scenario are satisfied concurrently, the decision provided by relay is 1 which is correct.

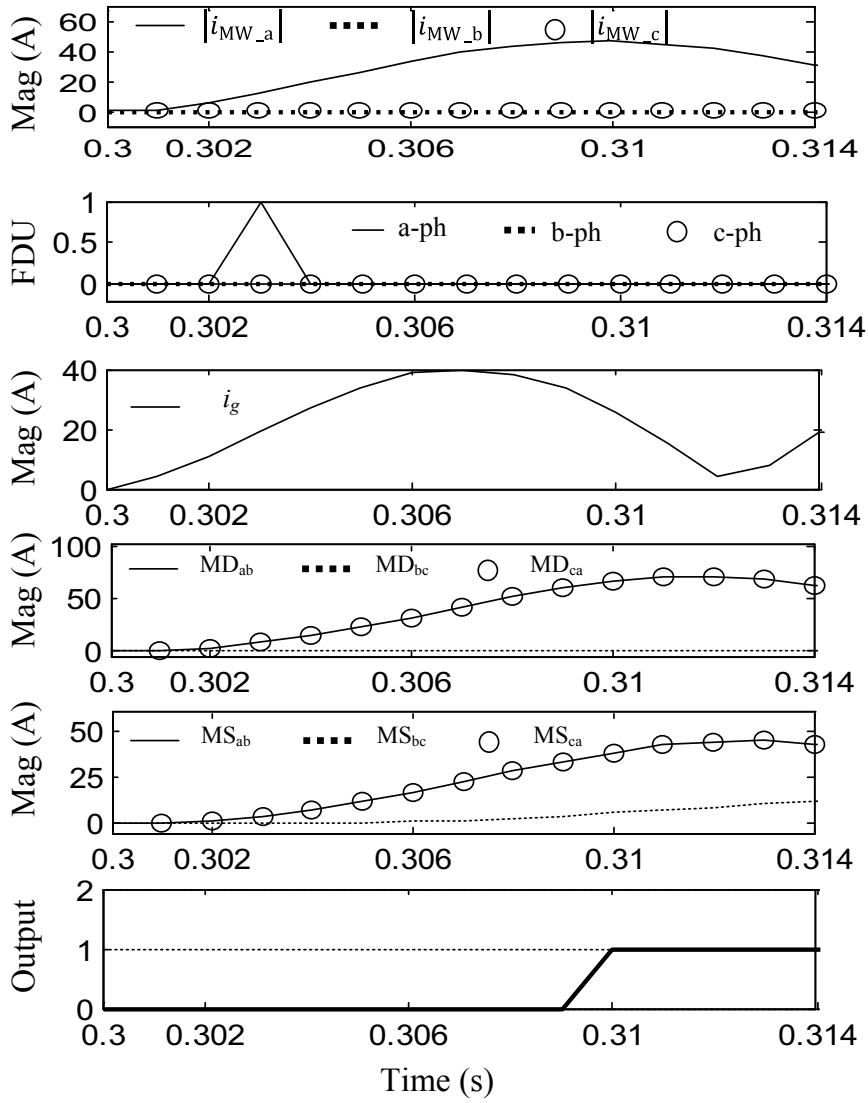


Fig. 3 Performance of FDU and proposed method during a–g fault

3.2 High Resistance Double-Line-to-Ground Fault

Next, to observe the response of proposed IMWAT technique for high resistance double-line-to-ground fault scenario, ab-g fault is created in-between bus 10–11 with a fault resistance of 150 Ohm. Fault is created at 0.3 s. The performance of relay for fault detector unit and proposed method is recorded and is depicted in Fig. 4. During far-end high resistance fault condition, the response of FDU for each phase is "1". As a consequence, the decision interpreted by the relay is incorrect. Hence, the information provided by moving window averaging approach is not adequate for correct faulty phase selection. In order to achieve the exact

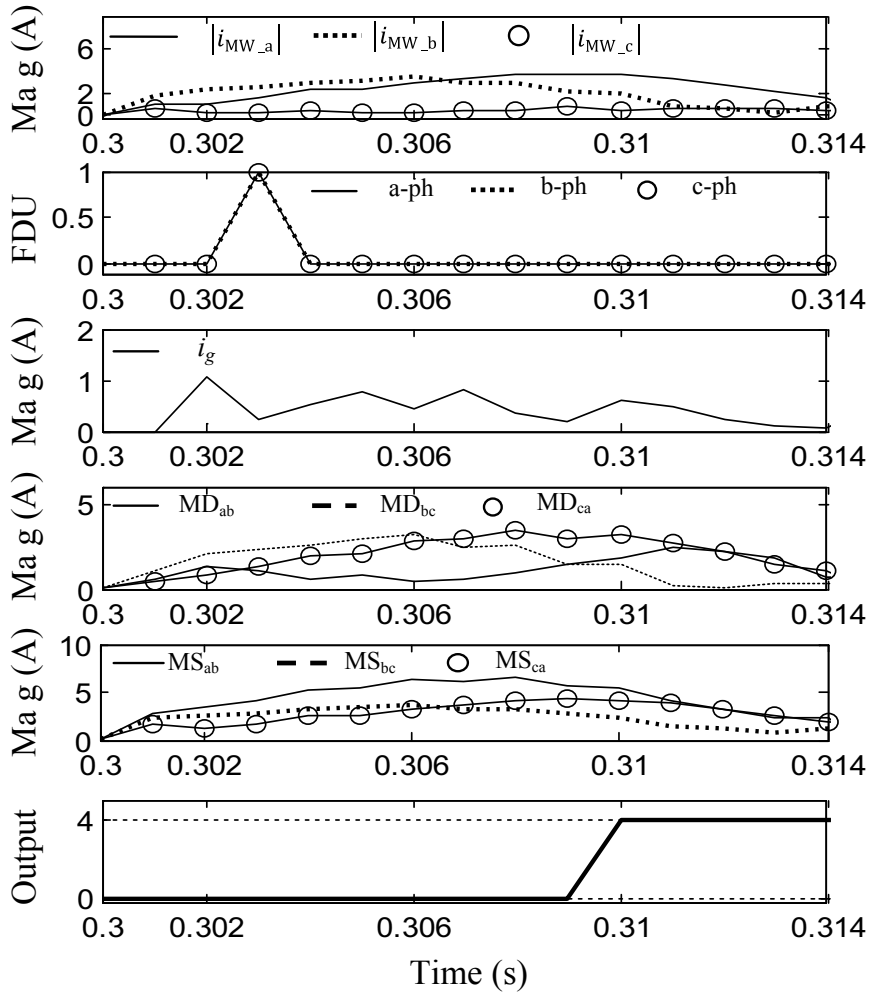


Fig. 4 Performance of FDU and proposed method during far-end high resistance a–b–g fault

classification of faults, the performance of moving window average-based FDU can be enhanced by employing the proposed IMWAT technique. So, by the integration of various proposed logics, the relay provides decision as "4", i.e., ab-g fault which is correct.

3.3 Double Line Fault

For the simulation of double line fault to observe the response of the proposed method, a b–c fault in-between bus 10–11 is created at 0.3 s. The fault resistance is considered as 5 Ohm. The response of the FDU and the proposed method are shown in Fig. 5. From the response of FDU it is cleared that, FDU only detects the

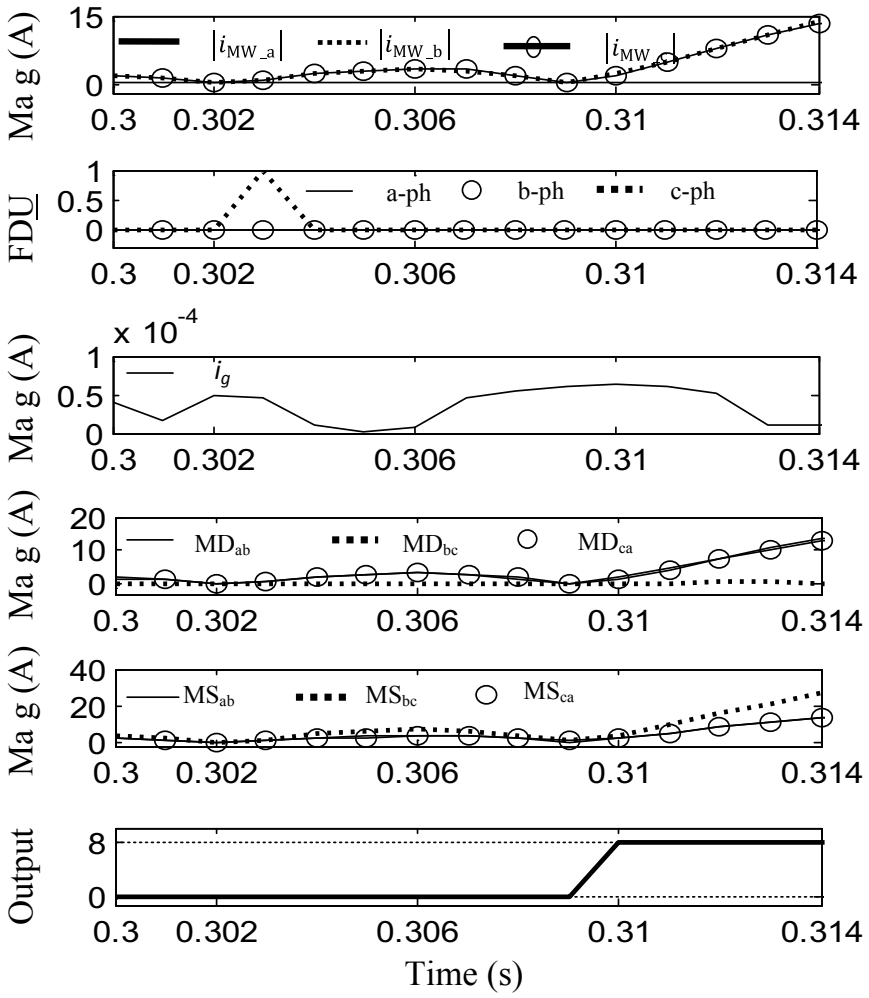


Fig. 5 Performance of FDU and proposed method during b–c fault

phase-c for b–c fault scenario. But the output of the proposed method is "8", which indicates the occurrence of b–c fault in microgrid system.

3.4 Close-in Fault

Close-in faults are more severe because the higher fault current magnitude deteriorates the performance of current transformer. So, relay finds more challenge during close-in fault. To simulate this condition, three-phase fault is simulated in-between bus 10–11 at 0.3 s. The fault resistance is 2 Ohm. By moving sum approach fault is detected at 0.303 s I all the three phases. The output of FDU for all the three phases will be [1]. From Fig. 6, $f1 \geq \beta_t$ AND $f2 \geq \beta_t$ AND $f3 \geq \beta_t$ AND $f4 \geq \beta_t$ this provides an output of 10, i.e., three-phase fault by IMWAT approach. Therefore,

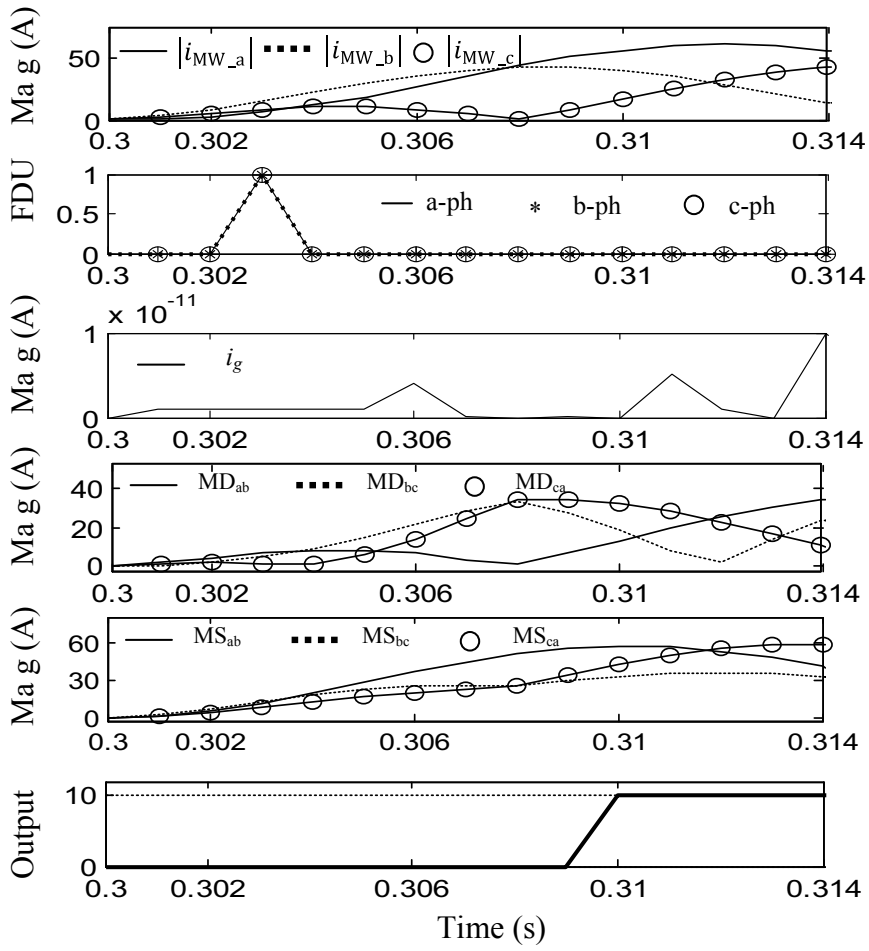


Fig. 6 Performance of FDU and proposed method during close-in three-phase fault

the algorithm can correctly classify close-in three-phase fault which is clear from the result.

3.5 Fault During Switching of Nonlinear Load

The application of electronic gadgets such as personal computer, laptop, printers, uninterrupted power supply, and music instruments, the generation of harmonics in distribution level is more. This may lead to false operation of fault classification algorithm. So, to verify the impact of the proposed method during switching of nonlinear loads, a 0.5 MW three-phase diode rectifier with resistive load is switched on at 0.3 s. At the same time, c-g fault is created in the line between bus 10–11.

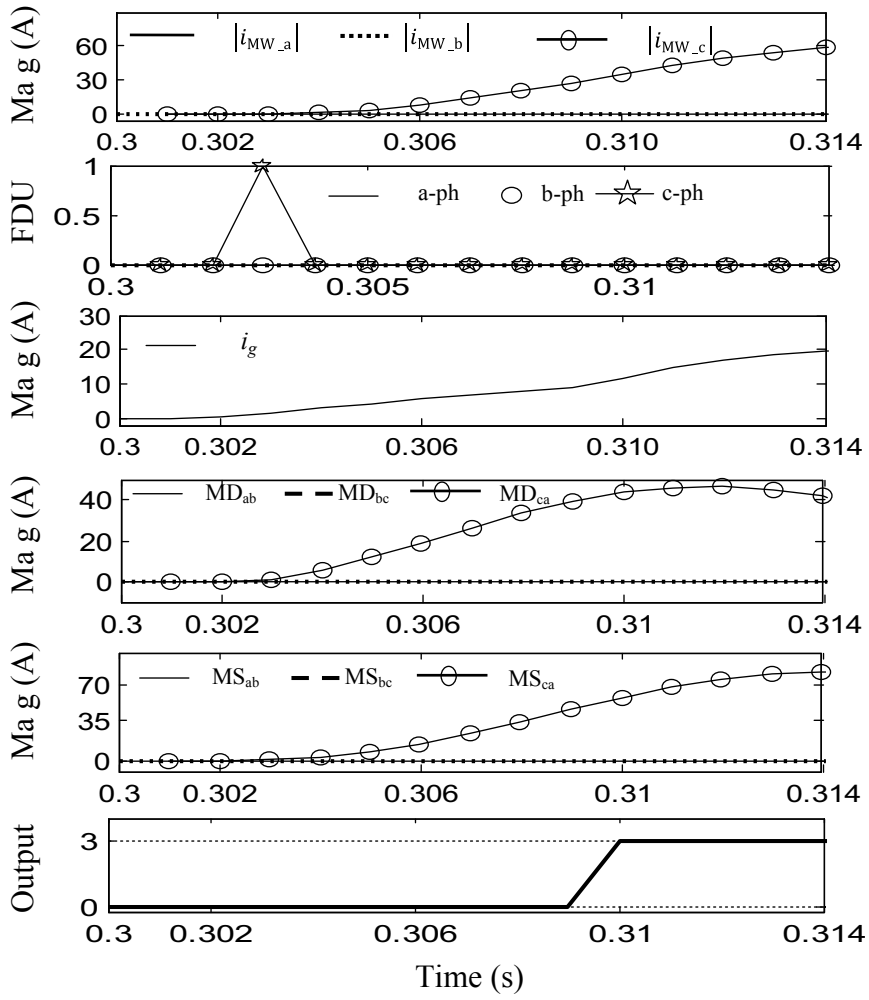


Fig. 7 Performance of FDU and proposed method during nonlinear load switching

As depicted from the output response of the proposed method shown in Fig. 7, the switching of nonlinear load has no significant effect on the proposed method.

3.6 Fault During Different Operating Modes

All the abovementioned results are provided for grid-connected mode. So, to verify the response of the proposed method during islanding mode, the PCC bus breaker, i.e., CB-4 is kept open. For this islanding case, an a–g fault is created at 0.305 s. The response of the proposed method for this case is shown in Fig. 8. From this figure, it is clear that the response of the method is also accurate for islanded mode as the final output by IMWAT technique is "1" which is correct.

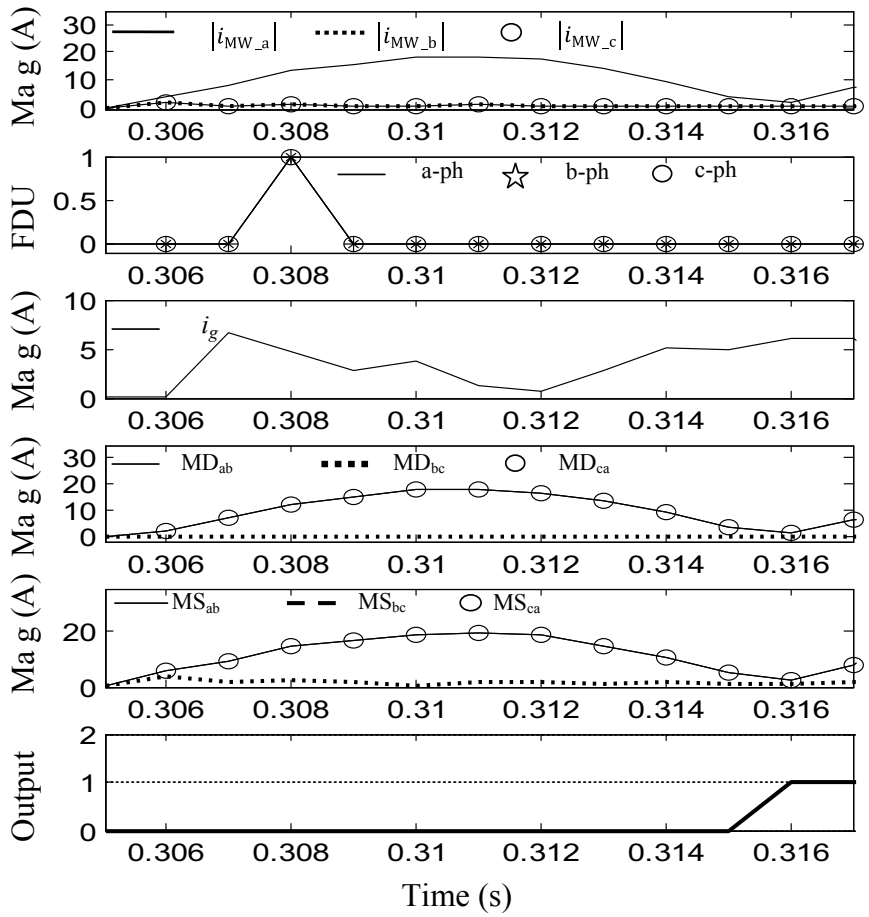


Fig. 8 Performance of FDU and proposed method for islanded mode of operation

So, the response of the method is verified for different critical cases and it is noticed that within one cycle time period accurate decision can be drawn irrespective of the critical operating conditions of fault and system.

4 Conclusion

Exact fault type selection is a challenging task in microgrid system due to the incorporation of renewable sources in the existing passive distribution network. The nonlinear operation of control devices, switching of nonlinear load ad changing mode of operation from grid-connected mode to islanded mode is the typical critical operating condition of microgrid system during which relay find challenges in discriminating the fault types. The other fault conditions such as close-in fault and high resistance fault may degrade the performance of existing protective relay

algorithms. To mitigate this technical and fault generated issues, a new integrated approach is proposed in this paper. In the integrated approach, the output of moving window averaging-based FDU is incorporated in the fault classification logic to obtain accurate decision. For ground fault detection residual current through ground s computed. Different indices are calculated for the fault class selection. The response of the proposed method is verified for various cases. It is observed from the results that within half-cycle time period exact fault type selection is possible in microgrid system.

References

1. IEEE Standard for Interconnecting Distributed Resources with Electric Power Systems. IEEE, Standard 1547-2003 (2003)
2. Impact of Increasing Contribution of Dispersed Generation on the Power System. CIGRE, Working Group 37.23 (1999)
3. G59/1 Recommendations for the connection of Embedded Generating Plant to the Regional Electricity Companies Distribution Systems, Electricity Association standard (1991)
4. Xu, W., Zhang, G., Li, C., Wang, W., Wang, G., Kliber, J.: A power line signaling based technique for anti-islanding protection of distributed generators-part I: scheme and analysis. IEEE Trans. Power Deliv. **22**(3), 1758–1766 (2007)
5. Wang, W., Kliber, J., Zhang, G., Xu, W., Howell, B., Palladino, T.: A power line signaling based scheme for anti-islanding protection of distributed generators—part II: field test results. IEEE Trans. Power Deliv. **22**(3), 1767–1772 (2007)
6. Ahmad, K.N.E.K., Selvaraj, J., Rahim, N.A.: A review of the islanding detection methods in grid-connected PV inverters. Renew. Sustain. Energy Rev. **21**, 756–766 (2013)
7. PVPS IEA: Evaluation of islanding detection methods for photovoltaic utility interactive power systems. Report IEA PVPS T5-09 (2002)
8. Velasco, D., Trujillo, C.L., Garcera, G., Figueres, E.: Review of anti-islanding techniques in distributed generators. Renew. Sustain. Energy Rev. **14**(6), 1608–1614 (2010)
9. Timbus, A., Oudalov, A., Ho, C.: Islanding detection in smart grids. In: Energy Conversion Congress on Exposition (ECCE), pp. 3631–3637 (2010)
10. Nale, R., Biswal, M.: Comparative assessment of passive islanding detection techniques for microgrid. In: Proceedings of International Conference *on* Innovations Information, Embedded and Communication Systems, Coimbatore, pp. 1–5 (2017)
11. Nale, R., Biswal, M., Kishor, N.: A transient component based approach for islanding detection in distributed generation. In: IEEE Transactions on Sustainable Energy (Early Access) (2018)
12. Ray, P.K., Kishor, N., Mohanty, S.R.: S-transform based islanding detection in grid-connected distributed generation based power system. In: Proceeding of IEEE International Energy Conference and Exhibition (EnergyCon), pp. 612–617 (2010)
13. Ray, P., Mohanty, S., Kishor, N.: Disturbance detection in grid connected distribution system using wavelet and S-transform. Electr. Power Syst. Res. **81**(3), 805–819 (2011)
14. Mohammadzadeh Niaki, A.H., Afsharnia, S.: A new passive islanding detection method and its performance evaluation for multi DG systems. Electr. Power Syst. Res. **110**, 180–187 (2014)
15. Mishra, M., Sahani, M., Rout, P.K.: An islanding detection algorithm for distributed generation based on Hilbert-Huang trabsform and extreme learning machine. Sustain. Energy Grids Netw. **9**, 13–26 (2017)

16. Mahat, P., Chen, Z., Bak-jensen, B.: Review on islanding operation of distribution system with distributed generation. In: Power and Energy Society General Meeting, pp. 1–8 (2011)
17. Keller, J., Kroposki, B.D.: Understanding fault characteristics of inverter-based distributed energy resources. Nat. Renew. Energy Lab., Golden, CO, USA, Tech. Rep. NREL/TP-550-46698 (2010)
18. Abddelagayed, T.S., Morsi, W.G., Sidhu, T.S.: A new approach for fault classification in microgrids using optimal wavelet functions matching pursuit. IEEE Trans. Smart Grid 9(5) (2018)
19. Biswal, M.: Faulty phase selection for transmission line using integrated moving sum approach. IET Sci. Meas. Technol. **10**(7), 761–767 (2016). https://doi.org/10.1049/iet-smt.2016.0081
20. Mahat, P., Chen, Z., Bak-Jensen, B.: A hybrid islanding detection technique using average rate of voltage change and real power shift. IEEE Trans. Power Deliv. **24**(2), 764–771 (2009)

Role of Data Analytics in Human Resource Management for Prediction of Attrition Using Job Satisfaction

Neerja Aswale and Kavya Mukul

Abstract The reputed management publications like Harvard Business Review (HBR) have started stressing upon the emergence of data-driven management decisions. The enhancing investments in data and analytics are underlining the aforementioned emergence. According to International Data Corporation, this investment is expected to grow up to $200 billion by 2020. In such a data lead management world collecting, managing, and analysing the human resources-related data becomes a key for any rather every organization. Human resource analytics is changing into necessary as strategic personnel designing is the need of the hour and helps organizations to investigate each side of HR metrics. HR analytics could be a holist approach. According to KPMG—India's Annual Compensation Trends Survey 2018–19 the average annual voluntary attrition across sectors is 13.1%. This is a considerably high percentage. Hence, antecedents leading to attrition are needed to be explored in order to propose appropriate HR policies, strategies, and practices. In relevance to these facts, this study focused on proposing a data-driven predictive approach that examines the relationship between the attrition (dependent variable) and other demographic and psychographic independent variables (Antecedents). The present study found that there is a strong relationship between job satisfaction and attrition. Further, there is a higher probability that the employees having work experience between 0–5 years may leave the organizations. Such data-based outcomes may offer help to HR managers in addressing the problems like attrition which intern may increase ROI. Thus, this paper underlines the emergence and relevance of analytics with special reference to human resource management domain.

Keywords Analytics · Human resources management · Attrition · Job satisfaction

N. Aswale (✉) · K. Mukul
MIT School of Technology Management, Vishwashanti Marg, Rambaug Colony, Kothrud, Pune, Maharashtra 411038, India
e-mail: neerjaaswale@gmail.com

K. Mukul
e-mail: kavya.mukul21@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_5

1 Introduction

The emergence of technologies has changed the way the business domains used to be operated. Taking these developments into consideration it has become important to accept that the various business functions have started deploying the technological advancements. The human resources management is no longer an exception to this phenomenon. Therefore, the "Human resource analytics (HR analytics) is a vicinity in the subject of analytics. The HR analytics refers to application of analytic techniques to the functions of human resource department. The major emphasis of HR department is to achieve enhanced productivity by increasing individual performance of the employee. The increased individual performance will subsequently result into a higher return on investment." The modern philosophy of human resource management has shifted focus from Human resources to Human Capital. This has added roles and responsibilities of HR professionals. The modern HR professionals are suppose to predict futuristic standpoints to the organization using analytics. This shift demands proper exploration, examination, and evaluation of individual performances. This approach will further help for talent management practices that focuse on retention and acquisition of human resources. In such cases, the data-based evidence would help HR managers in designing the HR strategies. These evidence/trends can be analyzed using analytics that includes application modern business analytics techniques like data mining to human resources data. In continuation to this introductory insights the next section details about the need for using HR analytics for strategic HR decision-making. Further, it also informs about studies conducted in HR domain for understanding the application of HR analytics.

Being a multidisciplinary method, HR analytics combines methodology that improves the people-related selections to enhance person and company's performance. HR analytics is also known as people analytics, workforce analytics, and talent analytics. The HR analytics introduces high-end predictive modeling approaches that forecast the effect of changing the policies. Further, the HR analytics has become an integral part of human resources functions like recruitment, training, development, succession planning, compensation, benefits, retentions, and engagement. Traditionally, analytics is being used to focus on the rate of turnover and the cost of hiring. Using complex statistical analyses, HR analytics enables corporations to measure the business impact of people policies and predicts the future of the workforce. This allows managers to measure the relationship between financial performance of the organization and the human resource practices.

Muscalu and Serban [1] examined the relevance of HR analytics for strategic human resource management. The authors deployed HCM: 21 model to know the effectiveness of the department. Momin and Taruna [2] reported that the growing HR-related problems such as succession planning, recruitment, and employee retention. Sujeet et al. [3] mentioned that Human Resource Predictive Analytics (HRPA) was found useful for all HR functions. Bindu [4] conducted a review emphasizing on HR analytics skills and as a result of focus group discussions reported that bad records would lead to negative HR analytics, whereas business

analytics provides a multidimensional method toward constructing effective HR strategies [5]. Globcon Technologies Pvt. Ltd [6] reported that application of HR analytics has opened emerging niche areas in HR analytics as certain specific problems get addressed by means of HR analytics. Rajbhar et al. [7] concluded that HR analytics helps in improving workers overall performance in a team. This study also mentions that HR analytics can be used for increasing the productivity. Shweta [8] concluded that applying analytic procedures to the human resource data help in enhancing worker performance and consequently getting a higher return on investment.

2 Methodology

It was observed during literature review that the role of HR analytics is very well explored in developed countries and hardly explored for developing countries like India. Further, the practical relevance of HR analytics is hardly explored. Even though these study mentioned about relevance of HR analytics, a few studies have explored a specific phenomenon like attrition. This has created a scope for the present study that emphasized on attrition and explored the various factors related to this phenomenon. To have a systematic approach this study defined certain research objectives. These objectives are as follows:

(a) To study the factors affecting attrition.
(b) To predict the factors influence attrition.
(c) To identify factors affecting job satisfaction.
(d) To make recommendations for retention of employees on the basis of analysis.

H0: Attrition is directly co-related to job satisfaction.
H1: Attrition is not directly co-related to job satisfaction.

HR analytics is defined as "a method for evaluating and understanding the causal relationship between HR practices and overall performance results of organizations (such as client satisfaction, sales, or profit)". It also helps in supplying authentic and reliable foundations for human capital choices for integrating enterprise strategy and performance. This integration could be achieved using statistical strategies and experimental tactics based on various metrics of efficiency, effectiveness, and competitive advantage. Therefore, this section details the relevance of using data analytics for the HR functions. The prime advantages reported by the past researchers are as follows:

(a) HR analytics is affordable and provides the data-based evidence for both strategic and operational decisions.
(b) The enhanced technological advances also provided the access to big data related to human resource which makes use of HR analytics relevant.

(c) It also helps in tackling the global talent warfare by enhancing both talent acquisition and retention.

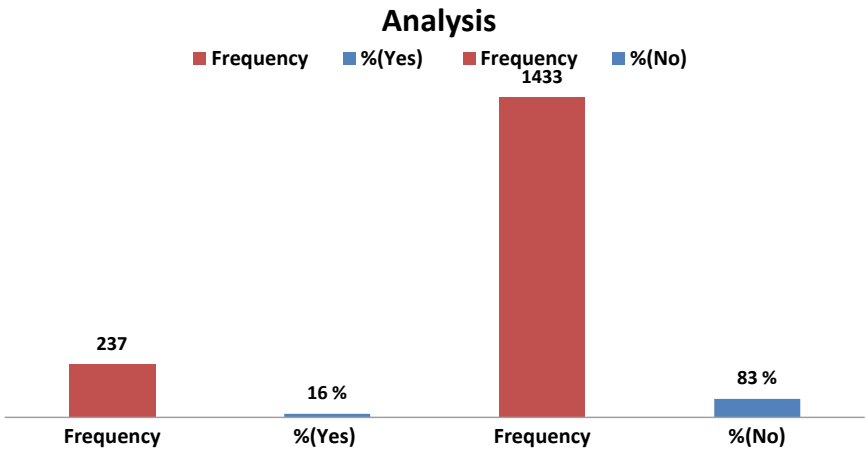
Further, there are three levels prominently used for analysis :

I. **Descriptive**: Descriptive HR analytics reveals and describes relationships. This type of analysis compares the present scenario with historical data patterns in order to generate insights. It consists of dashboards and score-cards; team of workers segmentation; statistics mining for main patterns; and periodic reports. Traditional HR metrics are mostly successful metrics (turnover rate, time to fill, value of hiring, volume employed and trained, etc.). The predominant center of attention is on cost reduction and process improvement.
II. **Predictive**: Predictive analysis covers an array of techniques (statistics, modeling, information mining) that use present day and historical facts to make predictions about the future policies, strategies, and practices. It deals with identifying probabilities and dealing with the elements of risk. The human capital (people) leverage intangible belongings to enhance employer performance, and hence the impact of various factors on human capital need to be explored. The predications related to such explorations may result in typical profiles and trends. These profiles and trends may be used for strategic human resources management decisions.
III. **Prescriptive**: Prescriptive analytics is used to analyze complicated facts to predict outcomes, provide resolution options, and exhibit desire impacts. The technique deals with reporting of HR metrics and prescriptive modeling of commercial organization practices. It involves in understanding the impact of knowledge of investments on the bottom line. Such kind of scenarios is rare in case of HR domain.

Taking the aforementioned logic into consideration this study used predictive approach in order to understand the relationship between attrition and job satisfactions in a better way. This study used a sizeable data set related to HR domain (fictional) created by IBM scientists.|The data set consist of 1470 data points and 35 variables. This data is specifically used for analysis of various HR-related attributes. The statistical tools like Statistical Package for Social Sciences (SPSS) and Microsoft Excel were used for analysis.

3 Data Analysis

The prime objective of this study was to identify the predictive indicators for identifying the typical profile of the employees who have left the organization. In order to develop these indicators, the data was divided into the employees left the organization (237) and still working in the organization (1433). This classification is presented as a Graph 1.



Graph 1 Classification of the employees

From the analyzed data it can be concluded that out of the total 1470 employees 16% employees have already left the organization. This number is 237 employees. The typical characteristics or behavioral pattern or the predictive indicators were identified on the basis of the employees who have left the organization.

Further, in order to identify the predictive indicators, the response or the characteristics of employees who left the organization were taken into consideration. The frequency analysis conducted for determining the predictive indicators is presented in Tables 1 and 2. Table 1 details about the factors related to an individual whereas Table 2 represents the prominent organizational factors identified by the researcher.

4 Individual Factors

The analysis presented in Table 1 indicates that the tendency to leave the organization is significantly higher among males, trainees, and unmarried employees. Further, it is moderate among sales representatives and people having life sciences as educational background. It is also there among lab technicians and people belonging to age group 31–35.

Similar kind of analysis is done for the organizational factors.

5 Organizational Factors

Researcher had observed that out of 20 factors, certain factors significantly influence the rate of attrition which are performance rating, years in current role, total years at company, years since last promotion, traveling in job, stock level, and percentage hike in salary. Further, the attrition may occur because of work–life

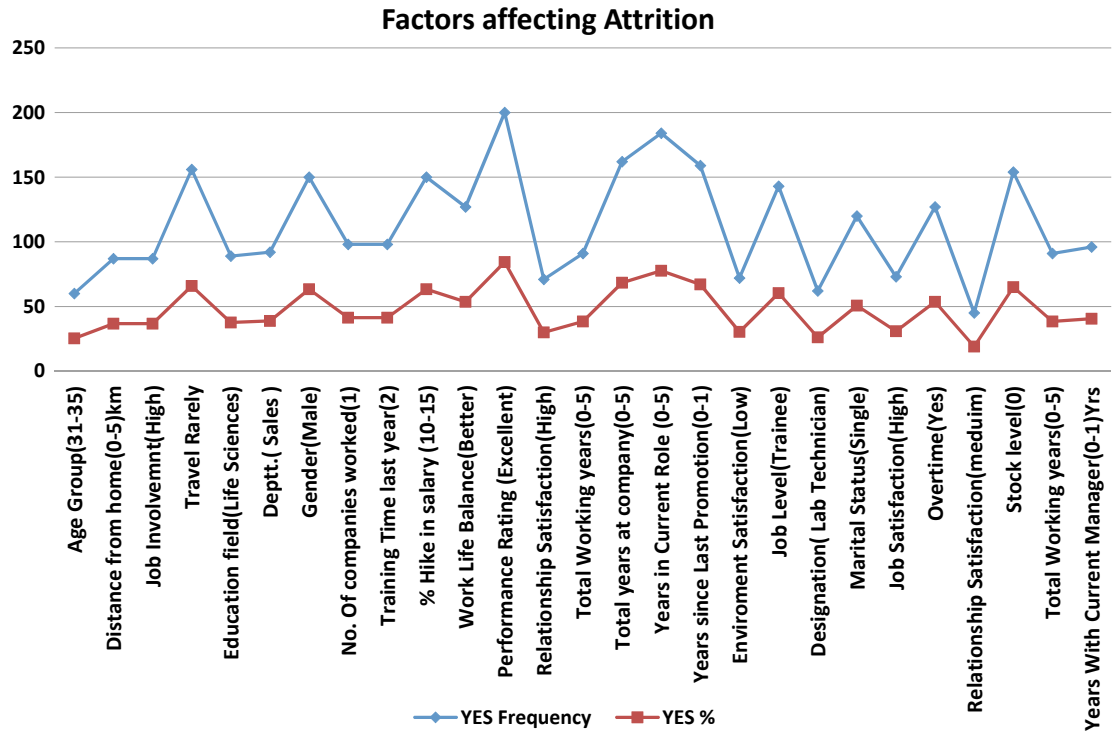
Table 1 Individual factors

Category	Frequency	Percentage
Gender (Male)	150	63
Job level (Trainee)	143	60
Marital status (Unmarried)	120	51
Department (Sales)	92	39
Education field (Life sciences)	89	38
Designation (Lab technician)	62	26
Age group (31–35)	60	25

balance, overtime, no. of companies worked, training time last year, years with current manager, total working experience, total working years, distance from home, job involvement, job satisfaction, environment satisfaction, and relationship satisfaction. Tables 1 and 2 are graphically represented as Graph 2.

Table 2 Organizational factors

Category	Frequency	Percentage
Performance rating (Excellent)	200	84
Years in current role (0–5)	184	78
Total years at company (0–5)	162	68
Years since last promotion (0–1)	159	67
Travelling in job (Rarely)	156	66
Stock level (0)	154	65
Percentage hike in salary (10–15)	150	63
Work life balance (Better)	127	54
Overtime (Yes)	127	54
No. Of companies worked (1)	98	41
Training time last year (2)	98	41
Years with current manager(0–1) yrs	96	41
Total working experience (0–5)	91	38
Total working years (0–5)	91	38
Distance from home (0–5) km	87	37
Job involvement (High)	87	37
Job satisfaction (High)	73	31
Environment satisfaction (Low)	72	30
Relationship satisfaction (High)	71	30
Relationship satisfaction (Medium)	45	19



Graph 2 Factors affecting attrition

Table 3 Results of hypothesis testing

	Attrition	Job Satisfaction
Attrition	1	
Job satisfaction	0.903481126	1

6 Hypothesis Testing

In order to test the hypotheses, correlation analysis is used. The results of hypothesis testing indicate that attrition and job satisfaction are strongly co-related. (Table 3)

Hence, from the above co-relation matrix, we can conclude that hence Ho Hypothesis is proved.

7 Multiple Regressions for Predictive Analysis

After confirming the impact of job satisfaction on attrition, the next step was to identify the parameters that impact the job satisfaction among the employees who left the organization. In order to identify these factors, the job satisfaction was

Table 4 Details of regression analysis

Regression statistics parameter	Details
Multiple R	0.339554231
R Square	0.115297076
Adjusted R Square	0.055249366
Standard error	1.086733162
Observations	237

Table 5 Results of ANOVA

ANOVA	*df*	SS	MS	*F*	Significance *F*
Regression	15	34.01409675	2.2676065	1.920091	0.022485206
Residual	221	260.9985615	1.180989		
Total	236	295.0126582			

treated as a dependent variable and 33 factors related to employees who left the organization as independent variables. The results of multiple regression are presented in Table 4.

In order to check the predictability of the factor, the multiple regression matrix was used. The threshold value for treading a factor as a predictive < 0.15 (Table 5).

From the above data, it is observed that four factors who have values less than or equal to 0.15 can be used to predict the job satisfaction. These factors are over time, stock option level, total working years, years at company, years with current manager. These predictors can be used for predicting the existing employee's tendency to leave the organization. Further, this will help the HR manager to design better strategies. The probability predications are presented in Table 6.

8 Probability Prediction for Attrition

From the above Table 7, it can be predicted that the employees who have been working for 0–5 years may leave the organization for growth, as the similar trend was observed with the employees who left the organization. This analogy can be applied for other factors like overtime years with current manager and total working experience.

Table 6 Predictors of job satisfactions

Particulars	Coefficients	Standard Error	*t* Stat	*P* value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	3.377592373	0.860301972	3.9260544	0.000115	1.682146867	5.0730379	1.682146867	5.073037879
Monthly Income	−3.08897E−05	3.04808E−05	−1.013415	0.31197	−9.09599E−05	2.918E−05	−9.096E−05	2.91805E−05
Monthly Rate	2.66044E−05	1.02223E−05	2.6025719	0.009879	6.45864E−06	4.675E−05	6.45864E−06	4.67501E−05
Num Companies Worked	0.004138138	0.030422756	0.1360211	0.891928	−0.0558177	0.064094	−0.0558177	0.064093975
Over Time	−0.258561383	0.1444734	−1.789682	0.074874	−0.543283242	0.0261605	−0.54328324	0.026160476
Percent Salary Hike	−0.036399008	0.031360677	−1.160658	0.247033	−0.098203258	0.0254052	−0.09820326	0.025405243
Performance Rating	−0.081905796	0.328406837	−0.249403	0.80328	−0.729115631	0.565304	−0.72911563	0.565304038
Relationship Satisfaction	0.053887991	0.064412541	0.8366071	0.403717	−0.073053427	0.1808294	−0.07305343	0.18082941
Stock Option Level	0.121833207	0.084938364	1.434372	0.15288	−0.045559607	0.289226	−0.04555961	0.289226021
Total Working Years	−0.028471292	0.019877602	−1.43233	0.153462	−0.0676452	0.0107026	−0.0676452	0.010702615
Training Times Last Year	−0.023684322	0.058022296	−0.408193	0.683527	−0.138032125	0.0906635	−0.13803213	0.090663482
Work Life Balance	−0.049626257	0.08927681	−0.55587	0.578862	−0.225569091	0.1263166	−0.22556909	0.126316577
Years At Company	0.070074927	0.028812772	2.4320787	0.015808	0.013291977	0.1268579	0.013291977	0.126857877
Years In Current Role	0.007748925	0.043790278	0.1769554	0.859706	−0.078551041	0.0940489	−0.07855104	0.09404889
Years Since Last Promotion	0.02579916	0.034231713	0.7536625	0.451854	−0.041663203	0.0932615	−0.0416632	0.093261522
Years With Current Manager	−0.063811836	0.040742793	−1.566212	0.11873	−0.144105951	0.0164823	−0.14410595	0.016482279

Table 7 Probability prediction for attrition

Predictive Analysis	Existing employees	
Parameters	Freq.	%
Years at company(0–5) yrs	614	50
Stock option (0)	477	39
Over time(Yes)	289	23
Years with current manager(0–1) yrs	243	20
Total working experience (0–5)	225	18

9 Conclusion

Researchers have observed that there is a strong relationship between the job satisfaction and attrition and failing of which it may lead to attrition. The employees will be satisfied with the organization when the company offers an interesting job, good policies, benefits, compensation. Hence, the management need to have an appropriate HR mechanism to deal with this.

The practice of overtime may be perceived as the source of additional income and a factor that may lead to physical stress. Further, it may exhaust the employees and may inversely affect the efficiency. From the data, it is ascertained that 54% of employees have left the organization due to overtime as one of the factors and it could be predicted from the analysis that 23% of employees may leave the organization.

Employees own a share in the company's success with stock option plans offered them a sense of belongingness. Holding of a stock automatically makes employees a contributor to the profit and loss of the company. The researchers have noticed that 65% of the employees who resigned were not owning stocks of the organization. This tendency may influence 39% of existing staff.

It is observed that the company recruits fresh people and offers them experience, skill, knowledge, and train them for their future. This staff expects higher salaries and growth opportunities. Based on the analysis it may be considered that employee between 0 to 5 years of experience may leave the organization. This demands better policy-making. The researchers observed that the role of supervisor/reporting authority is one of the factors that need to be considered.

Hence, as a result of this study, the researcher proposes to use predictive analysis techniques for better HR policy and strategy making.

References

1. Muscalu E, Serban A (2014) HR analytics for strategic human resource management. In: Proceeding of 8th international management conference
2. Momin, W.Y.M., Taruna, : HR analytics transforming human resource management. Int J Appl Res **9**(1), 2349–5869 (2015)
3. Sujeet, Mishra, Dev, Lama, Yogesh, Pal: Human resources predictive analytics (HRPA) For HR management in organizations. Int J Sci Technol Res **5**(5), 2277–8616 (2016)
4. Bindu, K.: A review of poor analytical skills. Int J Innov Res Technol **3**(1), 2349–6002 (2016)
5. Madhavi Lakshmi, P., Siva Pratap, P.: HR analytics- a strategic approach to HR effectiveness. Int J Hum Resour Manag Res **6**(3), 2249–7986 (2016)
6. Globcon Technologies Private Ltd (2016) Redefining HR through data analytics. Article published on www.analyticsindiamag.com
7. Rajbhar, A.K., Khan, T., Puskar, S.: A study on HR analytics transforming human resource management. J Invest Manag **6**(4), 2328–7721 (2017)
8. Shweta, T.: Analytics application in human resource context. Int J Sci Eng Res **6**(1), 2347–3878 (2018)

A Study of Business Performance Management in Special Reference to Automobile Industry

Gurinder Singh, Smiti Kashyap, Kanika Singh Tomar and Vikas Garg

Abstract In contemporary times, the automobile is one of the well-paid industries in the Indian market with an annual growth rate of 7.64% in the passenger-car market. The increasing disposable income of the people along with the ever-growing financial sector has led to this expendable growth. In accordance, there has been an increase in sales of passenger cars from 13.35% in –July 2018. This oligopoly market has fierce competition due to new entrants into the Indian markets. Thus, there emerges a need to grasp the knowledge to understand the ever-growing needs of the customers and dynamism in the technology-driven market. Every organization seeks to study consumer buying behavior and measure its business performance by analyzing customer perception toward the product. The understanding of customer's perception is an ongoing process to survive the cut throat competition. Taking this into consideration, the study provides insights about several attributes which drive a consumer behavior toward buying a product of a brand. The study has used theories and exploratory research design along with analytical tools to identify the major attributes of consumer buying behavior toward sedan cars within Delhi/NCR among high-end consumers. This knowledge helps the car manufacturers in market segmentation which enables the organization to plan the market strategies toward consumer retention and product upgradation.

Keywords Business performance management · Consumer buying behavior · Market segmentation · Factor analysis · Brand reputation · Purchase service · Physical appearance · Automotive techniques

G. Singh · S. Kashyap · K. S. Tomar (✉) · V. Garg
Amity University, Noida, Uttar Pradesh 201308, India
e-mail: kstomar@gn.amity.edu

G. Singh
e-mail: gsingh@amity.edu

S. Kashyap
e-mail: skashyap@gn.amity.edu

V. Garg
e-mail: vgarg@gn.amity.edu

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_6

1 Introduction

In India, the automobile is an industry which plays a major role in the growth of the Indian economy. The industry encompasses namely, passenger cars, two-wheelers, three-wheelers, commercial vehicles, multi-utility vehicles, and its components. The share of the Indian motorcycle manufacturers is the largest in the world, followed by other two-wheelers and tractor manufacturers. The Indian Commercial Vehicle manufacturers are the fifth largest across the global market and fourth largest in Asia. India has strength in manufacturing low-cost, fuel-efficient cars of various automobile companies namely, Maruti Suzuki, Volkswagen, Toyota, Hyundai, and Nissan.

Looking at the ever-growing share of the automobile industry in the growth, the government of India focuses on Research & Development to provide the infrastructure and world-class automotive testing to the industry. The paper throws light on the several determinants of consumer buying behavior of passenger cars which can be utilized by the industry for market segmentation, advertisement strategies, product differentiation in terms of technology upgradation, safety and user friendliness, and sales maximization.

1.1 Passenger Vehicles

A passenger vehicle is a four-wheeler vehicle used as a passenger carriage in which up to nine people including a driver can be seated. India is leading in the car market with a 9% growth rate. Japan is the only market which could come closer to this growth rate. China which is the largest market of passenger vehicles had registered a decline in the growth rate by 2.59% and the US market shrunk by 10%. The paper studies the consumer behavior of sedan cars in India. A sedan car is a 3-box configuration passenger-car with A: B: C pillars along with compartments for other things such as the engine of the car, its passenger, and cargo. The best sedan cars in India are Maruti Dzire, Honda Amaze, Honda City, Hyundai Verna, Maruti Ciaz, Toyota Yaris, Toyota Etios, Hyundai Xcent, Tata Tigor, C Class 2018, Sonata, Civic, A6 2019, A Class Sedan, etc.

1.2 Consumer Buying Behavior

The decision-making of a consumer, during a purchase is necessary to be understood to capture the market and tap the potential gains. The understanding is required in terms of what the dynamics of consumer preferences are.

Post this understanding, the automobile industry segments and works on the marketing strategies to improve business performance. In theory, there is a six-stages process of the consumer demand behavior which comprises:

- Exploratory research to understand the problem
- Assimilating the information for the research
- Evaluating the solutions and alternatives
- Decision-making
- The final purchase
- Post-purchase experience.

The decision-making of the consumer buying behavior can be studied at a psychological level, physiological level, financial level, and spatial level.

2 Literature Review

In the automobile industry, the factors which impact the buying pattern of new cars and second-hand cars have been studied extensively. It has been observed that the factors affecting the buying behavior of second-hand cars are very different from first-hand cars [20].

The purpose of studying the consumer buying behavior for new car entrants is to fill the gaps between customer expectations and the present products. This helps to depict the total quality management of the cars to tap the existing and potential buyers [12].

The question for managers in the service industries regarding consumer loyalty has been studied based on online or offline consumers. They have depicted the correspondence among customer satisfaction and loyalist in an online and offline environment. It shows that the consumer satisfaction level among online and offline customers is equal, however, the loyalty of the customer is more toward online service provider [16].

A pattern to study the pricing and advertisement expenditure using time series data has been made in 3D Innovation Diffusion Model with empirical evidence of sales to understand the growth pattern of sales [3].

The consumers have a bend toward sporting cars and speed cars as they consider it as the good of distinction. There are several constructs to these concepts. The variables like R&D by the technical fraternity, disposable income, and liberalization in terms of foreign exchange, high-end models and integration in automobiles pave way to the dynamic consumer behavior in the automobile industry as observed today [17].

The car segments in terms of easy finance availability of rural and urban sectors have been studied to understand the bend of consumer based on brand loyalty, which helps the existing companies to survive when new entrants come in the market. All market segments within the Indian market have been considered such as

value-for-money, safety and driving comforts, and brand image impacts on sales of passenger cars [5].

The logistic regression modeling has also been used to depict the protectionism adopted by the government of India to help the domestic car manufacturers to grow in the competitive world. The domestic car manufacturers use product differentiations, innovative products, and quality and reliability as the attributes to capture the market share. The companies such as Honda, Toyota, General Motors, Ford, and Hyundai used these strategies prior to liberalization and foreign direct transformation in the Indian market [14].

The increase in disposable income at the rate of 25% per annum and the changes in the sociocultural environment are the factors affecting consumer behavior in terms of luxury cars. Now, even middle-class people like to purchase luxury cars. So, the study shows how luxury cars now occupy the place in the daily life of the people within the Indian Economy [18].

Automotive Technology and Human Factor Research related to the past, present, and future is primarily driven in terms of the features of car control, display for the driver, driver workspace, driver's conditions, etc. [1].

The aim of this paper is to find the significant factors which influence the consumer buying behavior of sedan cars. Further, the paper provides recommendations to the automobile industry to upgrade their products, market segmentation, and advertisement strategies to maximize the sales and hence, capture the market.

3 Objectives

1. To evaluate the measures to improve the business performance of the sedan car industry to compete in the Indian metropolitan markets.
2. To determine the agents affecting the consumer buying behavior of sedan cars in Delhi/NCR.
3. To segment the market of sedan cars with reference to the consumer preferences.

4 Data Collection

This paper aims to evaluate the measures to improve business performance in the automobile industry. The business performance management is studied based on different perspectives namely, learning-growth perspective, financial perspective, customer perspective, and internalized business process [2].

This paper focuses on the learning and growth aspect of the automobile industry. The market segmentation is significant for the existing firms and new entrants to compete in the market [7]. With the review of the literature and pilot surveys among

the consumers of Delhi/NCR (30 consumers for the pilot survey were used), we have used the following variables which impact the consumer buying behavior in the sedan car market to do the market segmentation:

The demographic variables under study are age, sex, household income, and occupation.

Other factors:

4.1 Brand Reputation

This attribute influences consumer buying behavior. We have studied brand reputation using the following variables:

Brand image, its name, consistent and affordable pricing for high-end customers, and resale value are important in the decision-making process of the consumer [9].

4.1.1 Brand Name

It is the name given by the company to the product or a range of products. A consumer may associate themselves with a brand name or may build trust toward a specific brand name.

4.1.2 Price

It is the manufacturer suggested retail price of the car which influences the purchase of the product by the consumers given their income levels and preferences. This is also known as "Sticker Price".

4.1.3 Resale Value

It is a factor which is based on several attributes of the car which a customer keeps in mind while purchasing it. The customer likes to purchase a car with higher resale value.

4.2 Warranty

Sedan cars in India are sold by several companies such as Hyundai, Honda, Volkswagen, Maruti Suzuki, etc. The cars in the current era are more reliable and in cases of component failures, the companies provide for the replacement under the warranty period which is either in months/years or the kilometers driven from the date of delivery. A few companies even provide for extended warranty under the

name of "Forever yours" such as Maruti Suzuki. We have studied the warranty using the following variables:

4.2.1 Availability of Spare Parts

Availability of spare parts increases the longevity of the car and makes it economical.

4.2.2 Vehicle Availability

Availability of vehicle for use in the stipulated time of need.

4.3 Physical Appearance

The looks of a car play an important role in purchasing a car. We have studied physical appearance using the following variables:

4.3.1 Design and Exteriors

The design, shape, and surface are built by car manufacturers using upgraded technology which attracts the customer. The body of the car impacts the speed and safety of the car.

4.3.2 Comfort

It is a primary factor which influences consumer buying behavior. A comfortable vehicle with leg space, headroom, automatic starring are the definitive factors which a consumer investigates before buying the product.

4.4 Features

Apart from the looks, there are several other features which are kept into consideration while purchasing a car. We have used the following variables:

4.4.1 Interior Look

The well-defined body, neatly-cut lamp, installed lights, style, inside cabin, gearbox design, and several other interior designs are investigated by a consumer while purchasing a car.

4.4.2 Boot Space

The boot space ranges from 235 to 510 L in a sedan car which signifies how spacious a car is.

4.5 Technical Features

All the consumers look for an upgraded product along with user friendliness. We have used the following variables:

4.5.1 Technology

The sedan cars with high-tech features such as gadgets installed in the car, the automatic gear box, apple care play, compact hatch back, gas-electric hybrid technology, etc., are considered by the consumers prior to buying a car.

4.5.2 Engine

The cubic centimeters of the cylinder of the car engine are kept in mind by the consumer prior to purchase.

4.6 Value Addition

The additional services are provided by the car distribution channel to add value to the product. We have used the following variables:

4.6.1 Waiting Time

This is a time gap between the booking and delivery of the car.

4.6.2 Discount

The reduction in the price of the car makes the purchase attractive.

4.6.3 Safety

Several features in a car like airbag, crash rates, and platform, are important to a consumer and for auto-insurances.

4.6.4 Value for Money

If the utility received from the car is not in equilibrium with the pricing, the consumer tends to restrict the purchase.

4.7 Performance

The utility of the sedan car to a consumer can be studied using the following variables:

4.7.1 Mileage

The distance in kilometers covered by a car per liter of fuel is considered before purchase.

4.7.2 Fuel Economy

The fuel-efficiency of a car is amajor performance parameter.

5 Research Methodology

This is an exploratory research where we try to identify the major factors which determine the consumer buying behavior of sedan cars in Delhi/NCR. For the study, we have used high-end consumers under the income bracket of 7 Lac and above (per annum). The primary data was collected using the survey method [8].

We have built a questionnaire under several heads as per the variables. The responses to the questionnaire were collected from the consumers who were the potential buyers of sedan cars.

We have divided the Delhi/NCR majorly into four regions namely, north, south, west, and east to represent the consumers (with income of 7 Lac and above annually) of sedan cars. The information regarding the consumers was collected using the data available at the showrooms. We have used the sample size of 306 consumers from all the four regions to truly represent the target population and performed personalized interviews to collect the information.

6 Data Analysis

To identify the significant factors which affect consumer buying behavior, we have collected the data using the questionnaire method. Every question was analyzed to understand consumer preferences using bar graphs, pie charts, tables, and statistical tools.

6.1 Demographic Distribution of Consumers

6.1.1 On the Basis of Gender

The data collected shows that out of the sample, 20% of the consumers of sedan cars are females and 80% of the consumers are males (refer to Fig. 1). This interpretation helps the automobile industry to focus on the male counterpart of the society while designing the marketing strategies and features of the product.

The data collected shows that out of the sample size, 20% of the consumers are females and 80% are males.

6.1.2 On the Basis of Age

During the data analysis, it was found that sedan cars are purchased by consumers of different age groups. The maximum percentage of users, i.e., 50% of the users lie

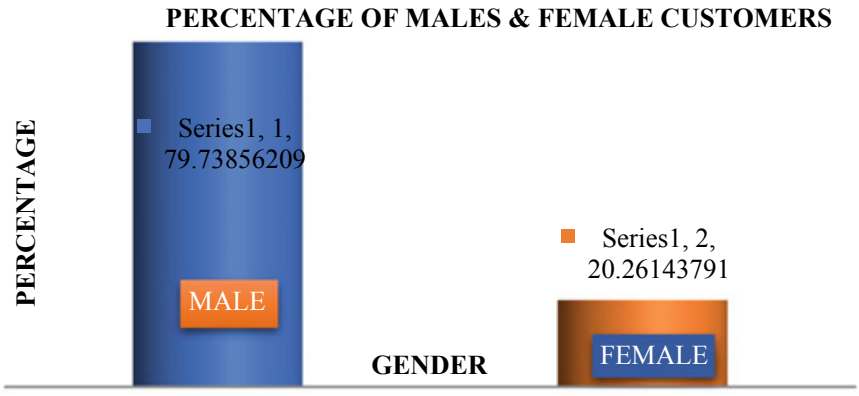


Fig. 1 Percentage of users of among males and females

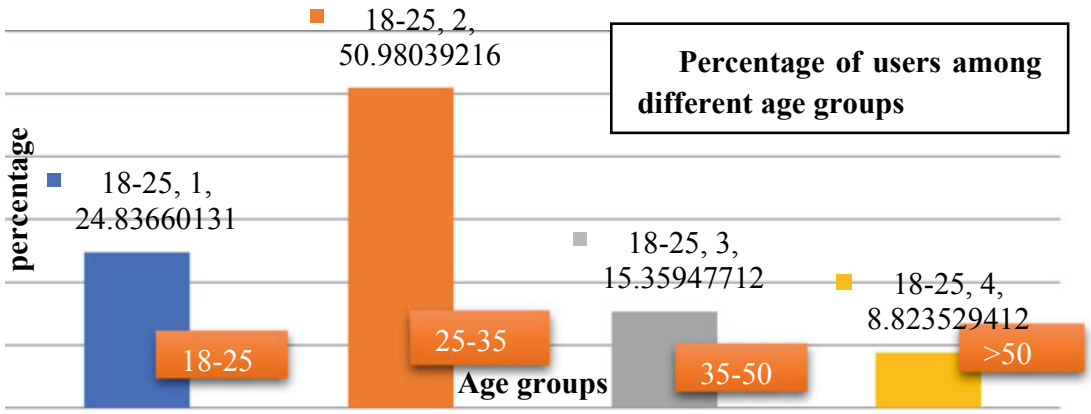


Fig. 2 Percentage of users among different age groups

in the age group of 25–35 years followed by the age group of 18–25 years constituting nearly 25% of the users (refer to Fig. 2). The rest 15% fall in the age group of 35–50 years and above. So, the companies should target the advertisements on the age group of 25–35 years to increase sales.

The consumers between 25–35 years are the people who demand 50% of the sedan cars in the market. The people between 35–50 years demand 15% of sedan cars in the market and 18–24 years demand 25% of the sedan cars in the market.

6.1.3 On the Basis of Occupation

The results show that the maximum users of sedan cars fall in the business class in terms of occupation constituting nearly 63%. The service class constitutes approximately 20% of the total users followed by students and others. (refer to Fig. 3). The automobile firms must focus on their marketing strategy on the business class.

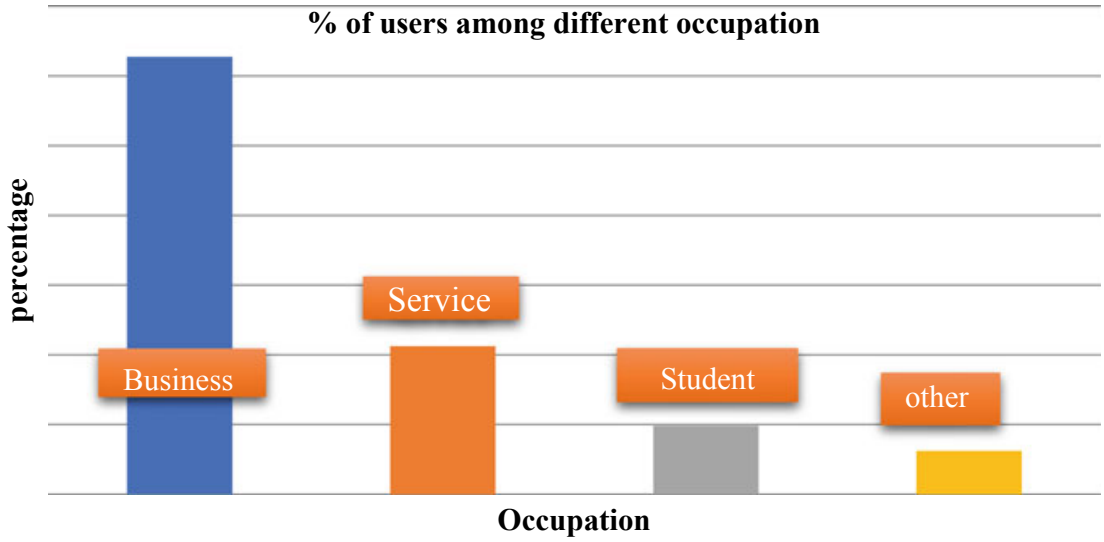


Fig. 3 Percentage of users among different occupation

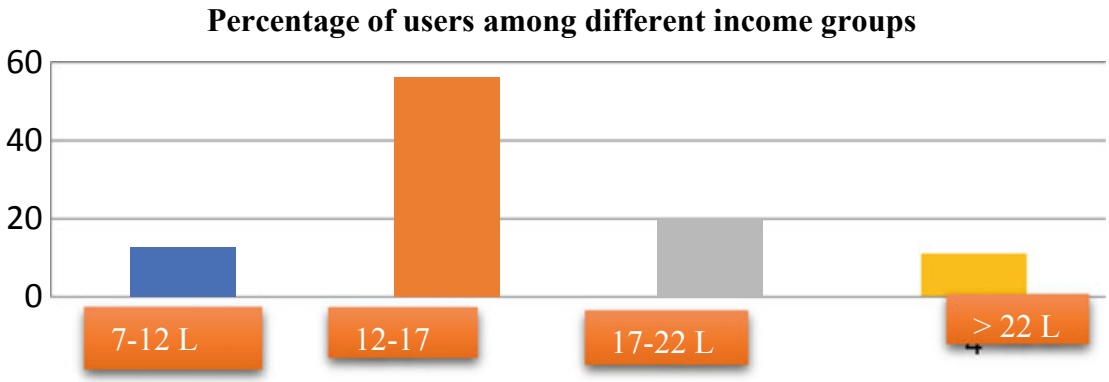


Fig. 4 Percentage of users among different income groups

Among the buyers, the maximum people are businessmen—approximately 63%. Nearly 20% are servicemen, 10% are students, and the rest 8% falls in the "others" category which includes self-employed people.

6.1.4 On the Basis of Income Groups

The data collected helps us to find that the maximum users lie in the income group of 12 to 17 Lacs constituting of up to 57%. 20% of the users fall in the income group of 17–22 Lacs and 12% of the users fall in the income group of 22 Lac and above and 11% fall in the income group of 7–12 Lacs (refer to Fig. 4). The companies' pricing, products, and marketing strategies must focus on users falling under the age group of 12 to 17 lacs.

According to my study, 57% of the buyers are of the income group 12–17 lacs per annum and 12% people of income above 22 lacs buy sedan cars, as people of this group would rather go for sedan cars.

6.2 Motivational Factors of Consumers

6.2.1 On the Basis of Daily Drive

The results show that on an average 30% of the consumers in Delhi/NCR drive under 50 kms, 25% of the consumers drive up to 50–80 kms, 34% the consumers drive up to 80–100 kms and rest 11% of the consumers drive above 100 kms (refer to Fig. 5). As on an average, maximum consumers drive between 80–100 kms, they would prefer more economical cars on roads such as diesel engine, good mileage along with luxury and comfort.

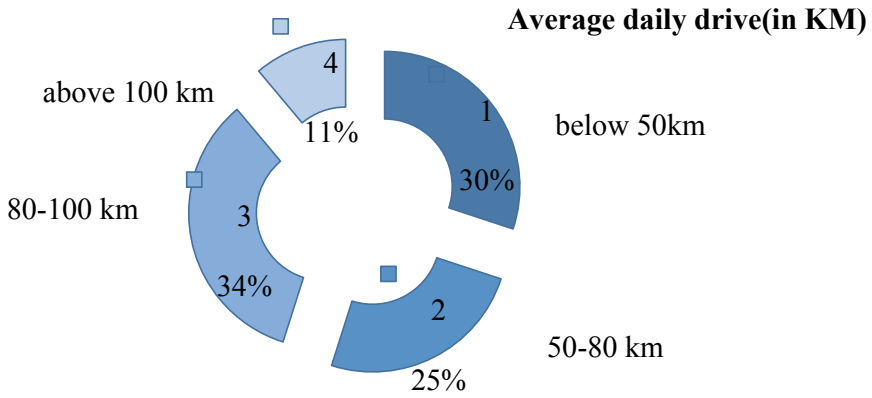


Fig. 5 The average daily drive of the consumers

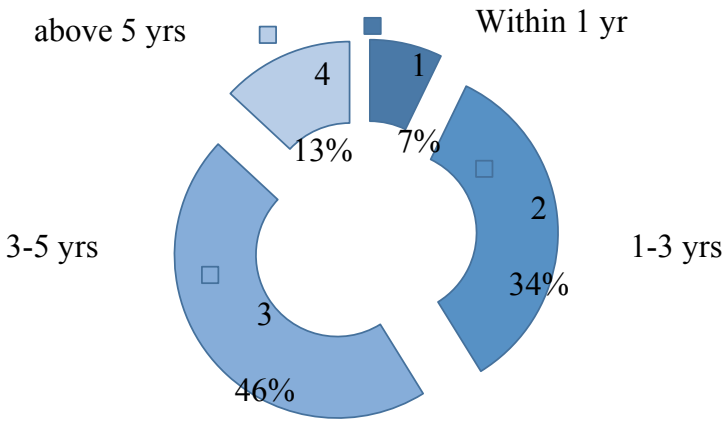
According to my survey, 34% of the buyers include people who drive 80–100 kms daily on an average and therefore they may be interested in a diesel engine sedan class car which would give them good mileage, luxury, and comfort.

6.2.2 On the Basis of Time Span Between Car Switch

The results show that on an average out of the total consumers of sedan cars in Delhi/NCR, 7% change their car within a time span of 1 year, 34% change it in 1 to 3 years, 46% change it in 3 to 5 years, and rest 13% change it in a span of above 5 years (refer to Fig. 6). The companies can use this data and intimate their consumers about the new releases within these time spans to retain the consumers.

According to my study, 46% people purchase/change their cars in 3–5 years. By keeping a track of such customers, who change their cars in every 3–5 years, the companies could send them invitations regarding their new releases and retain their customers.

Fig. 6 Time span in which the consumers change their cars



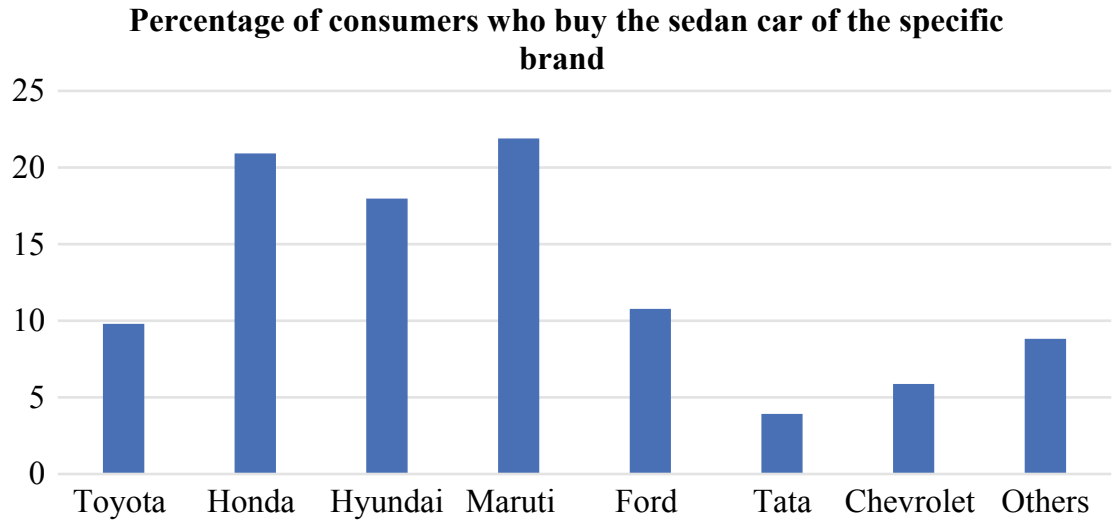


Fig. 7 The different brand sedan cars which consumers purchase

6.3 Preference of Consumers Toward Different Brands

The results show that in Delhi/NCR, 22% of the consumers buy Maruti sedan cars, 20% of the consumers buy Honda sedan cars, 18% of the consumers buy Hyundai sedan cars, 10% buy Ford, and the rest of the consumers prefer other brands like Toyota, Chevrolet, Tata, etc. (refer to Fig. 7). This shows the cut throat competition among the oligopolistic automobile industry.

According to the study, the consumers buy (cars) of other brands such as Honda, Hyundai, Maruti, and Ford which constitute 20, 18, 22, and 10%, respectively.

6.4 Consumer Preferences Based on Features of Sedan Cars

The study has used the 5-point Likert scale to analyze consumer preferences while buying a sedan car. The scale: 1 as Excellent, correspondingly 2 as Very-Good, followed by 3 as Satisfactory, 4 as Non-Satisfactory, and 5 as Poor. The study observes that when consumers purchase a sedan car, 20% is based on the comfort level of the car, 17% on the basis of engine strength, and 14% on the basis of luxury. Table 1 shows the weightage of different features while purchasing a sedan car.

Table 1 Frequency of consumer preference

Features	Frequency	Weightage (Percentage)
Comfort	64	20.91503
Luxury	43	14.05229
Engine	55	17.97386
Sunroof	36	11.76471
Cruise control	18	5.882353
Premium upholstery	24	7.843137
Push entry passive start (PEPS)	21	6.862745
Good mileage	39	12.7451
Others	6	1.960784

6.5 Factor Analysis

We have used the factor analysis technique using SPSS software. The factor analysis technique is used to extract the most significant variables from the list of several variables which influence the dependent variable. This tool calculates the maximum common variance of all variables to find the common score which determines the magnitude and direction of the impact of the independent variable on the dependent variable.

Factor analysis is a multivariate technique which is used to identify the underlying factors and then, eliminate the redundant variables [11].

We have calculated the eigenvalue which is used to consolidate the variance, characteristic roots, and latent values. Additionally, the scree plot of eigenvalue is used to find the significant variables of the study. All the variables on the scree plot before it becomes flat are taken into consideration [13]. An eigenvalue of either 1 or more than 1 is considered a good explanatory feature.

6.5.1 Descriptive Statistics

The data has been first analyzed to measure the central tendencies and dispersion using tools such as mean and standard deviation as depicted in Table 2. It is used to summarize the data set prior to applying factor analysis. The mean shows the influence of each factor on the consumer buying behavior of sedan cars. The weights which are closer to 0, hold no relevance while the larger the weights, larger is the influence on the dependent variable. The standard deviation shows the variation of these factors from the central tendencies. This data set is of total N (306) observations.

Table 2 Descriptive statistics

	Mean	Std. Deviation	Analysis N	Missing N
Technology	2.09	0.943	306	0
Brand_name	2.10	1.062	306	0
Price	3.67	1.133	306	0
Interior	2.32	0.659	306	0
Design_and_exterior_looks	2.42	0.630	300	6
Comfort_and_convenience	2.19	0.618	306	0
Safety_and_security	1.49	0.830	306	0
Engine	2.63	0.821	306	0
Mileage	1.84	0.733	306	0
Boot_space	1.51	1.031	306	0
Value_for_money	1.53	0.623	306	0
Warranty	3.18	0.747	306	0
After_sale_services	2.29	0.719	306	0
Resale_value	5.54	1.116	306	0
Availability_of_spare_parts	2.14	1.066	306	0
Discounts_or_special_offers	2.19	0.988	306	0
Fuel_economy	2.14	0.737	306	0
Vehical_availability	2.53	0.874	306	0
Waiting_time	2.63	0.762	306	0

6.5.2 KMO and Bartlett's Test

The KMO and Bartlett's Tests were performed to check the adequacy of the data set for factor analysis. The value of Kaiser test lying between 0.5 and above can be considered as middling to perform factor analysis and hence, acceptable. Additionally, Bartlett's test is used to find how useful it is to apply the factor analysis on the data set. If the p value is less than 0.05, then it is considered suitable for further operations. The results in Table 3 show that the data set is suitable for factor analysis.

Table 3 KMO and Bartlett's test

Kaiser–Meyer–Olkin measure of sampling adequacy.		0.585
Bartlett's test of sphericity	Approx. Chi-square	1944.462
	df	171
	Sig.	0.000

Table 4 Communalities

	Initial	Extraction
Technology	1.000	0.619
Brand_name	1.000	0.592
Price	1.000	0.775
Interior	1.000	0.727
Design_and_exterior_looks	1.000	0.652
Comfort_and_convenience	1.000	0.773
Safety_and_security	1.000	0.635
Engine	1.000	0.649
Mileage	1.000	0.800
Boot_space	1.000	0.591
Value_for_money	1.000	0.695
Warranty	1.000	0.726
After_sale_services	1.000	0.720
Resale_value	1.000	0.671
Availability_of_spare_parts	1.000	0.688
Discounts_or_special_offers	1.000	0.701
Fuel_economy	1.000	0.772
Vehical_availability	1.000	0.666
Waiting_time	1.000	0.826

Extraction method: Principal component analysis

6.5.3 Communalities

We have calculated communalities to find the proportion of the variance of the variable that can be explained by each factor. The Initial values indicate the estimates of variance on the account of all factors; it is always 1 for principal component extraction. The extraction value is the estimate of the variance of the variable by all components. A high extraction value indicates that all the variables are represented well. The corresponding results are shown in Table 4.

6.5.4 Total Variance Explained

The purpose is to extract the variables which explain the maximum variance of the dependent variable. To extract the total variance explained by the factors, we have run the factor analysis. The number of factors used for factor analysis is 19. The initial eigenvalue depicts the variance; the total values show the variance in each factor in descending order; the percentage of variance shows the total variance accounted by each factor and the cumulative percentage shows the variance accounted by the current and all the preceding factors. The result in Table 5 shows that the first seven variables explain approximately 70% of the total variance.

Table 5 Total variance explained

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	2.884	15.176	15.176	2.884	15.176	15.176	2.446	12.875	12.875
2	2.707	14.247	29.423	2.707	14.247	29.423	2.156	11.346	24.220
3	2.093	11.016	40.439	2.093	11.016	40.439	2.009	10.573	34.793
4	1.553	8.172	48.611	1.553	8.172	48.611	1.926	10.136	44.929
5	1.418	7.462	56.073	1.418	7.462	56.073	1.661	8.743	53.672
6	1.332	7.011	63.084	1.332	7.011	63.084	1.653	8.700	62.372
7	1.292	6.802	69.886	1.292	6.802	69.886	1.428	7.514	69.886
8	0.889	4.679	74.565						
9	0.821	4.321	78.886						
10	0.678	3.566	82.452						
11	0.630	3.314	85.766						
12	0.551	2.897	88.663						
13	0.523	2.753	91.416						
14	0.433	2.281	93.697						
15	0.408	2.148	95.845						
16	0.279	1.470	97.315						
17	0.202	1.062	98.376						
18	0.181	0.951	99.327						
19	0.128	0.673	100.000						

Extraction method: Principal component analysis

The extraction sum of squared loadings calculates the value of retained factors using common variance. This shows that a total of 7 factors are retained from a total of 19 factors. The rotation sums of squared loading are used to maximize the variance of each of the 7 factors in order to redistribute the total variance [19].

6.5.5 Scree Plot

The scree plot is the graph where the eigenvalue is plotted against the number of factors used for the study. The factors from where the line becomes flat show the magnitude of variance becoming smaller and smaller. Therefore, the graph shown in Fig. 8 depicts that the maximum variance is explained by the first seven variables. [15].

To further validate the extraction of the seven variables from the 19 variables of the study, we have calculated the correlation and covariance of each factor with these seven factors. The results show the high correlation and the appropriateness of the variables extracted. The Component Matrix (Table 6), Rotated Component Matrix (Table 7), Component Transformation Matrix (Table 8), Component Score Coefficients (Table 9), and Component score covariance matrix (Table 10) are shown in the appendix for reference.

6.5.6 Factor Loadings

The factors retained after performing factor analysis are brand reputation; purchase service, physical appearance, features, technical features, value additions, and

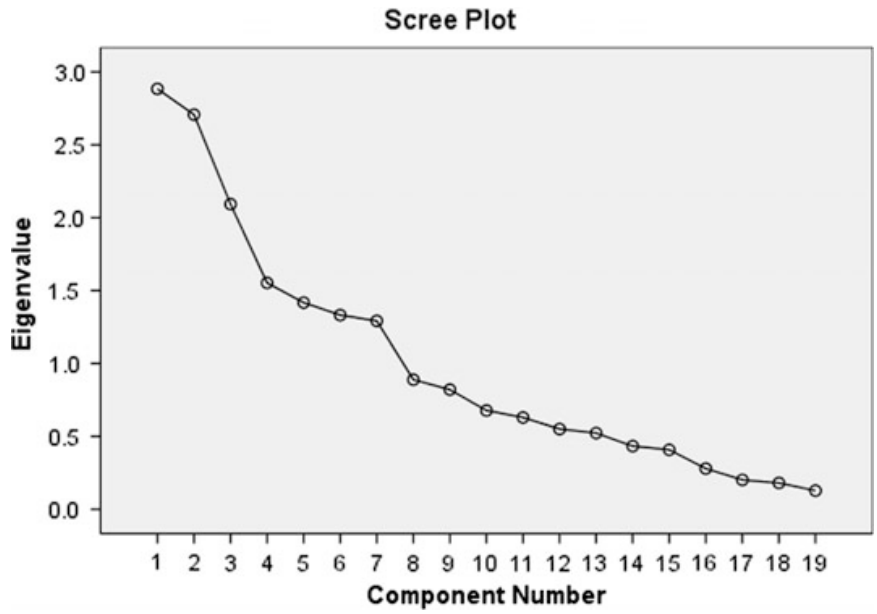


Fig. 8 Scree plot

Table 6 Component matrix (a)

	Component						
	1	2	3	4	5	6	7
Technology	−0.283	0.534	0.331	−0.056	0.286	−0.228	0.087
Brand_name	0.597	−0.234	−0.267	0.078	−0.128	−0.276	−0.104
Price	−0.077	0.383	0.181	0.561	−0.412	−0.137	0.294
Interior	0.342	−0.440	0.233	0.126	0.304	−0.227	0.450
Design_and_exterior_looks	0.639	−0.102	−0.298	0.200	0.252	−0.194	0.048
Comfort_and_convenience	0.574	−0.252	−0.274	−0.195	0.442	0.171	−0.207
Safety_and_security	0.159	0.470	0.251	0.055	0.460	0.259	−0.210
Engine	0.020	0.599	0.195	−0.399	0.086	−0.271	0.112
Mileage	0.497	0.715	−0.093	0.003	0.038	−0.157	0.088
Boot_space	−0.400	−0.216	0.071	0.276	0.066	0.476	0.269
Value_for_money	0.441	0.331	0.614	−0.005	−0.030	−0.103	0.045
Warranty	−0.275	0.100	−0.063	0.487	0.630	0.054	−0.017
After_sale_services	0.506	0.057	0.074	0.045	−0.186	0.477	0.437
Resale_value	−0.374	0.425	−0.485	0.161	0.211	0.126	0.170
Availability_of_spare_parts	0.376	−0.335	0.562	0.285	−0.039	0.073	−0.175
Discounts_or_special_offers	0.420	0.274	−0.498	0.361	−0.069	0.009	0.259
Fuel_economy	0.221	0.517	−0.415	−0.166	−0.262	0.364	−0.235
Vehical_availability	0.326	0.219	0.426	0.163	−0.051	0.448	−0.316
Waiting_time	−0.120	0.096	−0.084	0.605	−0.144	−0.283	−0.573

Extraction method: Principal component analysis. 7 components extracted

performance. Several attributes are studied under these variables. The factor loadings of these are shown using the component matrices [10].

The major results from the component matrices are summarized from Tables 11, 12, 13, 14, 15, 16, and 17:

Factor 1 comprises brand reputation provided by the company. Here price has maximum factor loading of .863 so it becomes an important factor.

Factor 2 comprises purchase services provided by the company. Here warranty has maximum factor loading of .812, so it becomes an important factor.

Factor 3 comprises physical appearance provided by company. Here design and exterior looks have maximum factor loading of .768, hence it becomes an important factor.

Here, factor 4 comprises features. Here, interiors have maximum factor loading of .754 so it becomes an important factor.

Factor 5 comprises technical features of the car provided by the company. As we can see, engine has maximum factor loading of .783, so it is an important factor.

Factor 6 comprises value addition. Here, waiting time has maximum factor loading of .783, so it becomes an important factor.

Factor 7 comprises performance provided by the company. Here, fuel economy has maximum factor loading of .863 which makes it an important factor.

Table 7 Rotated component matrix (a)

	Component						
	1	2	3	4	5	6	7
Technology	−0.284	0.637	−0.014	−0.070	0.305	0.157	−0.095
Brand_name	0.652	−0.122	0.043	−0.079	−0.344	−0.113	−0.117
Price	0.863	0.093	0.103	0.016	0.069	0.073	−0.015
Interior	0.294	−0.034	0.057	0.754	0.009	−0.060	0.254
Design_and_exterior_looks	179	−0.005	0.048	−0.157	0.051	0.768	0.012
Comfort_and_convenience	0.477	−0.100	0.148	0.005	0.052	0.705	0.116
Safety_and_security	−0.003	0.389	0.025	0.177	0.465	−0.184	0.549
Engine	−0.099	0.783	−0.072	0.106	−0.067	0.023	0.072
Mileage	0.505	0.159	0.122	0.308	0.053	628	0.112
Boot_space	−0.352	−0.448	−0.004	0.689	0.388	0.172	0.280
Value_for_money	0.084	0.515	0.122	−0.163	−0.154	0.166	0.574
Warranty	0.019	−0.036	0.812	−0.157	0.022	−0.003	−0.202
After_sale_services	0.270	−0.095	0.315	0.089	−0.073	0.153	0.708
Resale_value	0.554	0.067	−0.444	0.369	0.014	0.157	0.040
Availability_of_spare_parts	0.064	−0.191	0.691	−0.383	−0.143	0.018	−0.058
Discounts_or_special_offers	0.184	−0.028	−0.115	0.246	0.161	0.277	0.700
Fuel_economy	0.203	0.095	0.046	0.060	0.863	−0.040	0.126
Vehical_availability	−0.034	0.027	0.775	0.234	0.064	0.007	0.073
Waiting_time	0.160	−0.169	0.166	0.146	0.132	0.305	0.783

Extraction method: Principal component analysis. rotation method: Varimax with Kaiser normalization a rotation converged in 10 iterations

Table 8 Component transformation matrix

Component	1	2	3	4	5	6	7
1	0.768	0.127	0.460	−0.029	−0.283	−0.196	0.250
2	0.040	0.712	0.065	0.556	0.269	0.324	0.021
3	−0.478	0.293	0.661	−0.458	−0.112	0.158	0.044
4	0.323	−0.375	0.262	−0.142	0.486	0.590	−0.286
5	0.086	0.243	−0.022	−0.334	0.713	−0.559	−0.016
6	−0.250	−0.426	0.372	0.453	0.287	−0.159	0.552
7	0.080	0.098	−0.374	−0.379	0.084	0.382	0.740

Extraction method: Principal component analysis. rotation method: Varimax with Kaiser normalization

Table 9 Component score coefficient matrix [6]

	Component						
	1	2	3	4	5	6	7
Technology	−0.089	0.316	−0.030	−0.125	0.146	0.027	−0.051
Brand_name	0.270	−0.033	−0.027	−0.036	−0.174	−0.007	−0.142
Price	0.079	−0.017	0.031	−0.047	−0.006	0.544	0.013
Interior	0.147	0.060	−0.059	−0.437	0.083	0.030	0.168
Design_and_exterior_looks	0.333	0.020	−0.033	−0.120	0.094	−0.064	−0.044
Comfort_and_convenience	0.153	−0.027	0.067	0.035	0.118	−0.420	0.025
Safety_and_security	−0.030	0.133	0.251	0.076	0.308	−0.188	−0.005
Engine	−0.050	0.392	−0.097	−0.029	−0.085	−0.035	0.035
Mileage	0.202	0.260	−0.002	0.074	0.019	0.076	0.031
Boot_space	−0.137	−0.252	0.054	−0.014	0.254	0.108	0.265
Value_for_money	0.002	0.225	0.230	−0.112	−0.079	0.092	0.038
Warranty	0.071	−0.023	0.031	−0.133	0.520	−0.049	−0.109
After_sale_services	0.055	−0.114	0.123	0.064	0.004	0.141	0.488
Resale_value	0.050	−0.003	−0.192	0.126	0.300	0.053	0.078
Availability_of_spare_parts	−0.001	−0.105	0.349	−0.136	−0.027	0.040	−0.080
Discounts_or_special_offers	0.315	−0.061	−0.094	0.071	0.110	0.207	0.115
Fuel_economy	0.028	−0.052	0.062	0.464	−0.069	−0.057	0.064
Vehical_availability	−0.080	−0.083	0.437	0.191	0.068	−0.029	0.014
Waiting_time	0.123	−0.115	0.147	0.089	0.045	0.165	−0.567

Extraction method: Principal component analysis. rotation method: Varimax with Kaiser normalization. Component scores

Table 10 Component score covariance matrix [4]

Component	1	2	3	4	5	6	7
1	1.000	0.000	0.000	0.000	0.000	0.000	0.000
2	0.000	1.000	0.000	0.000	0.000	0.000	0.000
3	0.000	0.000	1.000	0.000	0.000	0.000	0.000
4	0.000	0.000	0.000	1.000	0.000	0.000	0.000
5	0.000	0.000	0.000	0.000	1.000	0.000	0.000
6	0.000	0.000	0.000	0.000	0.000	1.000	0.000
7	0.000	0.000	0.000	0.000	0.000	0.000	1.000

Extraction method: Principal component analysis. rotation method: Varimax with Kaiser normalization. Component scores

Table 11 Factor 1 brand reputation

Attributes	Factor loading	Mean
Brand name	0.652	2.10
Price	0.863	3.67
Result value	0.554	0.554

Table 12 Factor 2 purchase service

Attributes	Factor loading	Mean
Warranty	0.812	3.18
Availability of spare parts	0.691	2.14
Vehicle availability	0.775	2.53

Table 13 Factor 3 physical appearance

Attributes	Factor loading	Mean
Design and exterior looks	0.768	2.42
Comfort	0.705	2.19

Table 14 Factor 4 features

Attributes	Factor loading	Mean
Interior look	0.754	2.32
Boot space	0.689	1.51

Table 15 Factor 5 technical features

Attributes	Factor loading	Mean
Technology	0.637	2.09
Engine	0.783	2.63

Table 16 Factor 6 value addition

Attributes	Factor loading	Mean
Waiting time	0.783	2.63
Discount	0.700	2.19
Safety	0.652	1.49
Value for money	0.863	1.53
After sales value	0.554	2.29

Table 17 Factor 7 performance

Attributes	Factor loading	Mean
Mileage	0.628	1.84
Fuel economy	0.863	2.14

7 Conclusion

The understanding of customer's perception is an ongoing process to survive the cut throat competition. Taking this into consideration, the study provides insights about the several attributes which drive a customer's behavior toward buying a product of a brand. The study has used theories and exploratory research design along with analytical tools to identify the major attributes of consumer buying behavior toward sedan cars within Delhi/NCR among high-end consumers. This

knowledge helps car manufacturers in market segmentation which enables the organization to plan the market strategies toward consumer retention and product upgradation.

We have the output of factor analysis for this problem—the un-rotated factor matrix, the final statistics comprising the communality for all 19 variables, and the eigenvalues of all factors having eigenvalues of more than 1. The seven factors are extracted using factor analysis namely, brand reputation, physical appearance, features, purchase services, technical features, value addition, and performance. These variables capture approximately 70% of the total variance information content out of the 19 variables under study.

Under the Brand Reputation, price is the most influential factor having factor loading of .863, followed by warranty under Purchase Services having factor loading of .812, design and exterior looks under physical appearance with factor loading of .768, interior look under features with factor loading of .754, engine strength under technical features with factor loading of .783, value for money under value addition with factor loading of .863, and fuel economy under performance with factor loading of .863.

The automobile companies must focus on their male consumers falling under the age group of 25–35 years who lie in the business class within an income group of 12–17 lacs. The marketing strategy, advertisement, and price strategy should be specially designed for this segment of the consumers based on good mileage, luxury, comfort, and engine strength and fuel economy.

The automobile companies should also pay attention to additional factor, i.e., the time span of car switch among the customers to retain them. The customers upgrade their cars usually between a period of 3–5 years by either purchasing a new car of their trusted brand or switching the brand. So, to retain the customers, it is important that the companies keep a focus on R&D to develop new upgraded cars and intimating their existing customers from time to time.

Acknowledgements This manuscript is supported by different individuals (students and faculty) associated with Amity University campuses: Ref. No. AUGN/Cert./VP(P) & R/2018/10/01. All procedures performed in the study involving human participation were in accordance with the ethical standards of the institution, and comparable ethical standards. Prior public disclosure approval to publish the study results have been taken by the university. The anonymity of the respondents of the questionnaire and data has been treated strictly confidential and not disclosed at any level.

Appendix

Tables 6, 7, 8, 9, 10.

References

1. Amy Van Looy, A.S.: Business process performance measurement: a structured literature review of indicators, measures and metrics. **5**(1) (2016)
2. Banerjee, S.: Study on consumer buying behavior during purchase of a second car. J. Mark. Commun. **6**(2) (2010)
3. Child, D.: The Essentials of Factor Analysis. Continuum International Publishing Group, New York (2006)
4. Field, A.P.: Discovering Statistics Using SPSS for Windows: Advanced Techniques for the Beginner. Sage, London (2000)
5. Field, D.A.: Chapter 14: Factor Analysis Using SPSS; Discovering Statistics Using SPSS, 2nd edn. Sage, London (2005)
6. Gabay, J.: Brand Psychology: Consumer Perceptions, Corporate Reputations. Kogan Page, London (2015)
7. Jadczaková, V.: Review of segmentation process. ACTA UNIVERSITATIS AGRICULTURAE ET SILVICULTURAE MENDELIANAE BRUNENSIS 1215–1224 (2013)
8. Lorenzo-Seva, U.: How to report the percentage of explained common variance in exploratory research. Tarragona (2013)
9. Mathur, D.: Consumer buying behaviour of cars in India - a survey. In: 1st International Conference on New Frontiers in Engineering, Science & Technology, New Delhi, India, pp. 468–473 (2018)
10. Menon, B.: A study on consumer behaviour of passenger car segments through logistic regression modelling 20–32 (2017)
11. Monga, N.: Car market and buying behavior- a study of consumer perception. Int. J. Res. Manag. Econ. Commer. **2**(2) (2012)
12. Motoyuki Akamatsu, P.A.: Automotive technology and human factors research: past, present and future. Int. J. Veh. Technol. **27** (2013)
13. Pearce, A.G.: A beginner's guide to factor analysis: focusing on exploratory factor analysis. Tutor. Quant. Methods Psychol. 79–94 (2013)
14. Raju, P.T.: Factors influencing consumer behaviour for buying luxury cars. Int. J. Civ. Eng. Technol. **292** (2018)
15. Sachdeva, N.: An innovation diffusion model for consumer durables with three parameters. J. Manag. Anal. **3**(3) (2016)
16. Seunggeun Lee, F.Z.: Convergence and prediction of principal component scores in high dimentional settings. Ann. Stat. 3605–3629 (2010)
17. Shende, V.: Analysis of research in consumer behavior of automobile passenger car customer. Int. J. Sci. Res. Publ. **4**(2) (2014)
18. Stebbins, R.A.: Exploratory Research in the Social Sciences. Sage, London (2001)
19. Venkatesh Shankar, A.K.: Customer satisfaction and loyalty in online and offline environments. Int. J. Res. Mark. 153–175 (2003)
20. Watkins, D.S.: The matrix eigenvalue problem: GR and Krylov subspace methods. SIAM (2007)

Secure Online Voting System Using Biometric and Blockchain

Dipti Pawade, Avani Sakhapara, Aishwarya Badgujar, Divya Adepu and Melvita Andrade

Abstract Elections play an important role is democracy. If the election process is not transparent, secure and tamper proof then the reliability and authenticity of whole process is at stake. In this paper, we have discussed online voting system which fulfills all the above system requirements. We have addressed the issue of user authentication through iris recognition. We have used One Time Password (OTP) to have additional security check. We have also taken care that one valid user should not cast multiple votes. Use of Blockchain is the another security measure implemented in order to provide decentralized, tamper proof storage of data related to users biometric, personal details and votes casted by them. Thus we are not only focusing on user authenticity but also data security is also taken into consideration. The performance of the system has been tested for users from different age group and different background and its inference is presented.

Keywords Online voting · Blockchain · Sidechain · Iris recognition

1 Introduction

In democratic country like India, Election is considered as vital process. Right from choosing the Prime minister to the ward member, every election is carried out using voting system. Earlier, the ballot paper voting system was used in election process. Here people need to visit the appropriate Polling booth and need to show the voter ID card to prove their authenticity. Casting of vote is done by stamping on ballot paper secretly and then ballot paper is folded and deposited in ballot box. After the voting is done, the ballot boxes are sealed and moved to the counting center under security and counting of votes is carried out manually. But there were some incidences of manipulation with ballot boxes. Also this complete process took longer time for result declaration, required tremendous human resources and thus incurred high

D. Pawade (✉) · A. Sakhapara · A. Badgujar · D. Adepu · M. Andrade
Department of IT, K.J. Somaiya College of Engineering, Mumbai, India
e-mail: diptipawade@somaiya.edu

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_7

costing. Because of all these reasons, Electronic Voting Machine (EVM) is introduced in 1982 in 70-Parur Assembly Constituency of Kerala on trial basis [1]. At mass level it is used in 1998 assembly elections in Madhya Pradesh, Rajasthan and the NCT of Delhi. After that, EVM is being used in all the elections conducted by Election Commission of India. EVMs are under purview of Election Commission of India only and cannot be used for other elections like university senate elections, society election or elections carried out in various organizations. Thus they are still using the traditional ballot paper base election process.

In this paper we have addressed to the various issues related to the university senate election, where electoral roll of graduate teachers is made electronically but still ballot paper is used for vote casting. It has been observed that percentage of votes casted is very low as compared to the total number of eligible voters. One of the reasons for this reluctance is that for casting vote one needs to travel to the particular place. Most effective solution to this is like preparing electoral list and conducting election through e-portal. But due to security concerns, this option is ruled out. Hence we are proposing the concept of online voting system where voter identification is done using biometric iris recognition. As biometric trait is the basic parameter under consideration to verify the genuine voter, security of iris template database is also our concern. For that purpose we are storing all the details of voter including iris feature on Blockchain, so that it cannot be tampered. The system takes care of the issue related to multiple votes casted by single voter. Once the pre-set time of election conduction is over the portal gets locked and result statistics are automatically emailed to the election candidates and concerned authority.

Further outline of the paper is as follows: Sect. 2 consists of related work in the area of applications of online elections and research in iris recognition system. In Sect. 3 overview of Blockchain and related technologies is given. In Sect. 4, the design and implementation of the proposed online voting system is discussed. In Sect. 5 results are presented and in Sect. 6 conclusion and future work is stated.

2 Related Work

2.1 Literature Survey on E-Voting Application

Khasawneh et al. [2] discussed the implementation of biometric based secure e-voting system. For voter identification, initially they have used official voter ID cards which need to insert in magnetic card reader to fetch the details of the voters from local or central database. If record is found then the individual's fingerprints are scanned and matched against the pre-stored database. Thus only authenticated person is allowed to vote and once voting is done then flag is set so that same person cannot do the voting again. Main concern of this approach is that the authentication is solely dependent on central Authentication server. If that server is down, whole voting system will be affected.

Petcu et al. [3] describes a mobile e-system consisting of various phases. Firstly one needs to enroll themselves by visiting physically to the authentication center. Here they need to provide some valid identity proof and if person is eligible to vote then his/her mobile number, voice and fingerprint samples are collected and key token is given to the voter. In authentication phase, the user has to access the web portal and has to provide the ID card details and key token provided at the time of enrolment. The system will then send a SMS key on successful collection of three voice samples and fingerprint samples for both the hands of a voter. Then one needs to provide the unique code, key token and SMS key which is verified by the authority and then vote access is given to that voter. Now voter can use e-voting module to cast the vote securely and vote counting is also done automatically. Here author have used so many key values which need to be provided, which make the system complicated to use.

Sridharan [4] considered the smart card, voter identification number (VID) and fingerprints to authorize and uniquely identify the individual. The complete system has four modules viz. Franchise Excising Terminal to recognize smart card, Authentication Terminal, the Distributed Databases and Central Controlling Server. Use of distributed database makes the system scalable. But still biometric data security concern exists.

Agarwal et al. [5] proposed the AADHAR ID based online authentication system where voter first needs to provide AADHAR ID which is being verified by authority and then password is mailed to the particular person. One needs to login into online voting system to cast the vote and it is accepted only if fingerprint matching is done successfully. Author has not given the insight to how templates for fingerprint matching are collected and how the fingerprint server database is maintained.

Table 1 gives the overview of the different parameters used in various e-voting applications for voter authentication.

Table 1 Overview of available e-voting application

S. No.	Author	Parameter used for authentication
1	Khasawneh et al. [2]	Specially designed magnetic tape based Voter ID card, fingerprints
2	Petcu et al. [3]	Physical Identity proof, Voice sample, fingerprint of both hands, key token, 4 digit key, SMSkey
3	Sridharan [4]	Smart card, VID and fingerprints
4	Agarwal et al. [5]	AADHAR ID, Password, fingerprints

2.2 Literature Survey on Iris Recognition System

Garagad et al. [6] proposed iris identification system which is claimed to be invariant to distance and eye tiltation. They have Performed segmentation to take out the region of interest. It involves low intensity region recognition in input image, detecting iris and boundary inference point (edge detection) followed by selective image cropping followed by pupil localization and normalization. For feature extraction first cartesian plane data of an image is obtained through signature code generation and then it is converted to a binary iris signature code to optimize the memory utilization and speed up the processing. Finally for iris matching binary signature code of input iris image and template image are compared using XOR bitwise operation. The author [6] concluded that, this approach is simple to implement and effective way of iris authentication in case of scale and tilt variance.

Neda et al. [7] have suggested hybrid approach for iris recognition model based on a multi-layer perceptron NN (MLPNN) and particle swarm optimization (PSO) algorithms. Here, to separate the sclera and pupil, iris localization is carried out using circular Hough transform algorithm [8, 9]. Then normalization is carried out to convert image from the space of the Cartesian coordinates to polar coordinates using rubber-sheet model of Daugman followed by image segmentation [10]. 2D gabor kernel method is applied to the pre-processed image to extract the features. For pattern matching author have proposed a hybrid approach based on ANN-based MLP and PSO algorithm and called it as MLPNN–PSO algorithm. This hybrid approach achieves 95.36% accuracy rate which is much better when these approaches are used separately.

Muthana et al. [11] discussed two methods of iris feature extraction viz. the Fourier descriptors (FD) and Principle Component Analysis (PCA). Before feature extraction, iris segmentation is carried out using pupil and iris localization. The pupil localization consists of following steps: smoothing using non-linear median filter, binarization, and noise removal using morphological operation and finding the iris center. While the iris localization is done by applying Canny masks which is used to locate iris boundary and circular Hough transform which is used to get complete circular shape. Next step in pre-processing is normalization using rubber sheet model. Then for feature extraction FD and PCA is used and their results are classified using Manhattan, Euclidean and Cosine distance. According to their result discussion accuracy rate of FD with Manhattan distance is highest i.e. 96%.

Table 2 gives the overall summary for different iris recognition system based on various parameters.

Table 2 Summary of iris recognition system

Database	Preprocessing	Feature extraction	Classifier	Recognition rate	Peculiarity
High-quality UPOL iris images captured in constrained environment [6]	Segmentation	Gray signature code generation and binary signature code generation	Bit to bit comparisons of binary signature codes using XOR operation	99.64 (for Normal Test input images) 90.698 (for tilt & scale variation)	Approach is invariant to distance and tilt variations
(CASIA)-iris V3 and UCI [7]	Circular Hough transform algorithm, rubber-sheet model	2D Gabor kernel method	Hybrid MLPNN–PSO algorithm	95.36%	MLPNN and PSO algorithm gives better
CASIA v1 [11]	Pupil and Iris localisation, iris normalization	Fourier descriptors (FD) and Principle Component Analysis (PCA)	Manhattan, Euclidean and Cosine distance	96%	FD is better than PCA and FD with Manhattan gives highest accuracy

3 Overview of Blockchain Technology

Database is considered as a very important component when we are talking about full stack application development. Traditional databases follow the client-server architecture, where client/user has a provision to perform CRUD operations (Create, Read, Update, and Delete) and data is stored on a centralized server. As all clients have the rights to modify or update the data, there is a database administrator who authorizes the client and makes sure that the data is not being compromised. But here as data is maintained in centralized database, the whole system gets hampered if database server crashes. Also the load on one server is more thus the overall computation speed is low. To address this problem the concept of distributed database was introduced. A distributed database is considered as a single logical database where nodes are physically distributed across several geographic locations and communicate with each other over a network. This kind of database is normally suited for an organization, where all trust each other. Here in case if any node fails the system continues to work with other node.

Blockchain is said to be a decentralized database where data integrity and transparency is considered as parameter for public verifiability [12]. Opposite to traditional database, Blockchain allows only read and write operations to be performed on the data. In read operation user can read/query the data on Blockchain

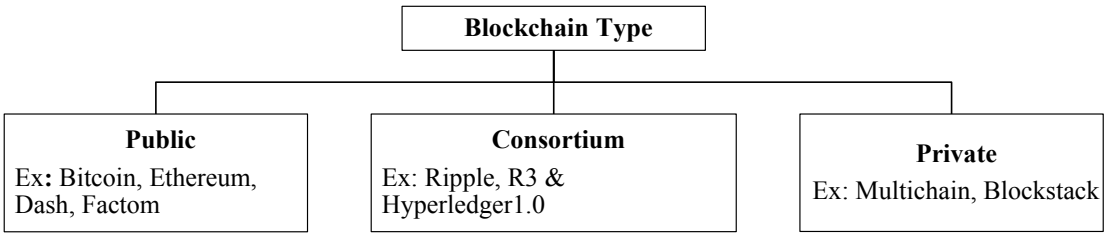


Fig. 1 Types of Blockchain

while write operation used to add the data on Blockchain. Whenever write operation is performed an additional block is appended in Blockchain. All previous data is available in block which cannot be altered. If it is being altered then it is added as new block and the original block data remains the same [13]. Thus it is considered as scam free, transparent system.

Figure 1 shows different types of Blockchain. Public Blockchain is also called as permissionless Blockchain, as there is no authority to sanction any Blockchain transaction. This means process of consensus building is public. Second type of Blockchain is consortium which is also known as public permissioned Blockchain. Here consensus process is monitored by a pre-selected group of nodes and for transaction verification minimum clearance is required. Private/private permissioned Blockchain is the third type of Blockchain where write operation rights are given to set of specific nodes while read rights can be public or restricted to specific nodes [14].

3.1 Sidechain

A sidechain is considered as separate blockchain that is attached to the blocks on the main blockchain. Main chain is also referred to as "Parent Blockchain" while sidechain is referred to as "childchain". The assets can be interchanged between the main chain and the sidechain at a predetermined rate. The major advantage of using sidechain is that one can easily develop and deploy a quickly scalable Blockchain solutions that too at lower cost and under high level of security [15].

3.2 Ethereum Platform

For our application development we are using Ethereum public Blockchain platform which allows the development of decentralized application (dapp). Ethereum platform has two building blocks viz. Ethereum Virtual Machine (EVM) and peer-to-peer network protocol. In order to get consensus for transaction, each node

in network runs the EVM and processes the instruction. In simple words EVM is a run time environment which converts the smart contract into EVM bytecode and deploys it onto Blockchain for execution. Here smart contracts are the basic building blocks of a program which consist of contract instructions which are simply set of instructions to be executed upon certain event. We have written our smart contract in solidity language [15].

3.3 Metamask

Earlier the biggest issue with using dapps was, using normal browsers it was not possible to write to the Blockchain. There were special Blockchain browsers like Mist but using it was very tricky and thus the whole development process became tedious. The solution to this issue is provided in the form of Metamask which is an Ethereum web browser extension. Metamask basically acts like a Ethereum wallet which stores Ethereum related data like private keys, public addresses etc. MetaMask is considered as the easiest way to interact with dapps in a browser. It allows you to connect to an Ethereum network without running a full node on the browser's machine [16].

4 Implementation Approach

As shown in Fig. 2, system has three main modules running on client side,

- Voter enrollment module using which user needs to enroll themselves as a authorized voter
- Election authority interface through which the election process is initiated and election result calculation command is triggered.
- Election interface, which verifies the identity of the user as valid voter and provides interface to cast the vote.

In order to connect our web user interface with the Ethereum Blockchain we have used Metamask. The overall system implementation is carried out as follows:

4.1 Working of Proposed System

As per the AICTE rules, every institute affiliated to the university needs to send the details of their employees to the university. These details include information like name, AADHAR ID number (UID), Email ID, PAN number, address etc. So before university election, name and Email ID and UID of eligible candidates are collected

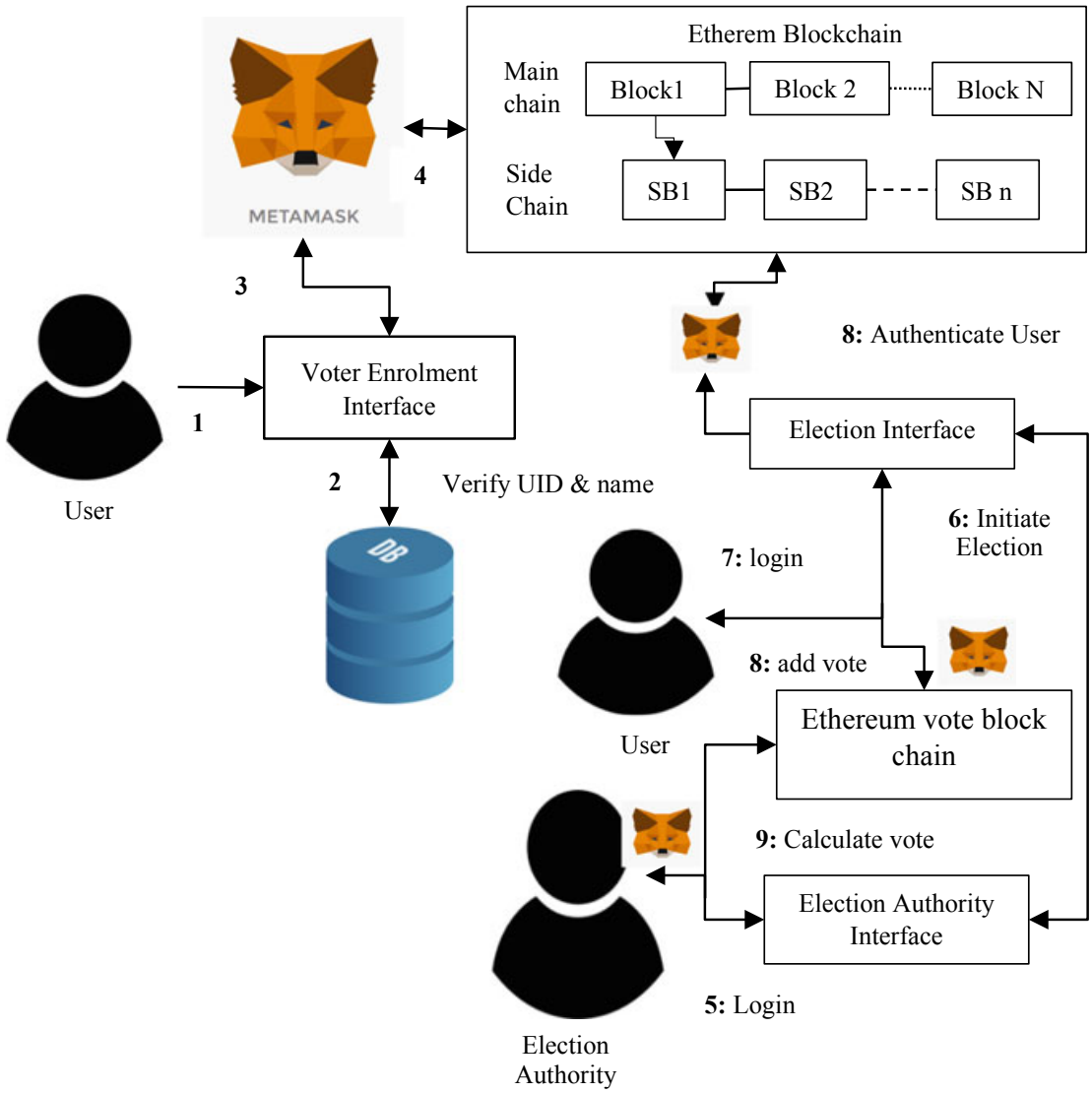


Fig. 2 Block diagram of proposed online voting system

from each institute and stored in database. A bot is designed to send auto-generated mail to notify eligible voters regarding voter enrollment deadline and guidelines. This bot will fetch the email IDs from the database and will send template mail to all specifying the link for registration and the description of the procedure to be followed.

Then the user needs to visit our election portal, for registration each user needs to provide basic information like name, address, valid mobile number and UID. This UID and name of the user is matched against the database for verification. If it gets matched then user has to choose unique username and password for them. This combination of username and password will be generated after successfully satisfying the required validation rules. For verification, link is sent on the respective Email Id. User needs to verify themselves by clicking on that link. Once verification is done, all the user data is pushed on Ethereum Blockchain through smart contract, then mining is carried out and genesis block is created [17].

The system will then ask the user to scan the iris. User can use webcam or front camera of mobile to take a picture of iris. The iris image is preprocessed and features are extracted. Preprocessing and feature extraction process is explained in Sect. 4.2. Once features are extracted, it is stored on associated sidechain along with other information like name, address, mobile number.

On the election day, the authorized person will instantiate the election portal. User have to login using the appropriate username and password which was generated during the enrollment phase. Then the user needs to provide his/her valid UID. System will then try to find the block on Blockchain with same UID. If the block is not found, then error message will be displayed otherwise system will scan the iris image through web-cam or front camera of mobile. Preprocessing and feature extraction is done using step in Sect. 4.2. These features are matched against the features stored on sidechain at the enrollment phase. If iris features do not match then the iris image is scanned again and matching is carried out. If match found, then One Time Password (OTP) is generated and sent to the mobile number stored on sidechain for that user. This OTP is valid for 40 s only. User needs to get verification done before that OTP expires. Because of some reason if OTP expires then the user can request for OTP again. Once OTP verification is done, e-ballot will appear on screen containing names and symbols of the candidates. In order to cast vote user needs to select the radio button in front of the particular candidate. Once voting is done flag is set for this user, so that same user should not cast the vote again. The vote casted by user is again pushed on Blockchain through smart contract and genesis block is created on voting Blockchain after mining is done.

The election authority person will disable the election portal once election period is over. On election authority console, calculate button is provided which will fetch all the votes from Blockchain, and calculate the number of votes gained by each candidate. Then the system will mail the result to the authorized members of election committee and the election candidates. Figure 3 depicts the flow of activities when user interacts with the election interface.

4.2 Iris Recognition System

Iris recognition system consists of following three steps:

- Preprocessing
- Feature Extraction
- Feature matching

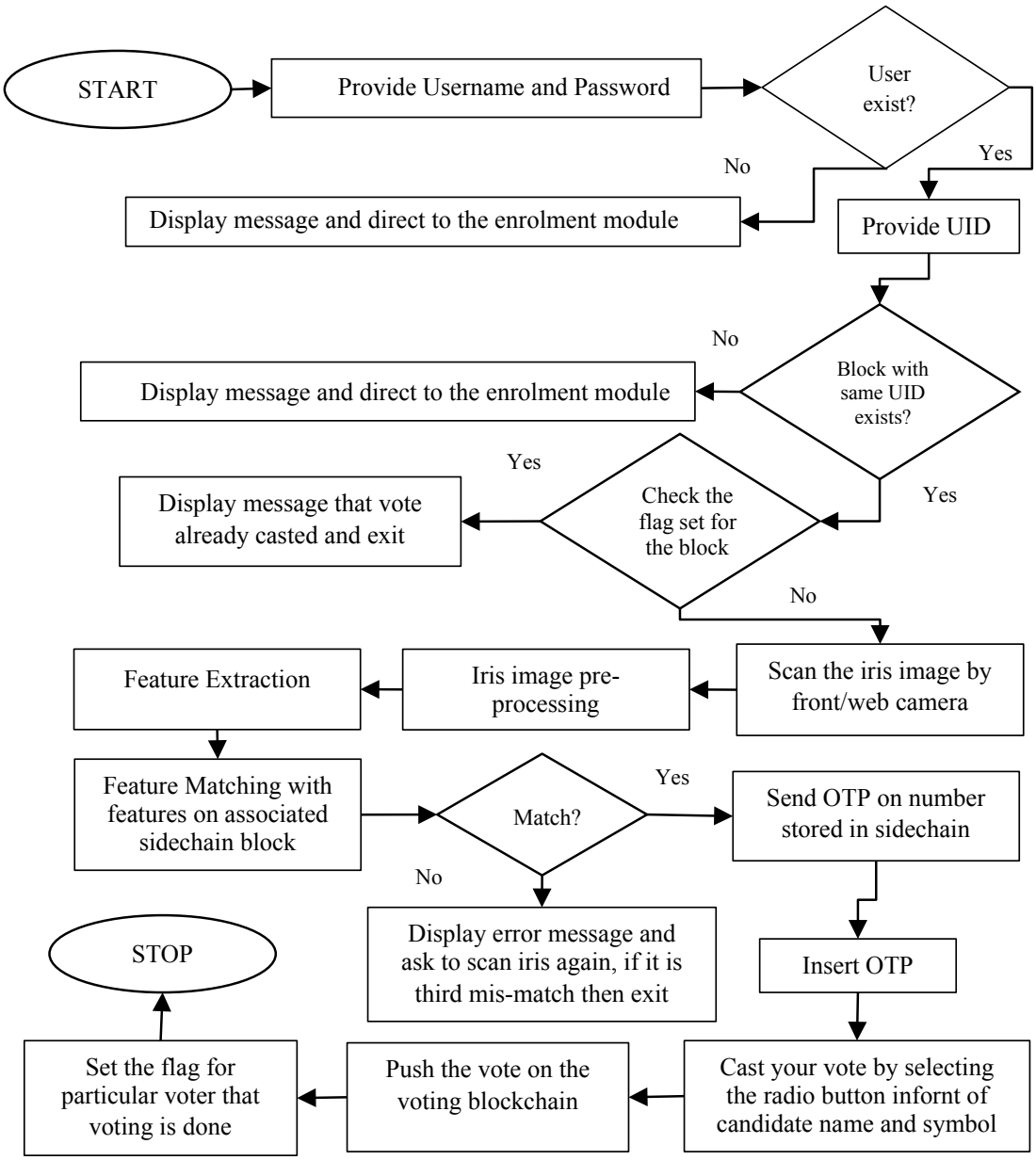


Fig. 3 Flowchart for online voting system for enrolled

4.2.1 Preprocessing

The first step is to convert the colored iris image to a greyscale image which contains only different shades of grey lying in the range of 0–255. This step is important because it reduces the unnecessary information in each pixel [18].

This grey color has the red, green and blue components in equal intensity in the RGB space. So henceforth, we only have one value in every pixel as compared to three values for every pixel of a colored image. Figure 4 depicts the process of color image conversion to greyscale image for any pixel at position (x, y) in colored image.

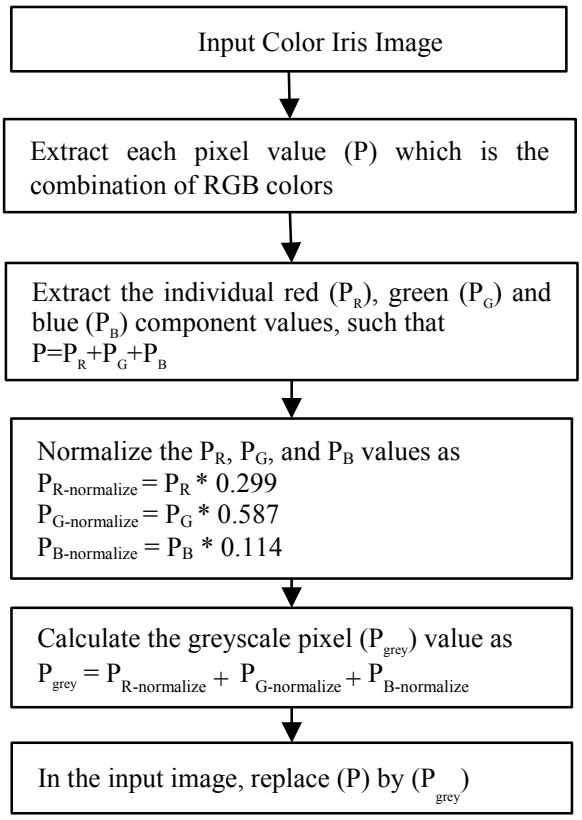


Fig. 4 Steps for color to greyscale Image conversion

The next step is to detect the edges of the iris and pupil. It works by detecting discontinuities in brightness. Sobel operator has been used for this purpose.

In sobel, we apply the derivative mask in horizontal and vertical direction to increase the edge intensity. The resultant is a faint boundary of the iris and pupil depending on the brightness of the image. It works by calculating the gradient of image intensity at each pixel within the image. It finds the rate of change of brightness.

The horizontal gradient Gx is the image that we get after applying the horizontal mask on the greyscale image (A) which is calculated as using the mask in Fig. 5a and the vertical gradient Gy is the image that we get after applying the vertical mask on the greyscale image (A) which is calculated by using mask in Fig. 5b. The masks contain positive and negative coefficients. This means the output image will also contain positive and negative values. The negative values are set to 0. The resultant image is the sum of the horizontal and vertical gradients.

(a)

$$G_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix} * A$$

(b)

$$G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A$$

(c)

p4	p3	p2
p5	p	p1
p6	p7	p8

Fig. 5 **a** Sobel vertical mask, **b** Sobel horizontal mask, **c** CN mask

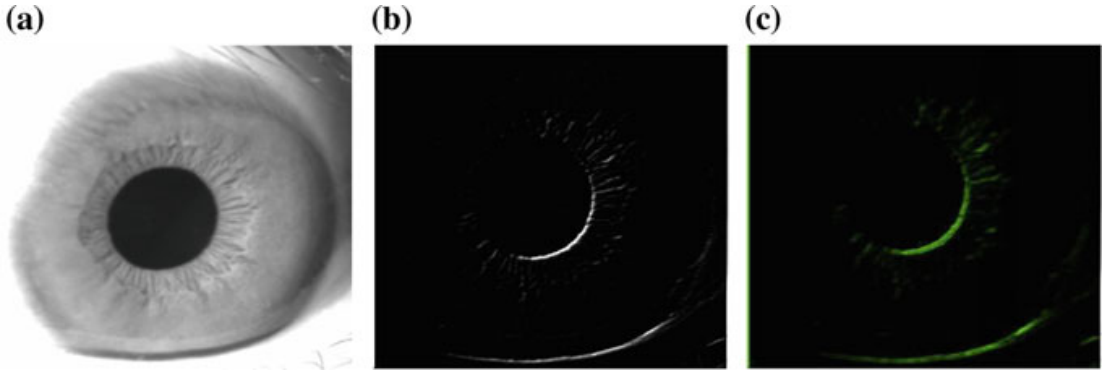


Fig. 6 **a** Sample input iris image, **b** after applying Gabor filter, **c** after applying CN mask

The next step is to use Gabor filter to analyze the texture of the iris. It basically does this by analyzing whether there is any specific frequency content in the iris in the localized region inside the iris circle. The image is convolved with a series of Gabor wavelets to perform a filtering operation. The buffered image is filtered and the resultant is produced as a rendered image [19, 20] (Fig. 6).

4.2.2 Feature Extraction

Minutiae representation is most widely used method of any biometric representation as it helps to determine uniqueness in an image. Minutiae or small details mark the regions of local discontinuity within a biometric image. For detecting minutiae points we have used Rutovitz's Cross Numbering technique.

This method extracts the minutiae points by identifying the bifurcation points (forking points) using a 3 × 3 mask on the local neighborhood pixels of the ridge pixel. These properties identify various properties (minutia) of iris like ridge endings, ridge bifurcations, etc. Figure 5c shows 3 × 3 Mask used for CN technique [21].

$$\mathrm{CN} = \frac{1}{2}\sum_{i=1}^{8} |P_i - P_{i+1}| \tag{1}$$

Using Eq. (1) we get iris feature vector. This output is then used for the matching algorithm.

4.2.3 Matching by Average Method

Step 1:
Here, each pixel is represented as 0xAARRGGBB (alpha, red, green, blue). By performing a bitwise-and with 0xFF, the individual RGB component is calculated as:

$$\begin{aligned} \text{red} &= (\text{rgbA} \gg 16) \,\&\, 0\text{xff} \\ \text{green} &= (\text{rgbA} \gg 8) \,\&\, 0\text{xff} \\ \text{blue} &= (\text{rgbA}) \,\&\, 0\text{xff} \end{aligned} \tag{2}$$

Step 2:
We calculate separate RGB values using Eq. (2) for both input and template iris images and calculate sum of difference between each of the values as

```
difference += Math.abs(redA − redB);
difference += Math.abs(greenA − greenB);
difference += Math.abs(blueA − blueB);
```

Step 3:
Then we calculate total number of pixels:
Number of pixels for each of the RGB component is calculated as given in Eq. (3) and then the summation of number of pixels for each RGB component is taken which gives the total number of pixels.

$$\text{Number of pixels for each RGB component} = \text{width} * \text{height} \tag{3}$$

Step 4:
For normalizing the value of different pixels and for accuracy we use average pixels per color component calculated as

```
average_different_pixels = difference/total_pixels
```

Step 5:
As there are 255 values of pixels(RGB) in total thus

```
Percentage = (average_different_pixels/255) * 100 * 10;
```

If percentage is greater than threshold value then it is iris mismatch otherwise it is considered as iris matched and the user is authenticated.

5 Results and Discussions

For the experimental purpose we have tested the system for 120 users. Now-a-days everyone has the basic knowledge of how to use the web based application. Still while using technology, comfort can be one of the key concerns in order to state the user friendliness of the application. Thus we have made 4 groups with 30 users in each group. Here users are the faculties of different age group and different background. The group details are as follows:

- Group A: contains users who are the faculties having age in between 27 and 42 years and belong to non-technical stream like Literature, Arts, Communication skill and Chemistry.
- Group B: contains users who are the faculties having age in between 27 and 42 years and belong to technical stream like Engineering, Polytechnic and Management.
- Group C: contains users who are the faculties having age in between 43 and 57 years and belong to non-technical stream like Literature, Arts, Communication skill and Chemistry.
- Group D: contains users who are the faculties having age in between 43 and 57 years and belong to technical stream like Engineering, Polytechnic and Management.

Table 3 shows the experimental results for users of different group based on age and their stream. For experimentation every group has used computer system and web camera of same configuration and same Internet speed. We have considered the standard environment with enough luminance and no tilt in iris image. Still it has been observed that there are few users whose iris matching is not done thus they

Table 3 Experimental results for users of different group based on age and their stream

Criteria	*A*	*B*	*C*	*D*
Average Time required for Enrollment (Sec)	4.38	3.89	5.99	4.46
Average Time required for Voting (Sec)	12.73	11.45	16.96	14.01
No. of user proceeding further with first OTP	24	25	19	25
No. of user not requesting OTP twice	02	02	03	00
No. of people not requesting OTP thrice or even more (max. 5 times)	00	00	01	01
No. of user for whom iris recognition is done in first attempt	16	21	10	18
No. of user for whom iris recognition is done in second attempt	08	06	07	03
No. of user for whom iris recognition is done in third attempt	02	00	06	05
No. of user for whom iris recognition is not done thus not casted the vote	04	03	07	04
No. of user casted the vote successfully	26	27	23	26

could not complete the voting process. We also found that technical expertise of the person and how well-versed the person is in dealing with web application and navigation has an impact on the total time required to complete the enrollment as well as voting process.

In second phase of experiment we have considered two groups viz. E and F each having 30 users. For group E, we asked the users to enroll during night with tilts and for group F enrollment is done during day time with tilts while voting for both the groups is carried out during day time. Here all the users are technically sound and between the age group of 27–48. For group E out of 30, only 19 users successfully casted the vote. For rest of the 11 users, biometric authentication failed. The reasons behind that is template iris features for this group is extracted and stored from image which is taken in night when luminance was poor. Thus the luminance plays an important role in the performance of system. For group F, out of 30, 23 people successfully did the iris authentication and casted vote. Thus even tilt in the iris image angle hampers the accuracy of the system.

Figure 7 gives the success and failure rate for all six groups. Same statistics can be considered for iris recognition accuracy and error rate of the system of the system. Thus the average accuracy of the system is 81.11% and average failure rate is 18.89%.

In Table 4, we have discussed the advantages of our proposed system over the existing system which are described in Sect. 2.1. In our system, we have completely eliminated the overhead of maintaining magnetic tape based voter ID cards or smart cards. This reduces the time as well as money required to complete the whole election process. Also there is no need for the voter to maintain these cards. We have also reduced the burden to generate and remember various security keys in order to use the system. Most of the researchers have used fingerprints for biometric

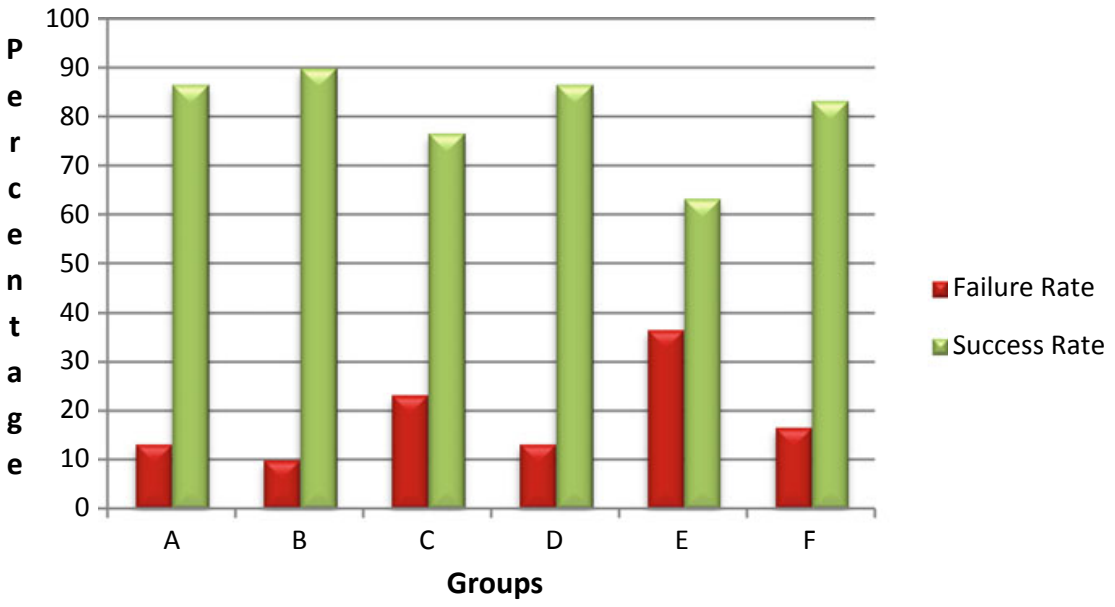


Fig. 7 Summary of success and failure rate for all five groups

Table 4 Advantage of proposed system over existing system

S. No.	Existing system discussed in Table 1	Advantage of proposed system over existing system
1	Khasawneh et al. [2]	No overhead of magnetic tape based Voter ID card, cost is low
2	Petcu et al. [3]	Easy to use as there is no need to generate and remember different keys
3	Sridharan [4]	Smart cards are not required to use the system
4	Agarwal et al. [5]	Biometric data storage is tamper proof and whole system is a decentralized, integrated and transparent

authentication in their system. Instead of fingerprints; we have used iris for biometric authentication as fingerprints are more vulnerable to attacks as compared to iris.

6 Conclusion and Future Work

This paper elaborates a technique of online voting empowered with biometric authentication and Blockchain Technology. By using iris recognition system, we make sure that only a verified person can cast the vote. Also as compared to the existing or proposed voting systems, our system does not store the captured iris image in the database. Instead, it extracts the features of the iris image and then removes the iris image from memory which makes the biometric data more secure. These extracted features along with all the other information of the user and the votes casted by them are stored on Blockchain. Thus system becomes transparent, decentralized and mishaps like information tampering or vote alternation can be easily detected. As now the average accuracy of the system is 81.11% and average failure rate is 18.89%. Accuracy rate is low because we have performed the experiment in low luminance and iris tilt variation. Thus capturing proper iris image with minimum noise is the prime requirement of the system. In future we are planning to improvise our iris recognition system, so that the overall system should work under non ideal condition without any constraint.

7 Acknowledgement

This research was carried out as an extension of final year dissertation titled “Fingerprint Authentication System Using Blockchain” for the degree of Bachelor in Information Technology of University of Mumbai, under the guidance of Prof. Dipti Pawade and Prof. Avani Sakhapara, at K.J. Somaiya College of Engineering, Mumbai, India. We express our gratitude to the Institute for its constant support and guidance. This research was carried out by adhering to the applicable ethical

standards and in accordance with the rules and regulations of the Institute. For this research, prior consent was taken from the participants to publish the results of research in public domain ensuring complete anonymity and confidentiality of the data.

References

1. http://eci.nic.in/eci_main1/evm.aspx
2. Khasawneh, K., Malkawi, M., Al-Jarrah, O., Barakat, L., Hayajneh, T.S., Ebaid, M.S.: A biometric-secure e-voting system for election processes. In: Proceeding of the 5th International Symposium on Mechatronics and its Applications (ISMA08), Amman, Jordan, May 27–29, 2008
3. Petcu, D., Stoichescu, D.A.: A hybrid mobile biometric-based e-voting system. In: 9th International Symposium on Advanced Topics in Electrical Engineering (ATEE), Bucharest, Romania, 7–9 May 2015, pp 37–42
4. Sridharan, S.: Implementation of authenticated and secure online voting system. In: 4th Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT) 2013, Tiruchengode, India No. 6, July 2013. IEEE—31661
5. Agarwal, H., Pandey, G.N.: Online voting system for India based on AADHAAR ID. In: Eleventh International Conference on ICT and Knowledge Engineering, Bangkok, Thailand, 20–22 Nov 2013
6. Garagad, V.G., Iyer, N.C.: A novel technique of iris identification for biometric systems. In: International Conference on Advances in Computing, Communications and Informatics (ICACCI), New Delhi, India, 24–27 Sept 2014
7. Ahmadi, Neda, Akbarizadeh, Gholamreza: Hybrid robust iris recognition approach using iris image pre-processing, two-dimensional gabor features and multi-layer perceptron neural network/PSO. IET Biom. **7**(2), 153–162 (2018)
8. Wildes, R.P., Asmuth, J.C., Green, G.L., Hsu, S.C., Kolczynski, R.J., Matey, J.R., McBride, S.E.: A machine-vision system for iris recognition. Mach. Vis. Appl. **9**(1), 1–8 (1996)
9. Ahmadi, N., Akbarizadeh, G.: Iris recognition system based on canny and LoG edge detection methods. J. Soft Comput. Decis. Support Syst. **2**(4), 26–30 (2015)
10. Daugman, J.: High confidence recognition of persons by rapid video analysis of iris texture. In: IET European Convention on Security and Detection, pp. 244–251 (1995)
11. Hamd, Muthana H., Ahmed, Samah K.: Biometric system design for iris recognition using intelligent algorithms. Int. J. Modern Educ. Comput. Sci. (IJMECS) **10**(3), 9–16 (2018). https://doi.org/10.5815/ijmecs.2018.03.02
12. David S., Harvey L., Vimi G., Alexander S., Stephen M., Tyler W.: What is blockchain? Blockchain Enigma. Paradox. Opportunity – Deloitte, 2016
13. Sloane B., Bhargav P.: Blockchain basics: introduction to distributed ledgers (2016). https://www.ibm.com/developerworks/cloud/library/cl-blockchain-basics-intro-bluemix-trs/index.html
14. Pawade, D., Jape, S., Balasubramanian, R., Kulkarni, M., Sakhapara, A.: Distributed ledger management for an organization using blockchains. Int. J. Educ. Manag. Eng. (IJEME) **8**(3), 1–13 (2018). https://doi.org/10.5815/ijeme.2018.03.01
15. https://www.ethereum.org/
16. https://metamask.io/
17. Pawade, D., Pawade, D., Sakhapara, A., Andrade, M., Badgujar, A., Adepu, D.: Implementation of fingerprint based authentication system using blockchain. In: The International Conference on Soft Computing and Signal Processing (ICSCSP-2018), Hyderabad, 22–23 June 2018

18. Pawade, D., Chaudhari, P., Sonkambale, H.: Comparative study of different paper currency and coin currency recognition method. Int. J. Comput. Appl. **66**(23), 26–31 (2013)
19. Adhau, A.S., Shedge, D.: Iris recognition methods of a blinked-eye in non-ideal condition. In: IEEE International conference on Information Processing, pp. 75–79. ISBN: 4673-7758 (2015)
20. Puhan, N.B., Sudha, N., Xia, H., Jiang, X.: Iris recognition on edge maps. In: IET Computer Vision, 5th September 2007, https://doi.org/10.1049/ietcvi:20080015
21. https://github.com/yoga1290/Fingerprint-Recognition

An Approach: Applicability of Existing Heterogeneous Multicore Real-Time Task Scheduling in Commercially Available Heterogeneous Multicore Systems

Kalyan Baital and Amlan Chakrabarti

Abstract Interest in design and use of heterogeneous multicore architectures has been increased in recent years due to the fact that the energy optimization and parallelization in heterogeneous multicore architecture are better than that of homogeneous multicore architecture. In heterogeneous multicore architectures, cores have similar Instruction Set Architecture (ISA) but the characteristics of the cores are different with respect to power and performance. Hence, heterogeneous architecture provides new prospects for energy-efficient computation and parallelization. Heterogeneous systems, furnished with different types of cores provide the mechanism to take actions with respect to irregular communication patterns, energy efficiency, high parallelism, load balancing, and unexpected behaviors. However, designing such heterogeneous systems for the different platforms like cloud, Internet of Things (IoT), Smart Devices, and Embedded Systems is still challenging. This paper studies the commercially available heterogeneous multicore architectures and finds out an approach or method to apply the existing work on heterogeneous multicore real-time task scheduling model to commercially available heterogeneous multicore architecture to achieve the parallelism, load balancing, and maximum throughput. The paper shows that the approach can be applied very efficiently to some of the commercially available heterogeneous systems to establish a generic heterogeneous model for the platforms like cloud, Internet of Things (IoT), Smart Devices, Embedded Systems, and other application areas.

Keywords Big core · big.LITTLE · Heterogeneous · Multicore · Octa-core · Real-time task · Small core · Xeon Phi

K. Baital (✉)
National Institute of Electronics and Information Technology, Kolkata Centre, Kolkata, India
e-mail: kalyan_baital@yahoo.co.in

A. Chakrabarti
A. K. Choudhury School of Information Technology, University of Calcutta, Kolkata, West Bengal, India
e-mail: amlanc@ieee.org

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_8

1 Introduction

Heterogeneous multicore systems are increasing day by day in the field of embedded systems as well as cloud environments to improve system throughput, parallelization, load balancing, and energy efficiency. Heterogeneous multicore system includes cores with various types or/and complexities to address throughput and productivity for different applications. Different types of heterogeneous multicore systems based on different application areas are commercially available in the market to offer a high degree of parallelism and energy efficiency. Thus, there is an increasing motivation to design a generic heterogeneous model for different application domains including the following technological domains which are in high demand in the market today:

i. Cloud and IoT infrastructure which is high-performance and high-resource domain and offering a high degree of parallelism.
ii. Smart devices and embedded domain which is low-performance and low-resource domain but offering great energy efficiency.

Therefore, our focus is to find an approach to establish a generic heterogeneous model which can be utilized as high-performance and high-resource domain (i.e., cloud and IoT). The same generic model can also be used as energy efficient and low resource domain (i.e., smart devices and embedded systems) along with other common application areas.

Heterogeneous multicore system has achieved significant importance in the area of High-Performance Computing (HPC) which offers a higher degree of parallelization, and hence it has enormous potential for future use in data centers and cloud platforms. In this regard, heterogeneous architectures such as Intel Xeon Phi coprocessor and Graphics Processing Unit (GPU) have significant roles in HPC systems.

Recently, heterogeneous architecture is being used in the mobile technology such as ARM big.LITTLE octa-core architecture. The usage pattern for smartphones and tablets varies drastically, for example, high-processing usage likes mobile gaming and web browsing and low-processing usage likes texting, e-mail, and audio. Octa-core heterogeneous architecture such as ARM big.LITTLE fulfills the demand of these different performance requirements by combining two different types of cores.

Hence, in our proposed approach, we apply different architectural parameters of Intel Xeon Phi coprocessor in our proposed generic model such a way that the generic model can additionally be used as cloud and IoT platform along with several common application areas. Further, we apply different architectural parameters of octa-core (ARM big.LITTLE) in the proposed model for using it as smart devices and embedded systems.

The following significant contributions are explained in this paper:

1. To study the architecture of some of the commercially available heterogeneous multicore architecture such as Intel Xeon–Xeon Phi platform and big.LITTLE mobile octa-core platform.
2. To find an approach to apply the existing heterogeneous multicore real-time task scheduling model as per work [1] into the commercially available heterogeneous multicore systems (Intel Xeon–Xeon Phi platform and big.LITTLE mobile octa-core platform) to achieve the parallelism, load balancing and maximum throughput.
3. To provide the performance comparisons between commercially available heterogeneous multicore systems (Intel Xeon–Xeon Phi platform and big.LITTLE mobile octa-core platform) and existing heterogeneous multicore real-time system [1] to highlight the achievements.

The organization of this paper is as follows: Sect. 2 explains the review of literature for commercially available heterogeneous multicore architecture, namely, Intel Xeon–Xeon Phi platform and big.LITTLE mobile octa-core platform and related works. In this section, it has been shown that (i) Intel Xeon–Xeon Phi system is a good candidate to be used in cloud and IoT infrastructure besides other application domains and (ii) octa-core system is a good candidate to be used in smart devices and embedded systems domain. Section 3 introduces an approach for applicability of the existing work [1] to commercially available heterogeneous multicore architecture such as Intel Xeon–Xeon Phi platform and big.LITTLE mobile octa-core platform. The section explains how the existing model [1] has been established as a generalized platform, where it can be used as a Xeon–Xeon Phi cluster in the cloud and IoT infrastructure as well as big.LITTLE mobile octa-core model in smart devices and embedded system domain. Section 4 summarizes our contribution.

2 Literature Review of Commercially Available Heterogeneous Multicore Systems and Related Works

Several heterogeneous multicore systems are available commercially in the market, namely, Intel Xeon processor–Intel Xeon Phi coprocessor multicore system platform, mobile octa-core processor, GPU multicore processor, etc. They are designed to be operated at specific types of application areas. Some of them are best suited for high-performance areas, some of them are best suited for graphics and image processing areas, and some of them are best suited for providing energy efficiency.

However, we study two heterogeneous multicore architectures, namely, (A) Intel Xeon processor–Intel Xeon Phi coprocessor platform and (B) Mobile octa-core processor due to the demand in current market scenarios as explained in Sect. 1. The architectures of both the systems are described below.

2.1 Intel Xeon Processor–Intel Xeon Phi Coprocessor Platform

In recent development of computer architecture, multicore is inclined toward many-core in high-performance computing applications and Intel Many Integrated Core (MIC) coprocessor which is known as Xeon Phi coprocessor is rising as single-chip many-core processor for high-performance computation [2]. The scaling capability of Intel Xeon processor-based system is increased with the association of Intel Xeon Phi coprocessor which offers additional power efficient scaling, vector support, memory bandwidth and maintains programmability and support associated with Intel Xeon processor [3]. Hence, Intel Xeon and Xeon Phi are generally integrated into a single platform and generally used for highly parallel applications which reach the scaling limits on Intel Xeon processors. A typical block diagram of the platform as per work [3] is shown as Fig. 1.

As per the existing work [3, 4, 5, 6, 7] the following architecture of the platform consisting of Intel Xeon and Intel Xeon Phi may be summarized:

- A platform cannot have only coprocessors.
- A typical platform consist of (i) 1–2 Intel Xeon processors (with a certain numbers of cores) with high frequency for the execution of highly serial tasks and (ii) 1–8 Intel Xeon Phi coprocessors per host with low frequency for execution of highly parallel tasks.
- Each coprocessor has around 60 cores (4 hardware thread on each core).
- All the cores composed of two levels of caches (L1 and L2).
- The Xeon Phi is attached with a host CPU through Peripheral Component Interconnect (PCI) express just like a Graphics Processing Unit (GPU).
- Both the Intel Xeon processor and Intel Xeon Phi coprocessor can be utilized by the specific application to achieve a higher degree of performance and parallelization. However, Xeon Phi coprocessor and Xeon CPU are different in performance characteristics with each other. The coprocessor may contribute

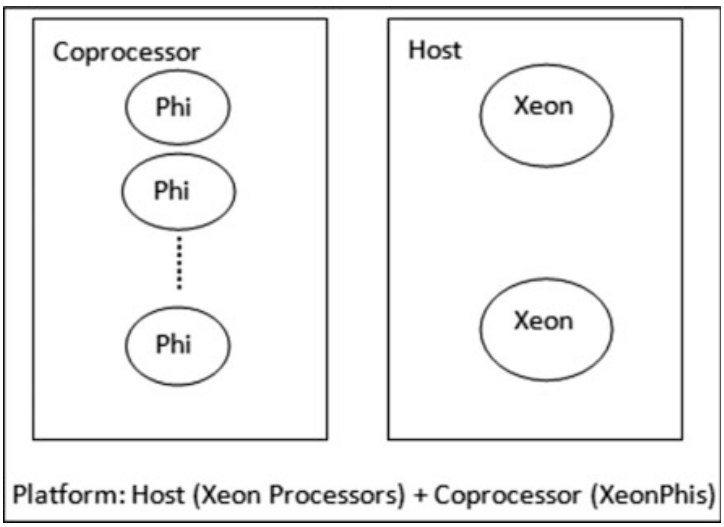


Fig. 1 A block diagram of platform consisting of Intel Xeon processor and Intel Xeon Phi coprocessor

considerably higher performance than the host processor for some applications, while the performance may be low for other applications.

- Preparing Intel Xeon Phi coprocessors should be done such a way that at first the performance of the Xeon Processor can be utilized fully for a given application. Additionally scaling of an application is very important to get the high performance and the application has to scale more than one hundred threads to achieve high parallelism.

Many research works are proposed to exploit the architecture of Xeon Phi to fit several applications. In paper [8], the authors presented new parallelization techniques that were used to search the biological sequence database in Xeon Phi-based architecture. The authors in paper [9] proposed a technique to enable the capabilities of Intel Xeon Phi for a virtual machine. In paper [10], the authors executed a well-known genomic aligner Burrows-Wheeler Aligner (BWA) in a heterogeneous system consisting of Xeon multicore processor and Xeon Phi manycore accelerator. More research work considering data analysis, performance, load balancing can be found in [11, 12, 13, 14].

From the above architecture on Intel Xeon–Xeon Phi system and all the existing research works based on the system, it may be concluded that Intel Xeon–Xeon Phi system is a good candidate to be used in cloud and IoT infrastructure besides other application domains for providing a high degree of parallelism and performance.

2.2 Mobile Octa-Core Processor

The functions of traditional desktop applications are steadily being taken over by mobile devices [15], and the end user uses it more closely than the Personal Computer (PC). Mobile devices have limited battery power and therefore, the thermal issue is more critical in mobile devices than the PC. More power is not required for a majority of tasks such as navigating through home screens, checking texts, and browsing the Web. However, power requirements become higher for tasks like HD video and gaming. Octa-core concept extends the battery life of modern handsets without compromising performance. Octa-core represents two separate operating setups of quad core integrated together on a single-chip for better energy efficiency. As per the work [16], it may be opinioned that ARM's big. LITTLE architecture is the basis of all octa-core mobile chips. As explained in the work [17], the architecture enables four (4) low-power Cortex-A7 cores (LITTLE cores) to operate with four (4) high-performance Cortex-A15 cores (big cores) and a simple block diagram as explained in the work [17] is shown as Fig. 2.

As per existing work [18, 19, 20] the following big.LITTLE architecture of ARM (mobile octa-core architecture) may be summarized:

- big.LITTLE octa-core architecture had been introduced to enhance the power optimization and performance simultaneously. big.LITTLE architecture consists

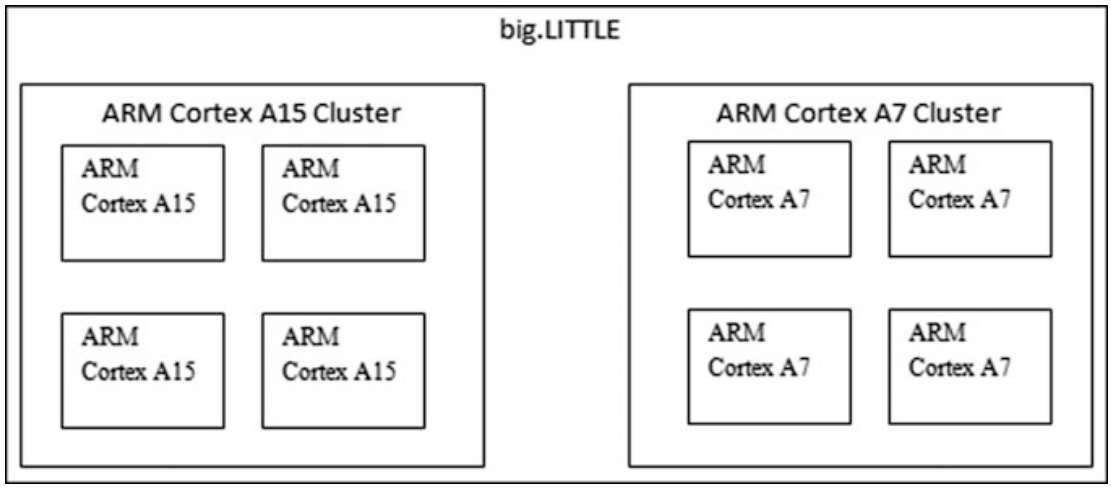


Fig. 2 Block diagram of big.LITTLE octa-core architecture of ARM

of performance-driven big core (but consumes more power) and power efficiency-driven LITTLE core (but delivers lower performance).

- Octa-core chips function as dual quad cores where tasks are balanced between them according to type. The low powered little set of cores is employed mostly while the faster big set of cores are used for advanced tasks. Different tasks have different performance and power requirements, and hence the tasks are required to be allocated to the correct processor according to their requirement.
- There are two major task scheduling techniques in big.LITTLE models as per below:

  (1) ***CPU Migration***: Here, each LITTLE core is paired with a big core and at any point of time, any single core in each pair is operational. The inactive core is powered off.

  (2) ***Global Task Scheduling***: The global scheduler is aware of the different capacities of big and LITTLE cores. The global scheduler keeps track of the performance requirement for each individual thread and decides which type of processor to utilize for each thread. Unused processors can be powered down. Global task scheduler is a flexible and popular model for all future development.

- Figure 3 shows power and performance and operating points of big (Cortex-A15) and LITTLE (Cortex-A7) processors [21]. As shown in *Fig.* 3, LITTLE core is more power efficiency than big core but performance is restricted than big core. Big core, on the other hand, provides high performance but with high power consumption. Altogether, big.LITTLE architecture in mobile environment can satisfy these two contradictory requirements—high performance with long battery life.

Many research works are found in the literature to use the big.LITTLE architecture in different applications. Paper [22] addressed the usage of low-power heterogeneous multicore such as ARM big.LITTLE architecture as micro servers

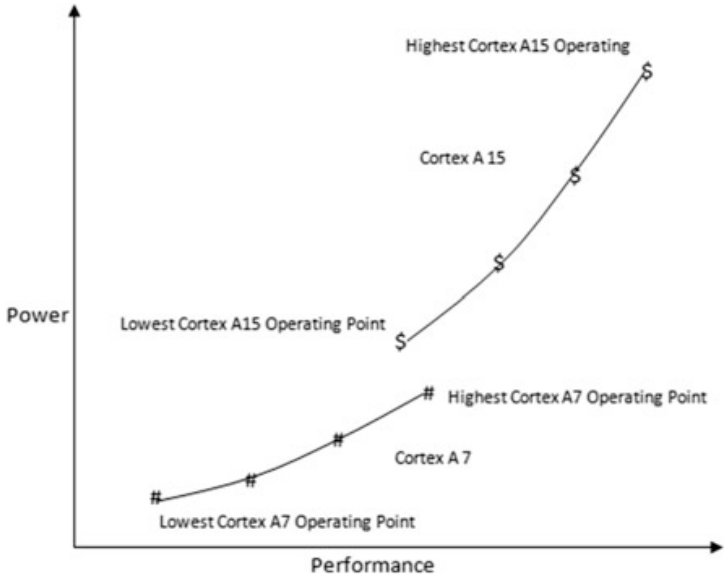


Fig. 3 Power and performance curve of ARM's big.LITTLE (Cortex-A15. Cortex-A7) architecture (from [21])

which were used as a web search application. The authors in [23] introduced hipster which is a method to merge heuristic and reinforcement learning to deal with latency critical workload using the ARM big.LITTLE architecture. More research work on big.LITTLE architecture considering the parameters of energy and performance can be found in [24, 25].

From the above architecture on octa-core system (big.LITTLE architecture) and all the existing research works based on the system, it may be concluded that octa-core system is a good candidate to be used in smart devices and embedded systems domain for providing greater energy efficiency with low resources.

From all the existing work on Xeon Phi as well as big.LITTLE octa-core architectures, it can be concluded that in near future there is a need to have a generalized heterogeneous multicore platform for execution of different applications in Xeon–Xeon Phi platform in the cloud and IoT infrastructure besides other application domains as well as mobile big.LITTLE octa-core architecture in the smart devices and embedded systems domain.

3 Applicability of Existing Work of Heterogeneous Multicore Real-Time Task Scheduling into Commercially Available Heterogeneous Multicore System

The objective of our work is to analyze the existing heterogeneous multicore scheduling model [1] and to establish the model as a generalized platform where it can be used as a Xeon–Xeon Phi cluster as well as big.LITTLE mobile octa-core model. The heterogeneous architecture of work [1] illustrated a dynamic real-time task scheduling model and the model as depicted in the work is shown as Fig. 4.

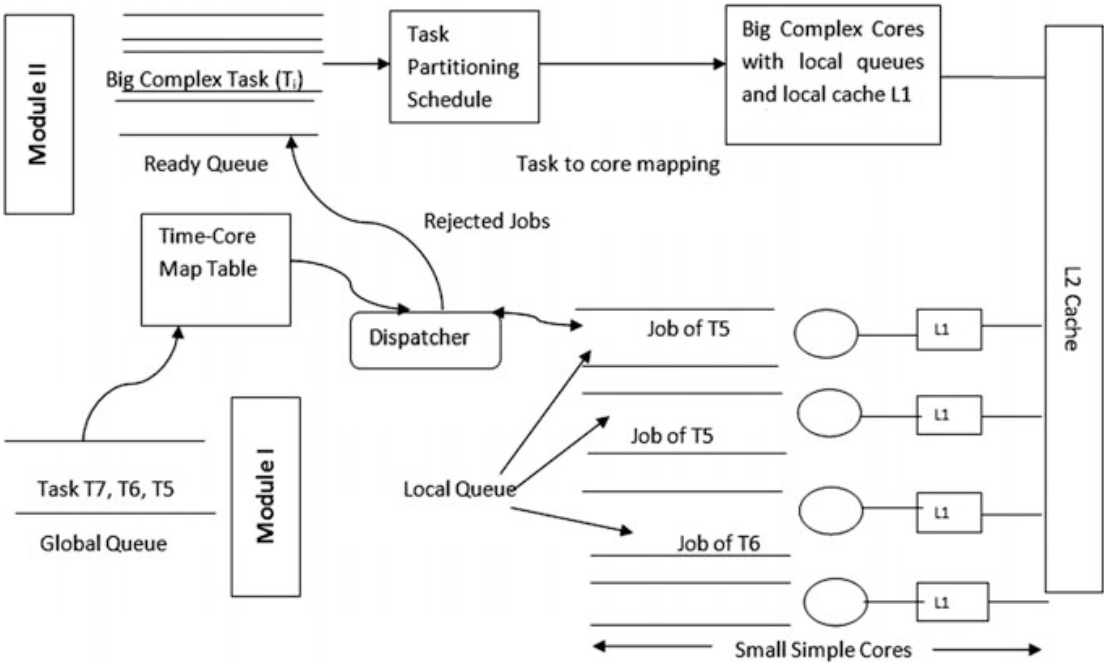


Fig. 4 Heterogeneous multicore real-time task scheduling model with different types of cores (from [1])

The heterogeneous architecture as explained in the work [1] may be summarized as per below:

- The architecture has two modules (module-I and module-II) along with a dispatcher to dispatch the matching or/and rejected job to/from particular core.
- The module-I is composed of low power small cores that are designed for the execution of low-power tasks to provide higher parallel system performance. The module-I is functional within a low-power range P (*where* $0 < P \leq P_l$ *and* P_l = *maximum power to activate the small cores*). The module has a global queue, where calculations of pseudo-release times of jobs (instances of low-power tasks) are completed and the time-core map table is searched with the calculated pseudo-release times. Consequently, run-time dispatcher forward the matching jobs to local queues of the corresponding cores. Subsequently, the jobs are assigned to cores from local queues meeting the deadline.
- The module-II is composed of high-power complex big cores that are incorporated for the execution of high-power tasks to provide higher serial system performance. The module-II is functional above a high power P (*where* $P_l < P \leq P_m$ *and* P_m = *maximum power to activate the cores*). The task instances (or jobs) of a big high-power task are stored in ready queue of particular big core. The rejected jobs from module-I are also accumulated in ready queues of module-II. Next task partitioning scheduling is used to forward all the task instances (or jobs) from ready queue to local queue and consequently into the corresponding big core.

- The model has C number of cores altogether, where $C = C_s + C_b$ (C_s = *Number of small cores;* C_b = *Number of big cores*). All the cores (small and big) consist of two levels of caches—L1 (own first level) and L2 (shared by all the cores). The work explained that low/high-power tasks are allocated to their corresponding small/big cores efficiently as per the requirements of serial and parallel performance.

We introduce an approach to apply this model [1] as Intel Xeon–XeonPhi cluster and mobile octa-core processors (big.LITTLE).

3.1 Applicability as Intel Xeon–Xeon Phi Multicore Platform

The high power big cores in module-II of work [1] are designed for highly serial tasks and can be treated Intel Xeon processors as the architecture of these cores are exactly similar to that of Xeon processors. The number of big cores of work [1] may be taken as n, i.e., $C_b = n$, *where* $n = i*j$; i = number of cores per Intel Xeon processor and j = number of Intel Xeon processor ($j \leq 2$).

Similarly, the low power small cores in module-I of work [1] are designed for highly parallel tasks and can be treated as Xeon Phi coprocessors as the architecture of these cores are exactly similar to that of Xeon Phi coprocessors. The number of small cores of work [1] may be taken as m, i.e., $C_s = m$, *where* $m = 60*k$; k = number of coprocessors ($k \leq 8$) and number of cores per Intel Xeon Phi coprocessor = 60.

Complete architecture of the model can act as a platform consisting of Intel Xeon processor and Intel Xeon Phi coprocessor that can be used in cloud and IoT infrastructure besides other application domains.

In the work [1] it has been shown that the task allocation percentage of the model is approximately near to 100% at a fixed time period. The model further shows that the CPU utilization% is very high (between 95 and 100%) when the number of cores is beyond four (4). Further, we compared the average utilization% of the model [1] with the average performance in terms of utilization of existing work [3] related to Xeon and Xeon Phi cluster. For the comparison, we considered average system utilization% with respect to a different number of cores which we get by varying the number of cores of the model [1] and by varying the number of threads and subsequently calculating the number of cores of the work [3]. The number of cores for the work [3] has been calculated from the assumption that each Xeon Phi core has been assigned four numbers of threads and each Xeon core has been assigned two numbers of threads as explained in the work [3].

We plotted the comparison result as shown in Fig. 5. From Fig. 5, we found that performance in terms of utilization for Xeon processor reached maximum (nearly 100%) when the number of cores was beyond eight (8) and for Xeon phi coprocessor utilization% reached high when the number of cores was beyond thirty-two

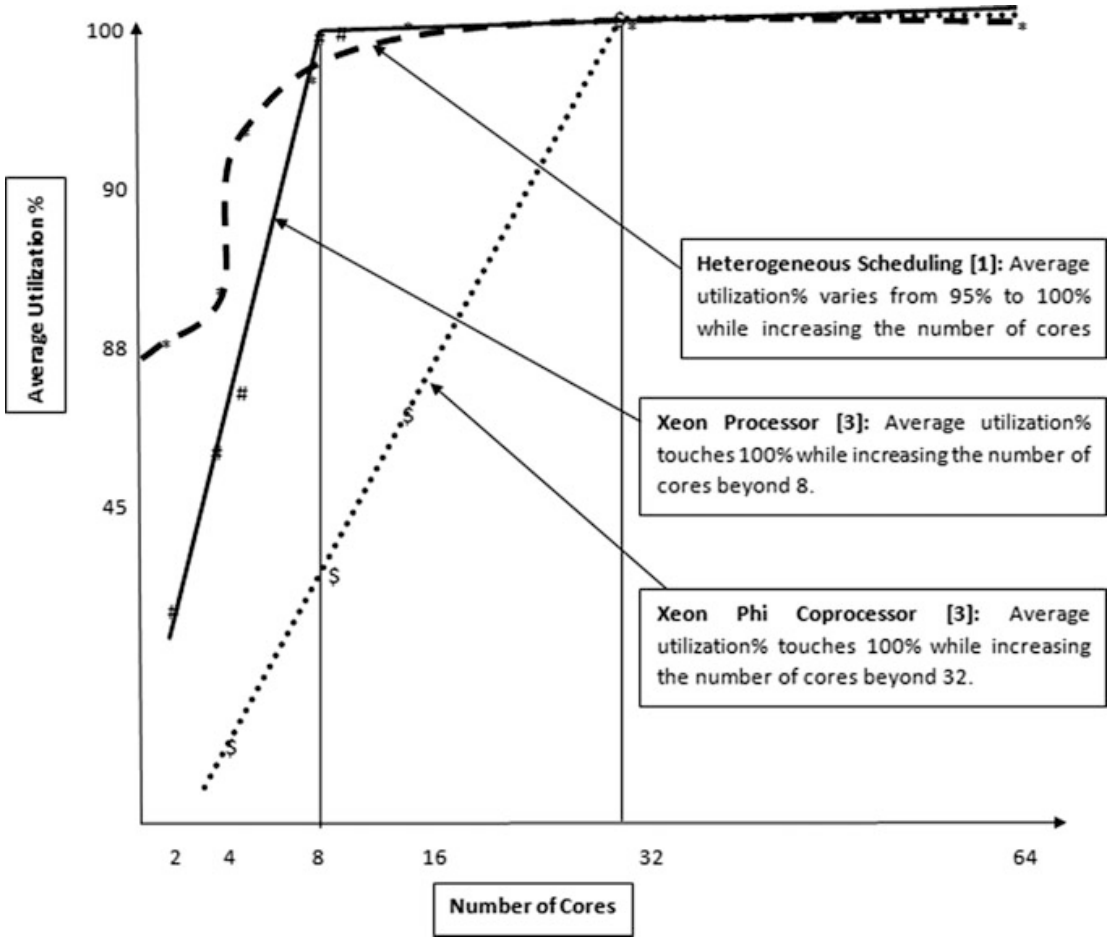


Fig. 5 Comparison of average system utilization: performance in terms of utilization for Xeon processor reached maximum (100%) when the number of cores was beyond eight and for Xeon phi coprocessor utilization% reached high when the number of cores was beyond 32. In the work [1], the average system utilization varied from 95% to 100% while increasing the number of cores beyond four. Hence, we can say that the model [1] can exactly fit into the system consisting of Intel Xeon processor and Intel Xeon Phi coprocessor

(32). Also, we found in the work [1], that the average system utilization varied from 95% to 100% while increasing the number of cores beyond four (4) only. Hence, we can say that the model [1] can exactly fit into the system consisting of Intel Xeon processor and Intel Xeon Phi coprocessor.

3.2 Applicability as Mobile Octa-Core Processor

The work of [1] exactly follows the architecture of ARM's big.LITTLE mobile octa-core architecture where each type of core is operational between certain power levels to optimize power and performance as explained in Fig. 3. The number of big cores (module-II) as well as small cores (module-I) of work [1] may be taken as four,

i.e., $C_s = C_b = 4$, so that the complete architecture of the model becomes the mobile octa-core processors that can be used in smart devices and embedded systems.

The work [1] followed the global task scheduling model of ARM's big.LITTLE architecture where a particular type of task including rejected task (small or big) was allocated to a particular type of core (small or big) through dispatcher maintaining constraints as explained previously.

The work [1] demonstrated that tasks of different types are allocated to cores of different types efficiently for optimization of power and performance maintaining very high task allocation percentage (100%) and CPU utilization percentage (varied from 95 to 100%). Further, we compared the job allocation% of the model [1] with the throughput in terms of job allocation% of existing work [26] related to mobile octa-core architecture (big.LITTLE). For comparison, we considered job allocation % with a fixed time period which we get by illustrating the variation in a number of jobs getting accommodated at a fixed time period.

We plotted the comparison result as shown in Fig. 6. From Fig. 6, we found that both the model accommodated 100% of the random task instances within a fixed time period. Hence, existing heterogeneous real-time task scheduling architecture [1] may be used as mobile octa-core architecture (big.LITTLE).

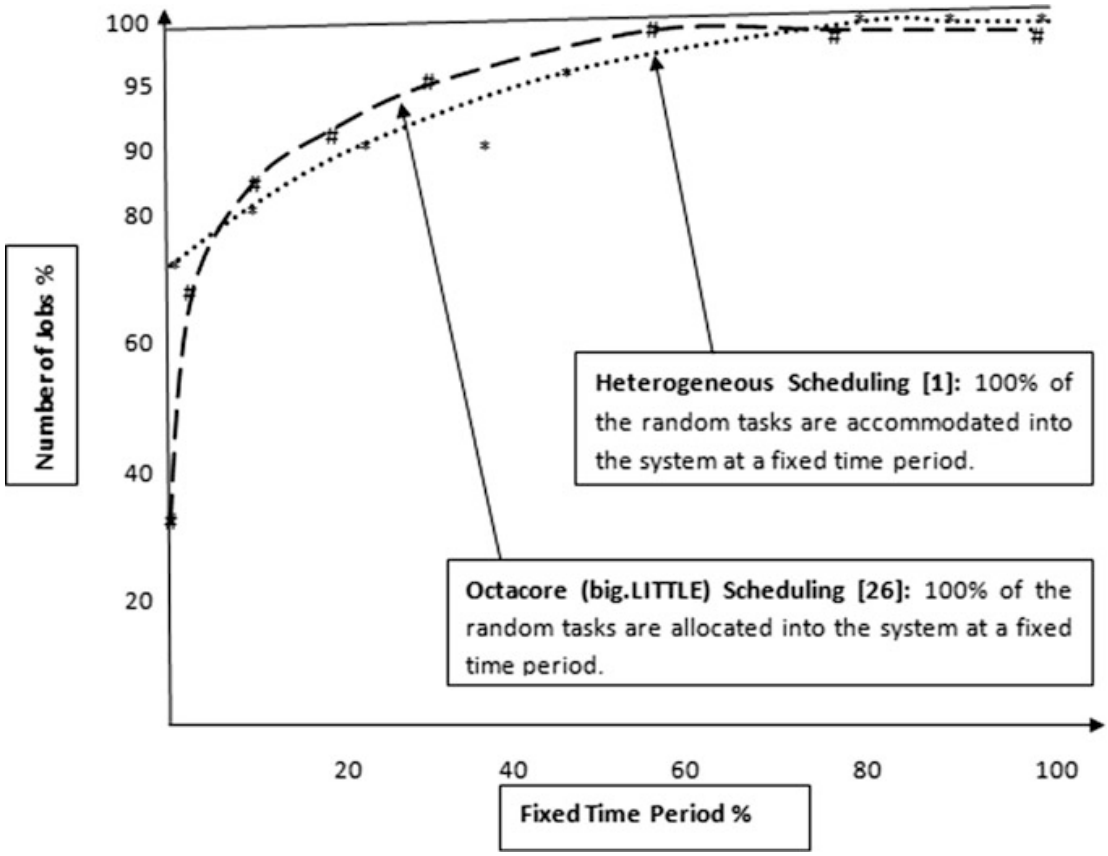


Fig. 6 Comparison of job allocation percentage between existing work [1] and mobile octa-core processor (big.LITTLE). The existing work explained that 100% of the random tasks are accommodated into the system at a fixed time period. As per work [26] 100% jobs are allocated in the ARM architecture of mobile octa-core

Therefore, from the above analysis considering the architectures of Xeon–Xeon Phi and big.LITTLE we can conclude that the existing heterogeneous multicore scheduling model [1] can be used as Intel Xeon–Xeon Phi platform in the cloud and IoT infrastructure besides other application domains as well as big.LITTLE mobile octa-core architecture in smart devices and embedded system domain maintaining the same power-performance benefit and parallelism.

4 Conclusion

This paper studies different parameters of commercially available heterogeneous multicore systems, namely, Intel Xeon–Xeon Phi cluster and mobile octa-core architecture. It also studies the scheduling algorithm for real-time tasks in existing heterogeneous multicore scheduling model [1]. It further finds out an approach or method to establish the existing heterogeneous system model [1] as a generic platform so that the generic model can be used as Intel Xeon–Xeon Phi processors and mobile octa-core processors as per requirements maintaining the same performance and parallelism.

The comparison results considering different parameters of existing model with Intel Xeon–Xeon Phi cluster and mobile octa-core architecture reveal that the existing model [1] can be applied as both the commercially available heterogeneous systems (Intel Xeon–Xeon Phi platform and mobile octa-core processors) for the use in cloud, IoT, smart devices, embedded systems, and other common application areas efficiently. Significant issues like resource sharing, energy enhancement, and dependency parameters will be incorporated in the existing model [1] to apply it as a more generic model.

References

1. Baital, K., Chakrabarti, A.: Dynamic scheduling of real-time tasks in heterogeneous multicore systems. IEEE Embed. Syst. Lett. **11**(1), 29–32 (2019)
2. Jha, S., He, B., Lu, M., Cheng, X., Huynh, H.P.: Improving main memory hash joins on Intel Xeon Phi processors: an experimental approach. Proc. VLDB Endow. **8**(6), 642–653 (2015)
3. Reinders, J.: An Overview of Programming for Intel® Xeon® processors and Intel® Xeon Phi™ Coprocessors, pp. 1–21. Intel Corporation, Santa Clara (2012)
4. Fang, J., Sips, H., Zhang, L., Xu, C., Varbanescu, A.L.: Test-Driving Intel Xeon Phi. In: Proceedings of 5th ACM/SPEC International Conference on Performance Engineering, NY, USA, 2014, pp. 137–148
5. Coviello, G., Cadambi, S., Chakradhar, S.: A coprocessor sharing-aware scheduler for Xeon Phi-based compute clusters. In: Proceedings of IEEE 28th International Parallel and Distributed Processing Symposium, Phoenix, AZ, 2014, pp. 337–346
6. Miyamoto, T., Ishizaka, K., Hosomi, T.: A dynamic offload scheduler for spatial multitasking on Intel Xeon Phi coprocessor. In: Proceedings of SASIMI 2013, pp. 261–266

7. Hirokawa, Y., Boku, T., Sato, S.A., Yabana, K.: Electron dynamics simulation with time-dependent density functional theory on large scale symmetric mode Xeon Phi cluster. In: Proceedings of IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), Chicago, IL, 2016, pp. 1202–1211
8. Lan, H., Liu, W., Schmidt, B., Wang, B.: Accelerating large-scale biological database search on xeon phi-based neo-heterogeneous architectures. In: Proceedings of IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Washington, DC, 2015, pp. 503–510
9. Gerangelos, S., Koziris, N.: vPHI: enabling xeon phi capabilities in virtual machines. In: Proceedings of IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), Lake Buena Vista, FL, 2017, pp. 1333–1340
10. Chen, S., Senar, M.A.: Improving performance of genomic aligners on intel xeon phi-based architectures. In: Proceedings of IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), Vancouver, BC, 2018, pp. 570–578
11. Xie, B., Liu, X., Zhan, J., Jia, Z., Zhu, Y., Wang, L., Zhang, L.: Characterizing data analytics workloads on Intel Xeon Phi. In: Proceedings of IEEE International Symposium on Workload Characterization, Atlanta, GA, 2015, pp. 114–115
12. Nookala, P., Dimitropoulos, S., Stough, K., Raicu, I.: Evaluating the support of MTC applications on Intel Xeon Phi many-core accelerators. In: Proceedings of IEEE International Conference on Cluster Computing, Chicago, IL, 2015, pp. 510–511
13. Pennycook, S.J., Hughes, C.J., Smelyanskiy, M., Jarvis, S.A.: Exploring SIMD for molecular dynamics, using Intel Xeon processors and Intel Xeon Phi coprocessors. In: Proceedings of IEEE 27th International Symposium on Parallel and Distributed Processing, Boston, MA, 2013, pp. 1085–1097
14. Misra, S., Pamnany, K., Aluru, S.: Parallel mutual information based construction of genome-scale networks on the Intel Xeon Phi™ coprocessor. In: IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 12, no. 5, pp. 1008–1020, 1 September–October 2015
15. Gao, C., Gutierrez, A., Rajan, M., Dreslinski, R.G., Mudge, T., Wu, C.: A study of mobile device utilization. In: Proceedings of IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS), Philadelphia, PA, 2015, pp. 225–234
16. Butko, A., Bruguier, F., Gamatie, A., Sassatelli, G., Novo, D., Torres, L., Robert, M.: Full-system simulation of big.LITTLE multicore architecture for performance and energy exploration. In: Proceedings of IEEE 10th International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSOC), Lyon, France, 2016, pp. 201–208
17. Greenhalgh, P.: big.LITTLE Processing with ARM Cortex™-A15 & Cortex-A7, ARM White Paper, 2011, pp. 1–8
18. Jeff, B.: big.LITTLE technology moves towards fully heterogeneous global task scheduling, ARM White Paper, 2013, pp. 1–13
19. ARM White Paper: big.LITTLE Technology: The Future of Mobile, 2013, pp. 1–12
20. Kamdar, S., Kamdar, N.: big.LITTLE architecture: heterogeneous multicore processing. Int. J. Comput. Appl. **119**(1), 35–38 (2015)
21. Yu, K., Han, D., Youn, C., Hwang, S., Lee, J.: Power-aware task scheduling for big.LITTLE mobile processor. In: Proceedings of International SoC Design Conference (ISOCC), Busan, 2013, pp. 208–212
22. Jain, S., Navale, H., Ogras, U., Garg, S.: Energy efficient scheduling for web search on heterogeneous microservers. In: Proceedings of IEEE/ACM International Symposium on Low Power Electronics and Design (ISLPED), Rome, 2015, pp. 177–182
23. Nishtala, R., Carpenter, P., Petrucci, V., Martorell, X.: Hipster: hybrid task manager for latency-critical cloud workloads. In: Proceedings of IEEE International Symposium on High Performance Computer Architecture (HPCA), Austin, TX, 2017, pp. 409–420
24. Holmgren, R.: Energy efficiency experiments on Samsung Exynos 5 heterogeneous multicore using OmpSs task based programming. MS Thesis, Norwegian University of Science and Technology, Department of Computer and Information Science, pp. 1–55 (2015)

25. Yoo, S., Shim, Y., Lee, S., Lee, S., Kim, J.: A case for bad big.LITTLE switching: how to scale power-performance in SI-HMP. In: Proceedings of HotPower'15 ACM Proceedings of the Workshop on Power-Aware Computing and Systems, NY, USA, 2015, pp. 1–5
26. Villebonnet, V., Costa, G.D., Lefevre, L., Pierson, J., Stolf, P.: Towards generalizing "big.Little" for energy proportional HPC and cloud infrastructures. In: Proceedings of IEEE Fourth International Conference on Big Data and Cloud Computing, Sydney, NSW, 2014, pp. 703–710

Analyzing the Detectability of Harmful Postures for Patient with Hip Prosthesis Based on a Single Accelerometer in Mobile Phone

Kitti Naonueng, Opas Chutatape and Rong Phoophuangpairoj

Abstract This research studies the use of a single accelerometer inside a smartphone as a sensor to detect those postures that may be risks for patients with hip surgery to dislocate their joints. Various postures were analyzed using Euclidean distances to determine the feasibility to detect eight postures that were harmful. With the mobile phone attached to the affected upper leg, it was found that there was one harmful posture that could not be detected due to its close similarity with a normal posture. Meanwhile, the other two harmful postures, although indistinguishable based on their measured data, were still detectable with the suitably selected threshold. The distance measure analysis is useful as an indicator as to which posture will be near to missing out in the detection process. This will form a guideline for further design of a practical and more robust detecting system.

Keywords Accelerometer · Euclidean distance · Hip prosthesis dislocation · Smartphone

1 Introduction

In a total hip replacement, there are basically four components that are used to replace the hip joint of the affected patient as shown in Fig. 1(left). The components are merged into an implant (center) that fits into the hip (right). With a sufficient

K. Naonueng
College of Engineering, Rangsit University, Pathum Thani, Thailand
e-mail: kittinn78@gmail.com

O. Chutatape
Department of Electrical Engineering, College of Engineering, Rangsit University, Pathum Thani, Thailand
e-mail: opas@rsu.ac.th

R. Phoophuangpairoj (✉)
Department of Computer Engineering, College of Engineering, Rangsit University, Pathum Thani, Thailand
e-mail: rong.p@rsu.ac.th

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_9

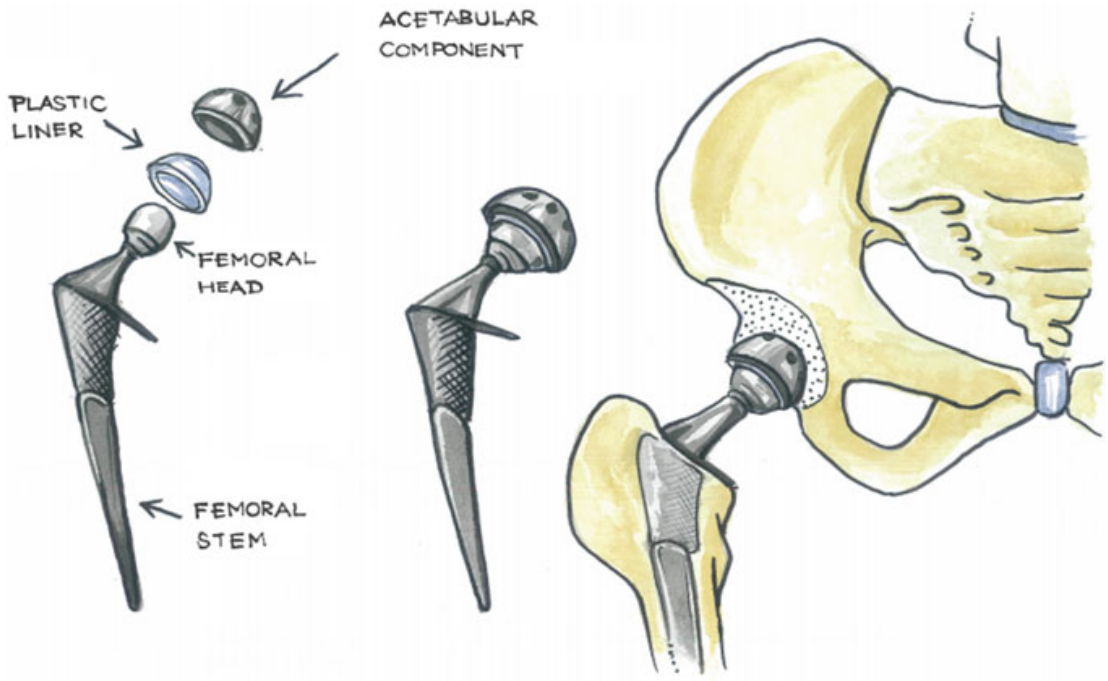


Fig. 1 Prosthetic components in total hip replacement

force from an inappropriate exercise or a repetitively unsuitable movement of the affected leg, the assembled joint can be dislocated.

For a patient who has undergone a hip replacement operation, during the recovery period which may last for a few months various postures and movements that are considered normal to ordinary people may be harmful to the person. These potentially harmful postures should be avoided to prevent possible dislocation of the joint that may lead to further complication and even require another surgery. Some of these forbidden postures or movements [1] are

- Cross the legs at the knees (during sitting, standing, or sleeping).
- Bring the knee up higher than the hip.
- Lean forward while sitting or as sitting down.
- Try to pick up something on the floor while sitting.
- Turn feet excessively inward or outward when bending down.
- Reach down to pull up blankets when lying in bed.
- Bend at the waist beyond 90°.

After being discharged from the hospital the person will need some help to get along with daily life at home. During this period a device that can help the person to stay more independently and safely will be most beneficial, and one such device, in this case, should be able to detect and give a warning against harmful movements or any postures that may cause a risk to the operated leg. Considering that the smartphone has already become an essential device for almost everybody's daily life and many applications are readily being built around its user-friendliness and myriad communication features, the next logical choice is to explore the

feasibility to incorporate these healthcare needs to it. Among various sensors embedded inside most smartphones are the most often used accelerometer in the form of a MEMS (Micro Electro Mechanical System) chip. The accelerometer is the sensor that is associated with the movement detection and therefore has applications in various core research areas such as gait analysis [2] human activity recognition [3–10], healthcare and fitness [11–17].

In this paper, the aim is to present the systematic approach to the collection of data and detection of improper postures based on a single accelerometer embedded in the smartphone, then to give the typical coordinate values of these harmful postures and of the relative distances that indicate the degree of detectability among themselves. This will simplify the remaining task of identifying the unrecommended postures and those that require further data collection to correctly identify them.

2 Related Work

Two major approaches have been used to detect and measure patterns of human motion in the three major research areas previously mentioned. One approach is based on the image processing techniques whereby video cameras were used to record 3D sequence of photographs of individual locomotion, which were then analyzed. Although the system is commercially available, highly accurate, and successfully implemented in a controlled environment, it is not publicly used and taken into everyday activities. Few main reasons are the relatively higher cost, the complexity of the procedures involved in the set-up, and the confinement to the controlled area to ensure proper operation. In another different approach, using accelerometers as movement and orientation sensors is more often the choice by most researchers. In recent research activities in the healthcare area, some new relevant initiatives have emerged. These are leg movement and hold detection for knee extension exercise [14], hip dislocation study [15], and upper arm posture and movement measurements [16]. According to the last report [16], an iPhone application was developed and validated against an optical tracking system (OTS). The small difference in comparison with the OTS indicated that the application which was based on sensors in a smartphone is a valid method. In another study, the method used to develop an Android-based mobile application which activated an alert signal and provided feedback in real time on the effectiveness of knee extension exercises was presented in [17]. Based on these reports it confirms that adopting a smartphone as a posture detecting device can give acceptable, valid and obviously more cost-effective results.

With a particular interest in the healthcare of the elderly, recent work on the detection of harmful postures leading to possible dislocation of the hip prosthesis was studied by using 3-axis accelerometer sensor present in any smartphone [15]. As far as the authors' knowledge, it was the first initiative to use a smartphone for the purpose of detecting (and further warning) the risky posture for hip-operated

patients. It was reported there that out of 14 postures understudy, there were three harmful postures that could not be classified correctly by using simply one smartphone. To give an insight into the details of these postures, this paper presents further study and report on the new analysis of the problem that could offer a better implementation.

3 Materials and Method

3.1 Instruments

The Android-based Sensor Fusion software set developed by Linköpins Universitet from Google Play Store was installed into Samsung Galaxy S6 smartphone and the application was used to record the sensor data. The smartphone was put into an extended armband and fastened to the thigh in the position as shown in Fig. 2.

For each typical posture selected as a reference, the 3-axis values of the accelerometer are recorded by the software. The deviation from this reference is measured using the Euclidean distance. To perform the detection analysis the application software was set to stream the data through the local area network to the program written in Microsoft Visual C# as shown in Fig. 3. This method is efficient and the data can be obtained in real time.

3.2 Proposed Method

For a smartphone testing, instead of using a traditional method that researchers have to develop an Android program, install and test it on a smartphone, this work used the publicly available software to read and send the sensor data and the PC-based program to study them. Changing a PC-based detection program is usually easier than changing the mobile program. It does not need to repeatedly load the written program to a smartphone installation and execution. Using this method an Android

Fig. 2 Position of the smartphone on the thigh

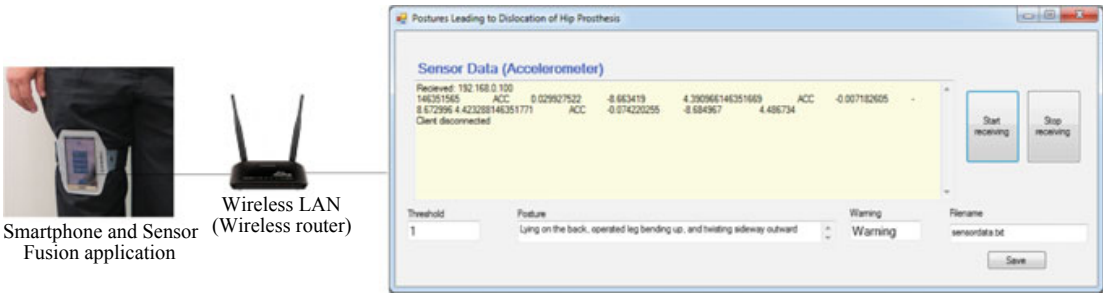


Fig. 3 Data collection of the test method

application to detect harmful postures can be developed after obtaining the desired test results.

The whole set-up worked as follows:

(1) Receive the accelerometer data from the installed application using the TCP socket communication.
(2) Extract the accelerometer *x*, *y,* and *z* values from the streaming data. Since the accelerometer values were received after the keyword "ACC", the program checked the keyword and then captured the string values of *x*, *y,* and *z*, respectively.
(3) Convert the string representation of the *x*, *y,* and *z* numbers to their equivalent real numbers.
(4) Compute the Euclidean distance between the receiving *x*, *y,* and *z* values and the selected values of each harmful posture.
(5) Detect harmful postures based on the Euclidean distance and a predefined threshold.

Steps 4–5 can be explained in details as follows:
Given that
P is the set of all detecting postures represented by

$$P = p_1, p_2, \ldots, p_i \quad 1 \leq i \leq N \tag{1}$$

N = the number of detecting postures,

p_i is selected accelerometer *x*, *y*, and *z* (reference) values for the *i*th posture represented by

$$p_i = (\mathrm{Ref}x_i, \mathrm{Ref}y_i, \mathrm{Ref}z_i), \quad 1 \leq i \leq N \tag{2}$$

O is an observation sequence (a sequence of accelerometer data) consisting of observations (o_t) at time 1 to time *T*.

$$O = (o_1\ o_2\ o_3 \ldots o_T) \tag{3}$$

$$o_t = (x_t, y_t, z_t) \quad 1 \leq t \leq T \tag{4}$$

(x_t, y_t, z_t) are accelerometer x, y, and z values occurring at time t. The Euclidean distance between two accelerometer data p_i and o_t is given by dist(p_i, o_t). The method to detect postures (DP*) at time t is described using the following equations:

$$\mathrm{DP}^* = \arg \min_{p_i \in P} \mathrm{dist}(p_i, o_t) \tag{5}$$

and

$$\mathrm{dist}(p_{\mathrm{DP}^*}, o_t) \leq \sigma \tag{6}$$

σ is a predefined detection threshold obtained from the experiments (e.g., 1).

By following the steps as indicated, manual intervention can be kept to a minimum and various results can be obtained for further analysis by using existing mathematical software packages.

4 Results

4.1 Standing, Walking, and Sitting

Figure 4 shows the accelerometer data x, y, z in the four scenarios, i.e., (1) straight standing, (2) walking, (3) sitting on a chair, and (4) sitting on a chair and slightly move the thigh up and down.

4.2 Reaching Down to Collect Object on the Floor or to Tie Shoestrings

Figure 5 shows the accelerometer data x, y, z when (1) sitting and bending up the leg to tie shoestrings, (2) sitting and stooping to collect an object on the floor or to tie shoestrings, and (3) stooping to collect object on the floor or to tie shoestrings with leg bending.

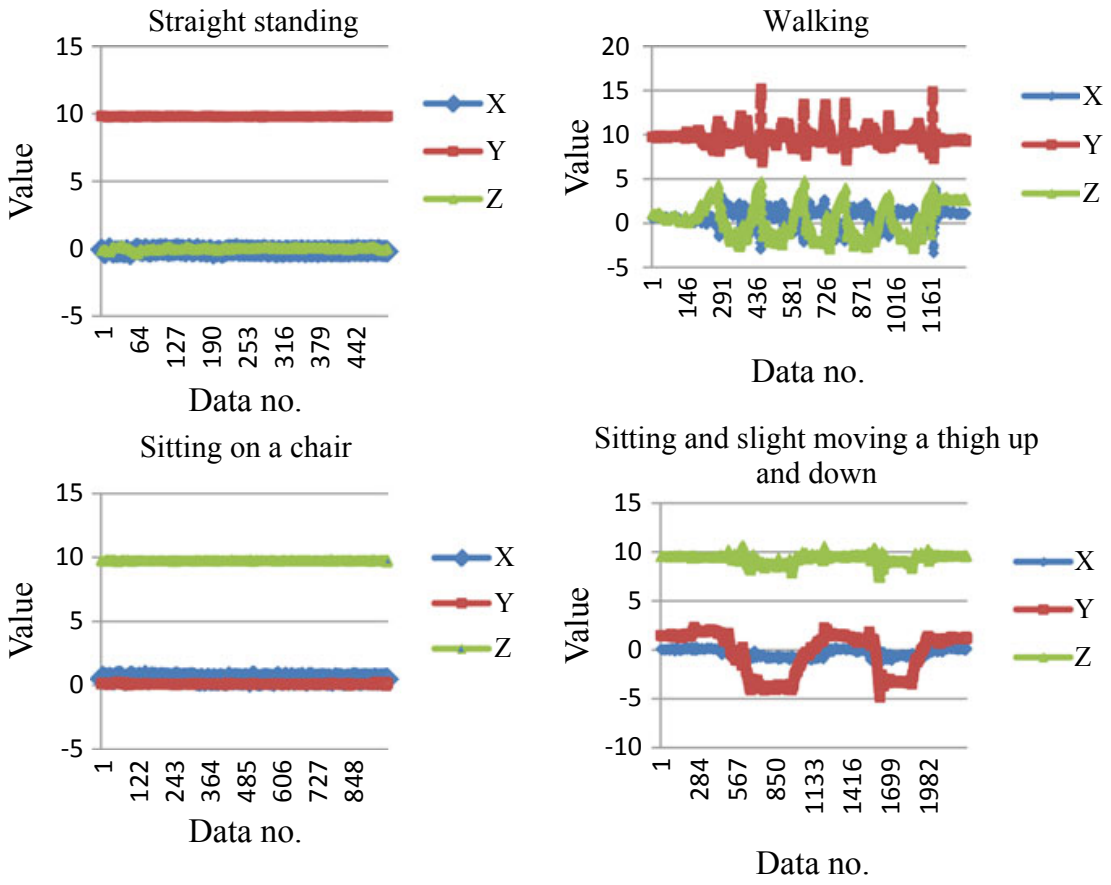


Fig. 4 Straight standing, walking, and sitting (with possibly a slight up-down thigh movement)

4.3 Lying on the Back

Figure 6 shows the accelerometer data *x*, *y*, *z* when (1) lying on the back, (2) lying on the back with the operated leg bent up and twisted sideways inward, and (3) lying on the back with the operated leg bent up and twisted sideways outward.

4.4 Twisting the Operated Leg Inward and Outward While Standing

Figure 7 shows the accelerometer data *x*, *y*, *z* when twisting an operated leg inward while standing, and outward while standing.

The next Fig. 8 shows the accelerometer data *x*, *y*, *z* when getting on a motorcycle.

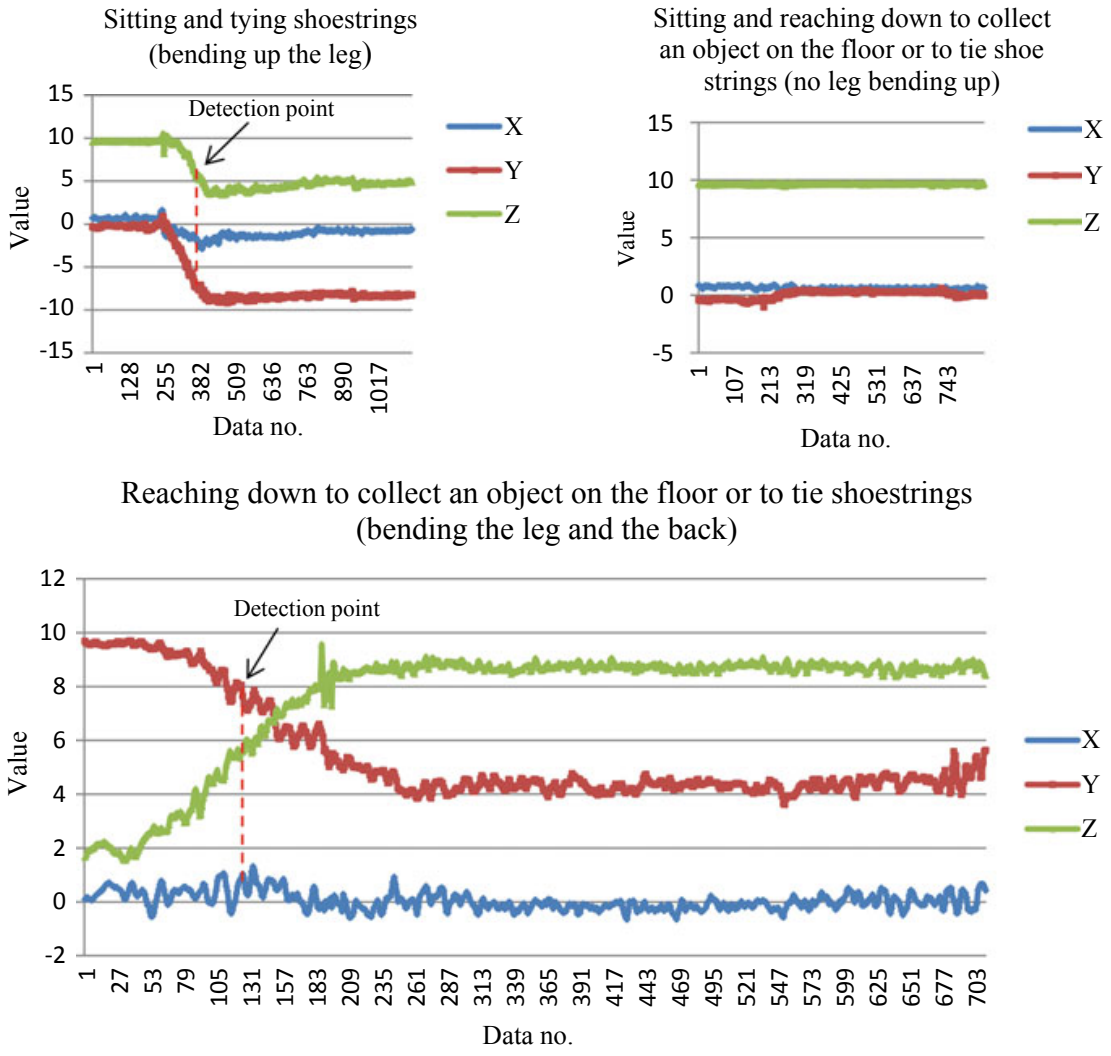


Fig. 5 Reaching down to collect object on the floor or to tie shoestrings

4.5 Accelerometer Outputs at Various Postures

The *x*–*y*–*z* values of the smartphone accelerometer at various different postures are shown in Table 1. In the last column, each posture is assigned either risk factor 0 (no risk), or 1 (with risk).

Euclidean distances between centroids are used to represent the relative differences among various postures as given in details in Table 2. Posture 2 (walking) and posture 3 (sitting) are two postures that have more different readings recorded. Notice that postures 1–4 are classified as "no risk" and postures 5–12 are classified as "at risk".

By setting the threshold of the Euclidean distance at unity, it was seen that most postures, particularly those considered being harmful, could be separated. These are cells with corresponding distance values more than one and displayed as empty cells as shown in Table 3. On the other hand, posture 5 and 8 could not be

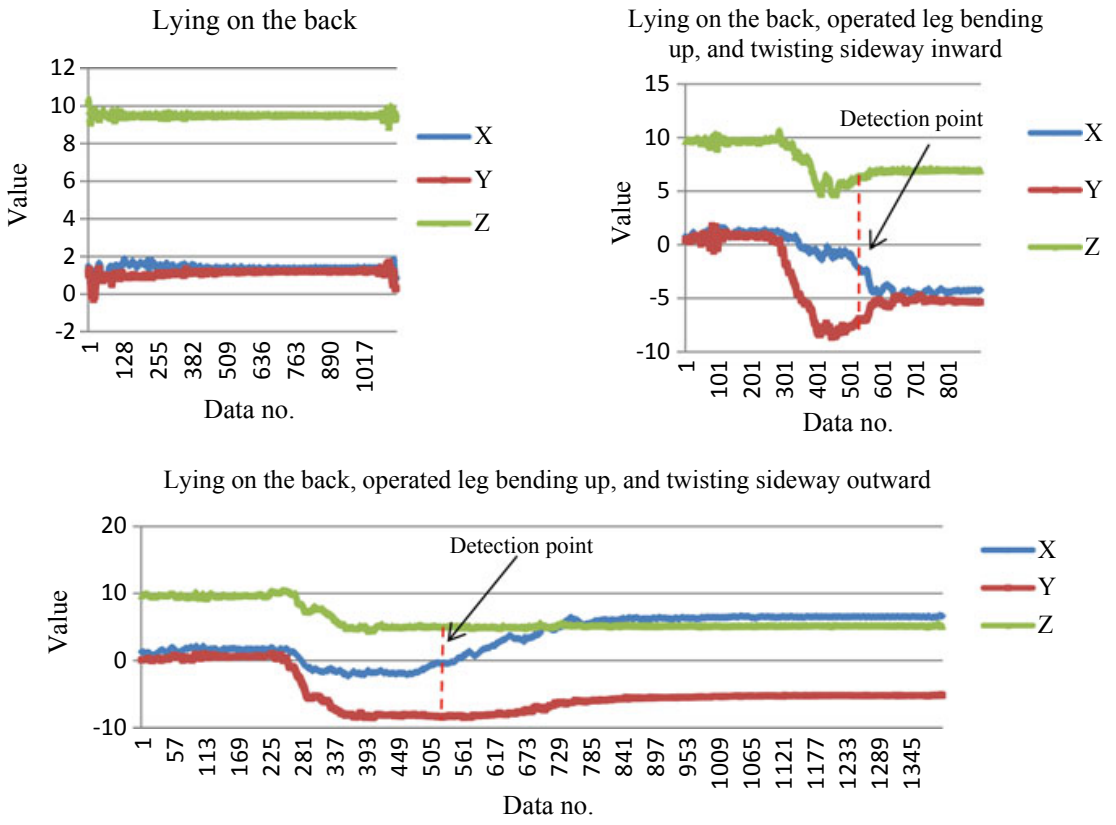


Fig. 6 Lying on the back posture

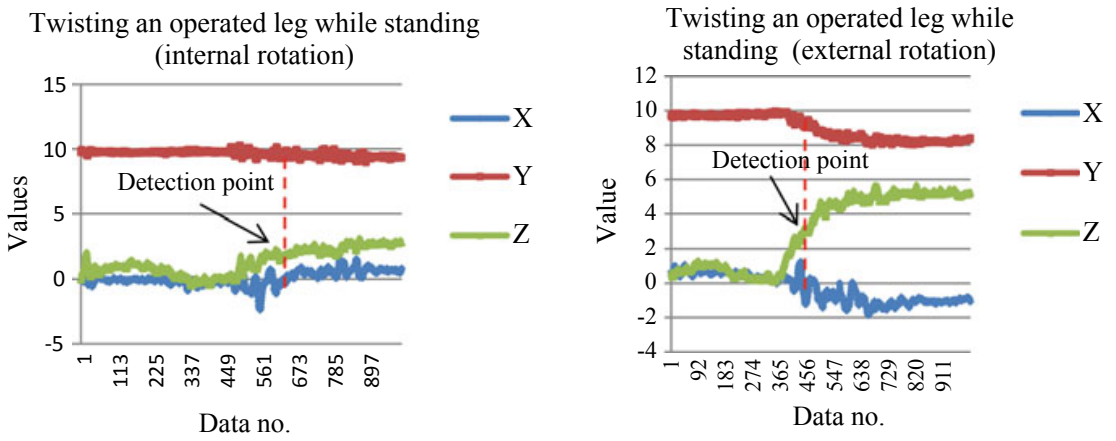


Fig. 7 Twisting an operated leg while standing (internal and external rotation)

distinguished from each other but they could be detected as harmful and give an alert signal since both were sufficiently at a far distance from the others. Posture 6 could not be recognized correctly because it was quite similar to posture 3 as far as the accelerometer values are concern. These cells in the table are labeled with the letter "N" which indicates that the two postures indicated by row and column could

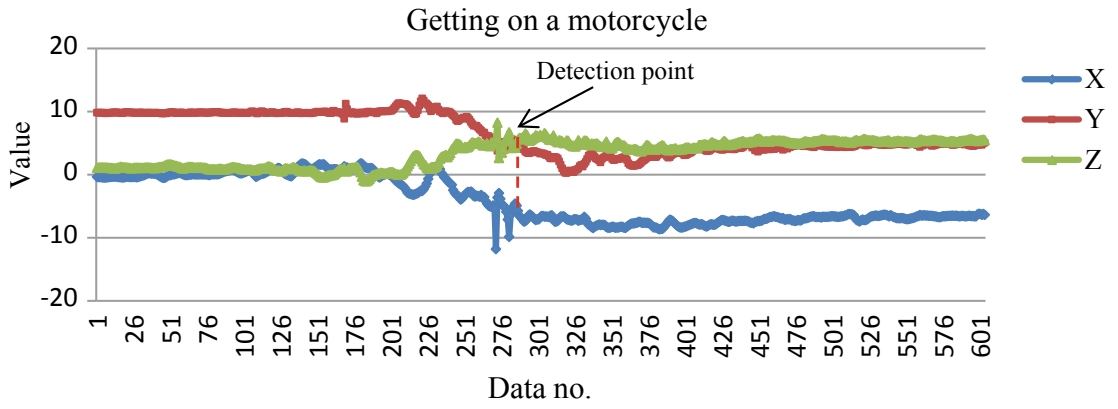


Fig. 8 Getting on a motorcycle

Table 1 *x*–*y*–*z* values of accelerometer at various postures

Posture	*x*	*y*	*z*	Risk factor
1. Straight standing	− 0.1545	9.802219	0.005719	0
2. Walking	0.501585	9.861717	0.52433	0
	0.659603	10.34175	0.802058	
	− 0.68235	10.52132	4.383783	
	1.740585	9.680955	− 1.14443	
	1.833959	9.56723	− 2.51152	
	− 2.23499	14.66808	2.592921	
3. Sitting	0.485402	0.094524	9.734577	0
	− 0.0826	1.447295	9.536105	
	− 0.90621	− 2.90297	8.707711	
	− 0.80325	− 3.95283	− 3.95283	
4. Lying on the back	1.377871	1.099392	9.50181	0
5. Sitting and tying shoestrings (bending up the operated leg)	− 1.99916	− 7.25922	5.436035	1
6. Sitting and reaching down to collect an object on the floor or to tie shoestrings (no leg bending up)	0.730254	0.020997	9.625291	1
7. Reaching down to collect an object on the floor or to tie shoestrings (bending the leg, and the back)	0.31364	8.137892	5.482722	1
8. Lying on the back, operated leg bending up, and twisting sideways inward	− 1.83994	− 7.34302	6.050148	1
9. Lying on the back, operated leg bending up, and twisting sideways outward	− 0.35674	− 8.43836	4.987122	1
10. Twisting an operated leg while standing (internal rotation)	− 0.20351	9.51336	2.000356	1
11. Twisting an operated leg while standing (external rotation)	− 0.50996	9.471462	2.958036	1
12. Getting on a motorcycle	− 5.89213	4.424485	5.469554	1

Table 2 Euclidean distance between postures

Posture	Euclidean distance																			
	1	2						3				4	5	6	7	8	9	10	11	12
1	0.00	0.84	1.26	4.47	2.22	3.22	5.89	13.76	12.67	15.42	14.33	12.97	18.00	13.75	5.74	18.26	18.91	2.02	2.99	9.58
2	0.84	0.00	0.58	4.09	2.09	3.33	5.91	13.43	12.34	15.23	14.58	12.58	17.99	13.41	5.25	18.22	18.86	1.67	2.66	9.74
	1.26	0.58	0.00	3.83	2.32	3.60	5.51	13.60	12.49	15.50	15.14	12.71	18.39	13.58	5.19	18.62	19.27	1.69	2.60	9.99
	4.47	4.09	3.83	0.00	6.09	7.40	4.78	11.78	10.45	14.11	16.70	10.92	17.86	11.82	2.81	17.98	18.97	2.63	1.78	8.09
	2.22	2.09	2.32	6.09	0.00	1.38	7.39	14.55	13.61	16.20	14.15	13.68	18.55	14.50	6.95	18.83	19.24	3.70	4.68	11.39
	3.22	3.33	3.60	7.40	1.38	0.00	8.28	15.54	14.65	17.00	13.85	14.71	19.00	15.48	8.26	19.31	19.63	4.95	5.95	12.24
	5.89	5.91	5.51	4.78	7.39	8.28	0.00	16.46	15.09	18.65	19.79	15.65	22.11	16.52	7.58	22.28	23.31	5.57	5.49	11.25
3	13.76	13.43	13.60	11.78	14.55	15.54	16.46	0.00	1.48	3.46	14.33	1.36	8.87	*0.28*	9.10	8.62	9.80	12.21	11.61	8.81
	12.67	12.34	12.49	10.45	13.61	14.65	15.09	1.48	0.00	4.50	14.55	1.50	9.81	1.64	7.83	9.62	10.89	11.04	10.39	7.69
	15.42	15.23	15.50	14.11	16.20	17.00	18.65	3.46	4.50	0.00	12.70	4.68	5.56	3.47	11.57	5.26	6.69	14.13	13.65	9.44
	14.33	14.58	15.14	16.70	14.15	13.85	19.79	14.33	14.55	12.70	0.00	14.54	10.03	14.23	15.38	10.61	10.01	14.74	15.10	13.60
4	12.97	12.58	12.71	10.92	13.68	14.71	15.65	1.36	1.50	4.68	14.54	0.00	9.89	1.26	8.18	9.67	10.69	11.38	10.79	8.95
5	18.00	17.99	18.39	17.86	18.55	19.00	22.11	8.87	9.81	5.56	10.03	9.89	0.00	8.83	15.57	*0.64*	2.07	17.22	16.98	12.32
6	13.75	13.41	13.58	11.82	14.50	15.48	16.52	*0.28*	1.64	3.47	14.23	1.26	8.83	0.00	9.12	8.58	9.71	12.21	11.63	8.97
7	5.74	5.25	5.19	2.81	6.95	8.26	7.58	9.10	7.83	11.57	15.38	8.18	15.57	9.12	0.00	15.64	16.60	3.78	2.97	7.23
8	18.26	18.22	18.62	17.98	18.83	19.31	22.28	8.62	9.62	5.26	10.61	9.67	*0.64*	8.58	15.64	0.00	2.13	17.41	17.15	12.46
9	18.91	18.86	19.27	18.97	19.24	19.63	23.31	9.80	10.89	6.69	10.01	10.69	2.07	9.71	16.60	2.13	0.00	18.20	18.03	14.01
10	2.02	1.67	1.69	2.63	3.70	4.95	5.57	12.21	11.04	14.13	14.74	11.38	17.22	12.21	3.78	17.41	18.20	0.00	1.01	8.38
11	2.99	2.66	2.60	1.78	4.68	5.95	5.49	11.61	10.39	13.65	15.10	10.79	16.98	11.63	2.97	17.15	18.03	1.01	0.00	7.79
12	9.58	9.74	9.99	8.09	11.39	12.24	11.25	8.81	7.69	9.44	13.60	8.95	12.32	8.97	7.23	12.46	14.01	8.38	7.79	0.00

Italic represents distinctly the low Euclidean distance between posture 3 and posture 6 (*0.28*) and also posture 5 and posture 8 (*0.64*), which made the system difficult to distinguish the postures accurately

Table 3 Separability at unity threshold

Posture	1	2	3	4	5	6	7	8	9	10	11	12
1. Straight standing												
2. Walking												
3. Sitting						N						
4. Lying on the back												
5. Sitting and tying shoestrings (bending up the operated leg)								N				
6. Sitting and reaching down to collect an object on the floor or to tie shoestrings (no leg bending up)			N									
7. Reaching down to collect an object on the floor or to tie shoestrings (bending the leg and the back)												
8. Lying on the back, operated leg bending up, and twisting sideways inward					N							
9. Lying on the back, operated leg bending up, and twisting sideways outward												
10. Twisting an operated leg while standing (internal rotation)												
11. Twisting an operated leg while standing (external rotation)												
12. Getting on a motorcycle												

not be distinguished from each other at the given threshold. For a higher threshold, more of such postures that are indistinguishable will surface up.

The Euclidean distance-based algorithm to identify harmful posture may be initiated by periodically reading the 3-axis accelerometer outputs from the mobile phone, then calculate the Euclidean distance between the sample posture and the known, selected values of the harmful postures in the database. If the distance is less than a chosen threshold value, then the sample posture is likely to close to the harmful one and the warning signal should be given.

There however remain certain postures that cannot be identified as normal or harmful, and these require further data from additional sensors to be attached at the other major parts of the body. One possibility is to attach a second sensor on the chest. This will give the information about the movement of the upper part of the body which can differentiate posture 5 from 8, and posture 6 from 3. Nevertheless, considering that each individual may not move exactly in the same pattern with the same range of movement for the same posture, and the position of the attached sensor can vary, the threshold value is an individual-dependent parameter that will require custom adjustment.

5 Conclusion

The analysis on the detectability of various harmful postures for hip-operated patient based on the use of a single accelerometer in the smartphone has been presented. This is the first initiative and a contribution made to the research area in addition to the proposed method for other similar developments. It was found that at the selected threshold, six out of eight harmful postures, including the major ones, could be detected and so the warning signal could be generated by the smartphone. In the future, the method to distinguish posture 3 (sitting) from posture 6 (sitting and reaching down to collect an object on the floor or to tie shoestrings) ought to be developed. This ambiguous posture could be solved by implementing an additional sensor to detect the movement of the upper part of the body. However, adding more sensors on the body is not the favorable option from a user convenience point of view, and therefore using a smartphone alone to detect all or as many of these harmful postures probably through better utilization of sensors and intelligent time series data analysis should be further investigated.

Acknowledgements We would like to thank Dr. Nathee Ruangthong of Ban Mi Hospital, Lop Buri, Thailand, for many of his valuable suggestions. One was that no real patient was required to participate as it might cause injuries and some difficulties. With this recommendation and to follow the ethical guideline for clinical trial, the participant who was a normal person in this trial was fully explained of the procedure and the use of the data collected and was later given a proper form of consent to sign voluntarily.

References

1. Activities After Hip Replacement—OrthoInfo—AAOS: American Academy of Orthopaedic Surgeons. https://orthoinfo.aaos.org/en/recovery/activities-after-hip-replacement/. Accessed 07 Sep 2018
2. Muro-de-la-Herran, A., Garcia-Zapirain, B., Mendez-Zorrilla, A.: Gait analysis methods: an overview of wearable and non-wearable systems. Highlighting Clin. Appl. Sens. **14**(2), 3362–3394 (2014)

3. Noury, N., Fleury, A., Rumeau, P., Bourke, A. K., Rialle, V., Lundy, J.E.: Fall detection: principles and methods. In: Proceedings of 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE EMBS), pp. 1663–1666 (2007)
4. Brezmes, T., Gorricho, J.-L., Cotrina, J.: Activity Recognition from Accelerometer Data on a Mobile Phone. Lecture Notes in Computer Science, vol. 5518, pp. 796–799. Springer, Berlin (2009)
5. Krishnan, N., Panchanathan, S.: Analysis of low resolution accelerometer data for continuous human activity recognition. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 3337–3340 (2008)
6. Long, X., Yin, B., Aarts, R.M.: Single accelerometer-based daily physical activity classification. In: Proceedings 31st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), pp. 6107–6110 (2009)
7. Mannini, A., Sabatini, A.M.: Machine learning methods for classifying human physical activity from on-body accelerometers. Sensors **10**(2), 1154–1175 (2010)
8. Huda, F.A., Tolle, H., Putra, K.P.: Human activity recognition using single accelerometer on smartphone put on user's head with head-mounted display. Int. J. Adv. Soft Comput. Appl. **9** (3), 239–249 (2017)
9. Mathie, M., Celler, B., Lovell, N., Coster, A.: Classification of basic daily movements using a triaxial accelerometer. Med. Biol. Eng. Comput. **42**, 679–687 (2004)
10. Kwapisz, J.R., Weiss, G.M., Moore, S.A.: Activity recognition using cell phone accelerometers. ACM SIGKDD Explor. Newsl. **12**(2), 74–82 (2010)
11. Putra, I.P.R.S., Brusey, J., Gaura, E., Vesilo, R.: An event-triggered machine learning approach for accelerometer-based fall detection. Sensors **18**(20), 1–18 (2017)
12. Gjoreski, H., Luštrek, M., Gams, M.: Accelerometer placement for posture recognition and fall detection. In: Proceedings of International Conference on Intelligent Environments, Nottingham, UK, pp. 47–54 (2011). https://doi.org/10.1109/ie.2011.11
13. Khawandi, S., Daya, B., Chauvet, P.: Automated monitoring system for fall detection in the elderly. Int. J. Image Process. **4**(5), 476–483 (2010)
14. Phoophuangpairoj, R.: Frame-based analysis of knee extension exercises using a smartphone accelerometer. In: Proceedings of 13th International Joint Conference on Computer Science and Software Engineering (JCSSE) (2016)
15. Chutatape, O., Naonueng, K., Phoophuangpairoj, R.: Detection of dangerous postures leading to dislocation of hip prosthesis by a smartphone. In: Proceedings of 14th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON) (2017)
16. Yang, L., Grooten, W.J.A., Forsman, M.: An iPhone application for upper arm posture and movement measurements. Appl. Ergon. **65**, 492–500 (2017). https://doi.org/10.1016/j.apergo.2017.02.012
17. Phoophuangpairoj, R.: Developing an accelerometer-based mobile application to aid knee extension exercise. Int. J. Adv. Soft Comput. Appl. **10**(2), 132–147 (2018)

Software Development Process Evolution in Malaysian Companies

Rehan Akbar, Asif Riaz Khan and Kiran Adnan

Abstract GSD is a phenomenon mainly associated with the outsourcing of software development projects to some offshore company. Reduction in software development cost increased productivity and advantage of multisite development with respect to time are the main benefits that software development companies (SDCs) get from GSD. Besides benefits, a number of challenges associated with GSD are also observed. Consequently, the traditional processes to develop software are evolving and being replaced with a new set of processes which are lightweight and outcome-based. The process evolution has been deeply investigated in the context of companies mostly in Europe, Australia, USA and mainly other countries in those regions. In this regard, limited research has been carried out on Malaysian companies. The present research investigates the process evolution phenomenon in Malaysian companies. The current software development processes and the reasons for the evolution of software processes in Malaysian software companies have been identified. A qualitative approach using structured interviews has been followed for the collection of data collection and its analysis. The findings explain that software processes in most of the Malaysia companies are increasingly evolving or have been evolved. The companies are overwhelmingly adopting agile methods because of their support to GSD. Some of the companies are using ad hoc approaches for software development. The size of the company and project has been found as one of the main factors behind using ad hoc approaches. Mainly the small and medium-size companies and projects are involved in this practice.

Keywords Agile · Global software development · Malaysian companies · Software evolution · Software processes

R. Akbar (✉) · A. R. Khan · K. Adnan (✉)
Faculty of Information and Communication Technology,
Universiti Tunku Abdul Rahman, 31900 Kampar, Perak, Malaysia
e-mail: rehan@utar.edu.my

K. Adnan
e-mail: kiranadnan@1utar.my

A. R. Khan
e-mail: asifriazkhan@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_10

1 Introduction

The Global Software Development (GSD) started in the late 1990's by mid-2000's it got an overwhelming response from the software companies [1]. GSD is a form of project outsourcing to offshore companies. GSD is carried out from different locations around the world that involved different teams for the development of software product [2, 3]. The availability of fast internet and latest communication tools are increasing the trend of developing software product from different locations [4]. GSD also developed the connections among different nations having different behaviors, social values, and different culture [5]. GSD helps the software companies to decrease the software development cost. Since, the GSD provides a number of benefits such as less development cost, development from different locations with respect to time, access to a large pool of experienced people and tax incentive [6–9]. Due to the GSD benefits, most of the SDCs have started the GSD to decrease the cost of software development and improve the software quality [10]. Moreover, GSD help to exchange the best expertise among teams [1, 8] and also provide access to skilled resources that would help the SDCs to increase their productivity and efficiency [1, 5, 8, 9]. In addition, the development of software with follow-the-sun (24/7) model also is an added advantage of GSD [8, 11].

Despite having advantages, a number of challenges are also associated with GSD [2, 5, 7, 12–16]. The main challenges being faced in GSD are related to communication and collaboration [5, 12, 13, 17–21], poor management [13, 22], project diversity and complexity,[12] and delayed response due to time zone difference [7, 12]. Although, the SDCs are using different communication applications like Skype, Instant Message (IM), Google chat, telephones, and emails [23, 24] to reduce these challenges. However, still, problems exist at certain levels and need to be addressed properly. To overcome the GSD challenges, many studies have proposed some strategies to meet the GSD challenges [6, 13, 25–28].

To deal with these challenges and to meet the market competition, software development processes have been evolving since the late 1990s and shifting towards new software development paradigms prominently agile methods [19]. The software project clients want to see the progress of the project on a weekly basis and also require to make necessary changes in requirements and deliverables anytime during the project. Consequently, to meet the client expectations the SDCs started following agile methods, as these approaches define procedures to provide an update to the client on a weekly basis [19, 29, 30]. According to Versionone survey [31] about 80% of the SDCs have been moved to the agile methodologies. The evolution in software development processes in companies operating in the UK, USA Australia, and other western countries have been investigated in different studies. However, limited work has been presented in the context of Malaysian companies. Most of the work presented on software development processes in Malaysian companies are very general and do not provide deep insight into process evolution and process paradigm shift. In the present research work, the aim is to check the trend of current software development processes in Malaysian companies

as a result of GSD. The evolution of software development processes, how the processes got affected and eventually changed in Malaysian companies, and reasons behind. The study also presents the expert views on the change in software development processes. In addition, some discussion on the software development practices in Malaysian companies has been made in the literature review section.

The section two of the paper presents the review of related literature while research methodology is presented in the third section. The results have been discussed in the fourth section and finally, conclusion has been summarized in section five followed by acknowledgment.

2 Related Work

A software process involves a set of procedures, tasks, and activities which are required to build a software [32]. The phenomenon to develop software from different locations is called GSD [19, 33]. Software development at different locations made the conventional software development process more difficult to implement. Consequently, most of the software development projects have been failed and over budget [34, 60, 61]. A number of factors behind the failure of GSD projects are misunderstanding of requirements, absence of face-to-face communication, absence of suitable software process, high coordination costs, lack of experience about distributed development, culture and languages problems, dependent modules in distributed development, and delay in responses due to time zone difference [4, 5, 7, 10, 12, 13, 32, 35, 36]. Numerous studies have suggested solutions and improvement to deal with these challenging factors [2, 6, 13, 25–28, 37–39].

The main purpose of using the software development process is delivering the project on time and within budget [32]. Therefore, software development companies need to follow a suitable software development process that helps to minimize these issues and deliver projects on time. The conventional approaches of software development have been deprecated because of less support to GSD and lightweight methodologies such as agile have appeared as popular practices by software development companies because of their support to frequent communication, short deliverables, and other disciplined practices [13, 18, 19, 26]. It is also observed that the application of agile methods in GSD environments has significantly reduced the failure rate of the projects [18, 26]. The existing studies describe that current software practices have been changed globally but are uncertain about Malaysian SDCs. As per our knowledge, there has not been carried out such study that investigates the current trends of software development in Malaysia as consequences of GSD. The research study presented in this paper accomplishes this research gap. Though a few studies have been conducted on Malaysian companies [40–46] but not in GSD perspective. Also, these studies insufficiently present the reasons behind the selection of the software development process as a result of GSD.

Almomani in a survey [41] investigated the difficulties in recent software processes improvement (SPI) practices in small and medium-size companies (SMEs) operating in Malaysia. The study has some limitations. Firstly, the study has been conducted through a questionnaire with small sample size, secondly, the target respondents are only those companies which are doing in-house development and are not involved in GSD. The study concludes that implementation of SPI is not very common in Malaysian SMEs because of inadequate knowledge, communication issues, and lack of resources. In addition, findings also show that many software development companies in Malaysia are using ad hoc approaches and agile methodologies. In 2005, Baharom [42] conducted a survey on software development practices in Malaysia. The survey has been conducted through a questionnaire and the findings of the survey explain that most of the software development companies are not following any software development process in their development environments. Moreover, software development companies have been facing problems in adopting standards and processes, software quality, over budget, and late delivery related issues.

In another research study [40], a survey using a questionnaire to observe the current practices of developing software in Malaysian companies has been conducted. The statistical analysis of the responses has been made to interpret the results. Findings show that the percentage usage of the methodology for managing the software project is still very low. Likewise, the use of the latest software development practices to develop software has also not been found satisfactory due to which the success rate of the projects is low. Another study conducted on "the adoption of software development methodologies among IT organization in Malaysia" has been found different from the present study in terms of (a) the study has been conducted on overall IT industry, not particularly on the companies involved in GSD, (b) the software development methodologies are defined and presented according to each phase of software development lifecycle (SDLC) such as requirements phase, design phase, development, and testing phase and (c) the focus of the study is too general [44]. Moreover, the results derived from these findings do not particularly list the names of software development processes rather it just generally discussed the software practices (without explicitly listing the software practices) being followed by Malaysian companies. However, it is found that few of the companies have developed their own processes while some of them are not following any software process. Furthermore, the author also proposed a framework with respect to each software development phase but sufficient discussion has not been made rather just discussed generally. The present study covers this gap and adequately investigates the most recent trends of software development practices in Malaysian companies as consequences of GSD and presents reason behind the process evolution.

3 Research Methodology

The research subject in present studies is software development companies operating in Malaysia which are practicing GSD in their software development environment. The qualitative research approach has been selected and data has been collected through structured interviews conducted face-to-face. Structured interviews are conducted to enhance the comparability of data [47]. Open-ended and pre-planned questions have been set for the interviews with the same replication in all the interviews [48]. The medium to conduct a structured interview could be face-to-face or online using the internet or via telephone. The interviews are divided into two sections such as (a) demographic data which is analyzed through statistical analysis and (b) open-ended structured questions which are analyzed qualitatively by using "general inductive approach" [49]. The general inductive approach is quite similar to grounded theory with the only difference is that it does not apply any coding technique for data analysis rather it analyzes the data more straight forward. [49] has discussed the general inductive approach in detail. The data from interviews have been managed and analyzed using qualitative data analysis software NVivo. Although, NVivo provides a platform to organize the qualitative data such as interviews and surveys at one place analysis of the data is totally dependent on the researcher that how he or she derived the result from it [50]. NVivo helps to reduce manual tasks and provide easy to manage, sort, and organize the qualitative data. The main features which NVivo provide are coding, querying text data, and representation of text with multiple charts and graphs in order to understand any particular phenomena. It is difficult to explain all the features of NVivo software, however, numbers of useful links are provided to understand the working of NVivo software in detail [51–57]. In NVivo, coding is performed in the 'Nodes' container. The information related to the particular question, theme or concept is stored under one node. In the present study, the answers of interviewees on current software development practices are stored in one node. It makes it easy to summarize the data and perform comparisons to derive the findings.

To conduct interviews for data collection, initially, 40 software development companies in Malaysia were selected. Only 25 software development companies agreed for the interviews out of which later 10 of companies refused to participate for interviews because of project deadlines and emergency meeting calls during the scheduled time. The privacy policies of the companies have been found as one of the big problems of low response rate as most of the companies also refused due to their company's privacy policy. Therefore, a total of 15 interviews could be conducted. During face-to-face interviews, it is ensured that all the questions have been clearly understood by the respondents in order to get proper answers. In face-to-face interviews, the researcher has the opportunity to clarify the confusion about the questions, if arise any.

Very few questions have been set for the interview. Mainly the questions asked were about the current software development practices in their company along with the follow-up question about reasons. The interviews have been conducted with

expert professionals such as project manager (PM), technical manager (TM), team leads, senior software engineers, senior system analysts, and software developers working in the Malaysian software companies.

4 Results and Discussion

The interviews have been divided into two sections as the first section is about demographic data comprising of name, email, company location, and respondent experience, and the second section is comprised of open-ended questions related to the current software development practices along with follow-up question on reasons to select the current practices.

In the first section, the questions have been asked to know the participants' qualification, experience level, and company profile. As shown in Table 1, 40% of the interviewees are senior software engineers, 33% software engineers, and 20% are project manager/technical manager. Only 7% were system analysts.

Most of the interviewees are quite experienced professionals in the software development industry as shown in Table 2. 20% respondents have experienced between 10 and 15 years while similarly, 20% of the respondents have more than 15 years of experience. 13.3% interviewees have experience of 3–5 years and 5–10 years. However, 33.4% of them have 1–3 year(s) of experience in software development.

As shown in Table 3, the response rate of companies with respect to their location. Kuala Lumpur, Penang, and Selangor are the main states in Malaysia where the big software development companies are operating their business [46, 58]. The majority of the respondents of the present study also belong to companies in Kuala Lumpur, Penang, and Selangor as given in Table 3. The data shows that 40% of the interviewee companies are from Penang, 26.7% from Selangor, and 13.3% companies are from Kuala Lumpur. 20% of the total companies are from Ipoh.

The size of the company has been estimated based on the number of employees (staff strength) in that company. As shown in Table 4, most of the respondents, i.e., overall 90% are from small and medium-sized companies. According to [41, 42, 59],

Table 1 Employee designation

Employee designation	Interviewees	Percentage (%)
Project Manager/Technical Manager	3	20
Project Team Lead	0	0
Senior Software Engineer	6	40
Software Engineer	5	33.3
System Analyst	1	6.7
Total	15	100

Table 2 Experience level in software development

Experience level	Interviewees	Percentage (%)
1–3 year(s)	5	33.4
3–5 years	2	13.3
5–10 years	2	13.3
10–15 years	3	20
More than 15 years	3	20
Total	15	100

Table 3 Locations of respondent companies

Company location	Interviewees	Percentage (%)
Kuala Lumpur	2	13.3
Selangor	4	26.7
Penang	6	40
Ipoh	3	20
Total	15	100

the majority of the software companies in Malaysia are small and medium-sized. Table 4 shows that 40% of the software companies have staff less than 20, while 26.7% companies have staff between 20 and 50. 20% of the companies have 50–100 employees while only 13.3% companies are enterprise-level big companies and have more than 200 employees.

In the second section, questions about the effect of GSD, changes in software development processes and reasons behind it have been asked. The analysis of the qualitative data shows that GSD has brought a significant change in Malaysian SDCs and it has also affected the software development paradigms. It is evident from the results that the majority of the SDCs operating in Malaysia have replaced their existing processes with lightweight processes predominantly agile methods. 6 out of 15 respondents particularly stated that their companies are using agile methods in their software development environment to deal with the challenges and problems faced due to GSD. According to a project manager, software development companies strive to launch their products early in the market and agile method provides support in terms of "better visibility and faster development." Another project manager added that the use of the agile method as "Managers can see the

Table 4 Company size w.r.t number of staff

Number of staff	Interviewees	Percentage (%)
Less than 20	6	40
Between 20 and 50	4	26.7
Between 50 and 100	3	20
Between 100 and 150	0	0
More than 200	2	13.3
Total	15	100

progress of projects and developers do the tasks with the given priority." Software engineers also believe that agile methods are the best software development practices to handle the GSD related issues such that "as far as agile is concerned, it helps this company. There are many potential benefits and issues that can arise from GSD. The most frequently cited issue is a communication problem, so agile helps to sort out the communication problem." A number of studies report that communication is the main issue in GSD [2, 4, 5, 12, 17, 20, 21]. The respondents of the current study also consider that communication is the main problem in GSD that can be overcome by following agile methods. Agile methods provide disciplined and best practices to develop the software product [19, 21, 26]. Among agile methods, Scrum is mostly being followed in the Malaysian SDCs because Scrum supports and defines processes for weekly updates and frequent communication among team members and clients [13, 18, 19, 26]. One of the software engineer respondents mentioned that "we use Scrum because we have daily meeting to discuss what we currently doing and possible problems we traced." Furthermore, a senior software engineer also supported using the Scrum as "we've been using Scrum in-house, and then we tried to use it in some outsourcing projects and it really works in terms of managing project resources and tracking project status." The analysis of the answers of the respondents shows that agile methods are suitable for the environments of software development companies and provide best practices to build the software as well as to manage the resources, tasks, and for better visibility of the project progress.

Subsequently, 3 out of 15 respondent companies have been found as following ad hoc approaches of software development. The small and medium-size of the companies has been found as the main reason for using ad hoc approaches in Malaysia [42]. Adding to this, one of the senior software engineers highlighted the reasons of using the ad hoc approaches in their software companies as "we don't use any process because we are a small and medium-size software company. That's why we don't use any process." Likewise, a project manager/technical manager also mentioned that "Because projects are not big usually 2 to 3 weeks long and this is the reason we don't follow any process because the use of the process is unnecessary."

Similarly, other studies conducted in Malaysia [40, 42–44] reports that most companies don't follow any software process; similar trends also found in the present study. Surprisingly, Malaysian SDCs are also following the software product line (SPL) due to GSD as 4 interview respondent companies are following the SPL. The common reason reported by the respondents is "to reuse the software components and for better management." According to [21], SPL and agile methods mostly share the common goals but achieved in different ways. However, few companies are still using traditional approaches of software development because the projects are at enterprise-level, as two of the participants are using the waterfall model and only one participant reported the ISO standards.

5 Conclusion

The research study presented in this paper successfully describes the trends of software development processes in Malaysian software development industry as well as the reasons from the experts. The findings show that GSD has brought a great evolution in Malaysian SDC and as a result software paradigm has been shifted. The findings also reveal that lightweight methodologies are the recent trends in the Malaysian SDCs. Agile and SPL methods are mostly being used as these provide fast development, better visibility of project progress, easy to manage resources and tasks as well as help to minimize the issues of GSD particularly related to communication and coordination. The small and medium-size of the companies as well as projects also urges them to follow ad hoc approaches. However, few of the respondent companies are still using traditional practices in their development environment as a result of GSD.

The results of the present study adequately explain the software development processes currently being used in Malaysian companies as a result of GSD. It also presents the views of IT experts in detail in order to know the reasons to adopt the current process and evolution in software development processes.

Acknowledgements This research project is supported under UTARRF; IPSR/RMC/UTARRF/2014-C1/R01, University Tunku Abdul Rahman, Malaysia. All procedures performed in the study involving human participants were in accordance with the ethical standards of the institution, and comparable ethical standards. Prior public disclosure approval to publish the results of the study have been taken from the university. The research study is led and supervised by Dr. Rehan Akbar as the sole member of the supervisory committee. The anonymity of the respondents of the questionnaire and data has been treated as strictly confidential, and not disclosed at any level.

References

1. Yu L., Mishra A.: Risk analysis of global software development and proposed solutions. Automatika. **51**(1), 89–98 (2010)
2. Niazi, M., et al.: Challenges of project management in global software development: a client–vendor analysis. Inf. Softw. Technol. **80**, 1–19 (2016)
3. Zafar, A., Ali, S., Shahzad, R.K.: Investigating integration challenges and solutions in global software development. In: Frontiers of Information Technology (FIT) (2011)
4. Alnuem, M.A., Ahmad, A., Khan, H.: Requirements understanding: a challenge in global software development, industrial surveys in Kingdom of Saudi Arabia, pp. 297–306 (2012)
5. Cho J.: Globalization and global software development. Issues Inf. Syst. **8**(2), 287–290 (2007)
6. Gomes, V., Marczak, S.: Problems? We all know we have them. Do we have solutions too? A literature review on problems and their solutions in global software development, pp. 154–158 (2012)
7. Javed, B., Minhas, S.S.: Process support for requirements engineering activities in global software development: a literature based evaluation (2010)
8. Conchúir, E.Ó., et al.: Exploring the assumed benefits of global software development (2006)
9. Nguyen, T., Wolf, T., Damian, D.: Global software development and delay: does distance still matter? pp. 45–54 (2008)

10. Khan, A.R., Akbar, R., Ten, D.W.H.: A study on Global Software Development (GSD) and software development processes in Malaysian software companies. J. Telecommun. Electron. Comput. Eng. **8**(2), 147–151 (2016)
11. Kroll, J., et al.: Handoffs management in follow-the-sun software projects: a case study. In: 2014 47th Hawaii International Conference on System Sciences (2014)
12. Holmstrom, H., et al.: Global software development challenges: a case study on temporal, geographical and socio-cultural distance (2006)
13. Hossain, E., Babar, M.A., Paik, H.-y.: Using scrum in global software development: a systematic literature review, pp. 175–184 (2009)
14. Gotel, O., et al.: Working across borders: overcoming culturally-based technology challenges in student global software development, pp. 33–40 (2008)
15. Mohagheghi, P.: Global software development: issues, solutions, challenges [cited 2016 10 Augest]. Available from: http://www.idi.ntnu.no/grupper/su/publ/parastoo/gsd-presentation-slides.pdf (2016)
16. Rao, N.M.: Challenges in execution of outsourcing contracts (2009)
17. Shah, Y.H., Raza, M., UlHaq, S.: Communication issues in GSD, p. 8 (2012)
18. Paasivaara, M., Lassenius, C.: Could global software development benefit from agile methods? p. 5 (2006)
19. Sriram, R., Mathew, S.K.: Global software development using agile methodologies: a review of literature (2012)
20. Kamaruddin, N.K., Arshad, N.H., Mohamed, A.: Chaos issues on communication in agile global software development (2012)
21. Ali Babar, M., Ihme, T., Pikkarainen, M.: An industrial case of exploiting product line architectures in agile software development (2009)
22. Mary, C.L., Joseph, W.R.: Effects of offshore outsourcing of information technology work on client project management. Strat. Outsourcing: Int. J. **2**(1), 4–26 (2009)
23. Kuusinen, K., Mikkonen, T., Pakarinen, S.: Agile user experience development in a large software organization: good expertise but limited impact (2012)
24. Sharp, J.H., Ryan, S.D.: Global agile team configuration. J. Strat. Innov. Sustain. **7**(1), 120–134 (2011)
25. Alzoubi, Y.I., Gill, A.Q.: Agile global software development communication challenges: a systematic review (2014)
26. Paasivaara, M., Durasiewicz, S., Lassenius, C.: Distributed agile development: using scrum in a large project, pp. 87–95 (2008)
27. Niazi, M., et al.: Establishing trust in offshore software outsourcing relationships: an exploratory study using a systematic literature review. IET Softw. **7**(5), 283–293 (2013)
28. Colomo-Palacios, R., et al.: Project managers in global software development teams: a study of the effects on productivity and performance. Softw. Qual. J. **22**(1), 3–19 (2014)
29. Guo, Y., Seaman, C.: A survey of software project managers on software process change, p. 7 (2008)
30. Martini, A., Pareto, L., Bosch, J.: Communication factors for speed and reuse in large-scale agile software development. In Proceedings of the 17th International Software Product Line Conference. ACM, Tokyo, Japan, pp. 42–51 (2013)
31. Versionone. State of agile survey. [cited 2016 25 Sep]. Available from: https://www.versionone.com/pdf/2011_State_of_Agile_Development_Survey_Results.pdf
32. Jain, R., Suman, U.: A systematic literature review on global software development life cycle. SIGSOFT Softw. Eng. Notes **40**(2), 1–14 (2015)
33. Lane, M.T., Agerfalk, P.J.: On the suitability of particular software development roles to global software development, pp. 3–12 (2008)
34. Betz, S., Makio, J., Stephan, R.: Offshoring of software development—methods and tools for risk management. In: International Conference on Global Software Engineering (ICGSE 2007) (2007)

35. Khan, A.A., Basri, S., Dominic, P.D.D.: A propose framework for requirement change management in global software development. In: 2012 International Conference on Computer & Information Science (ICCIS) (2012)
36. Niazi, M., et al.: GlobReq: a framework for improving requirements engineering in global software development projects: Preliminary results. In: 16th International Conference on Evaluation & Assessment in Software Engineering (EASE 2012) (2012)
37. Lopez, A., Nicolas, J., Toval, A.: Risks and safeguards for the requirements engineering process in global software development. In: 2009 Fourth IEEE International Conference on Global Software Engineering (2009)
38. Silva, F.Q.B.d., et al.: Challenges and solutions in distributed software development project management: a systematic literature review. In: 2010 5th IEEE International Conference on Global Software Engineering (2010)
39. Khan, A.A., Keung, J.: Systematic review of success factors and barriers for software process improvement in global software development. IET Softw. **10**(5), 125–135 (2016)
40. Ahmad, M.A., Ubaidullah, N.H., Lakulu, M.: Current practices in monitoring software development process in Malaysia (2014)
41. Ali, R.Z.R.M., Ibrahim, S.: An integrated software process assessment for Malaysia's SME organizations (2011)
42. Almomani, M.A.T., et al.: Software development practices and problems in Malaysian small and medium software enterprises: a pilot study (2015)
43. Baharom, F., Deraman, A., Hamdan, A.R.: A survey on the current practices of software development process in Malaysia (2005)
44. Kamaroddin, J.H., et al.: The adoption of software development methodologies among IT organization in Malaysia (2012)
45. Mansor, Z., et al.: A survey on cost estimation process in Malaysia software industry (2012)
46. Mohamed, S.F.P., Baharom, F., Deraman, A.: An exploratory study on current software development practices in Malaysia focusing on agile based software development (2013)
47. Patton, M.Q.: Qualitative Research & Evaluation Methods. SAGE Publications, Thousand Oaks (2002)
48. Woods, M.: Interviewing for research and analysing qualitative data: an overview [cited 2016 6 June]. Available from: http://owll.massey.ac.nz/pdf/interviewing-for-research-and-analysing-qualitative-data.pdf (2016)
49. Thomas, D.R.: A general inductive approach for analyzing qualitative evaluation data. Am. J. Eval. **27**(2), 237–246 (2006)
50. Burnard, P., et al.: Analysing and presenting qualitative data. BDJ **204**(8), 429–432 (2008)
51. Dixon, H.: Qualitative data analysis using NVivo [cited 2016 20 June]. Available from: http://www.slideshare.net/HelenDixon1/qualitative-data-analysis-using-n-vivo (2016)
52. QSR International Pty Ltd. NVivo 11—getting started guide [cited 2016 1 July]. Available from: http://download.qsrinternational.com/Document/NVivo11/11.3.0/en-US/NVivo11-Getting-Started-Guide-Starter-edition.pdf (2016)
53. Lane Medical Library. Qualitative data analysis with NVIVO [cited 2016 1 June]. Available from: https://www.youtube.com/watch?v=0YyVySrV2cM (2016)
54. QSR International Learning how to explore and visualize your data with NVivo | NVivo Brown Bag Webinar [cited 2016 1 Augest]. Available from: https://www.youtube.com/watch?v=Ez7PB6ZIA5I (2016)
55. Rowe, D.: NVivo 10 coding [cited 2016 20 June]. Available from: https://www.youtube.com/watch?v=4crQbeHKhtk (2016)
56. Stanford University. Using NVivo for qualitative data analysis [cited 2016 17 March]. Available from: http://web.stanford.edu/group/ssds/cgi-bin/drupal/files/Guides/UsingNVivo9_0.pdf (2016)
57. Toolis, E.: Analyzing qualitative data using NVivo: an introduction [cited 2016 10 July]. Available from: http://csass.ucsc.edu/images/NVivo.pdf (2016)
58. Asnawi, A.L., Gravell, A.M., Wills, G.B.: An empirical study: understanding factors and barriers for implementing agile methods in Malaysia (2010)

59. Khaleel, Y., Sulaiman, R.: A system development methodology for erp system in SMEs of Malaysian manufacturing sectors. J. Theor. Appl. Inf. Technol. **47**(2), 504–513 (2013)
60. Hastie, S., Wojewoda, S.: Standish Group 2015 Chaos Report—Q&A with Jennifer Lynch [Online]. Available: https://www.infoq.com/articles/standish-chaos-2015 (2015). Accessed 26 Sept 2016
61. THE STANDISH GROUP. The Standish Group Report Chaos [Online]. Available: https://www.projectsmart.co.uk/white-papers/chaos-report.pdf (2014). Accessed 12 April 2017

Automated Scheduling of Hostel Room Allocation Using Genetic Algorithm

Rayner Alfred and Hin Fuk Yu

Abstract Due to the rapid growth of the student population in tertiary institutions in many developing countries, hostel space has become one of the most important resources in university. Therefore, the decision of student selection and hostel room allocation is indeed a critical issue for university administration. This paper proposes a hierarchical heuristics approach to cope with hostel room allocation problem. The proposed approach involves selecting eligible students using rank based selection method and allocating selected students to the most suitable hostel room possible via the implementation of a genetic algorithm (GA). We also have examined the effects of using different weight associated with constraints on the performance of the GA. Results obtained from the experiments illustrate the feasibility of the suggested approach in solving the hostel room allocation problem.

Keywords Scheduling · Hostel space · Genetic algorithm

1 Introduction

A hostel management system is responsible for allocating students to available hostel spaces and managing students' resident and hostels information. The rapid rise in the student population of the tertiary institutions over the years in developing countries has become an inevitable challenge to university administration due to limited hostel spaces available in the university. Therefore, in order to cope with the issues above, optimization of the scheduling process for hostel room allocation is a must. Hostel room scheduling is the problem of allocating eligible students to hostel rooms while satisfying specified constraints and the scheduling process is

R. Alfred (✉) · H. F. Yu
Knowledge Technology Research Unit, Faculty of Computing and Informatics, Universiti Malaysia Sabah, Jalan UMS, 88400 Kota Kinabalu, Sabah, Malaysia
e-mail: ralfred@ums.edu.my

H. F. Yu
e-mail: roronuozoro@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_11

automated [2]. Automated scheduling is usually executed using a scheduling algorithm to generate a sequence of actions in order to achieve certain objectives. Some of the applications or problems which require automated scheduling are flow shop scheduling problem (FSSP) and job shop scheduling problem (JSSP) [1] including Hostel Space Allocation Problem [2]. Some examples of technique or method for dealing with scheduling problem are hybrid metaheuristic algorithm [3], relax and fix heuristics [4], and hierarchical heuristics approach [2]. The reason why these proposed techniques are advantageous is mainly because of the simplicity, flexibility, inexpensive computational cost, and derivation-free mechanism of heuristic or metaheuristic [5]. In this paper, a hierarchical heuristics approach will be proposed and implemented to optimize the scheduling process of hostel room allocation. Hierarchical heuristics approach is chosen because it is able to provide good results in dealing with multilevel allocation process of hostel space allocation problem [2].

The rest of the paper is organized as follow. Section 2 discusses related works with regard to implementation of GA-based heuristic in dealing with scheduling problems. Section 3 presents the design of the proposed hierarchical heuristics approach while the experimental setup is discussed in Sect. 4. This paper is concluded in Sect. 5.

2 Related Works

A large amount of research and study has been conducted in a variety of domains of scheduling problem such as JSSP [6], shelf space allocation problem (SSAP) [7], timetabling problem [8], knapsack problem [9], and bin-packing problem [10]. Yaqin et al. [6] studied on the job shop scheduling with multiobjectives-based on GA and according to the research, GA is relatively faster than other methods used for scheduling which are usually based on single point search as GA do multipoint searches on the whole population simultaneously. Their studies show that GA is feasible in optimizing NP-hard or combinatorial optimization problem. Castelli and Vanneschi [7] proposed a hybrid algorithm that combines GA and variable neighbor search (VNS) algorithm to increase the explorative ability of GA to prevent GA from being stuck in local optimum and results obtained from their experiments had shown the enhancement in explorative ability of GA. Yang et al. [9] introduced attribute reduction of rough sets into the crossover of GA to tackle multidimensional knapsack problem which is also known as the multi-constraints knapsack problem. The steps of the GA proposed by them are different than the steps in typical GA. If reduct is found in the genes of the parent chromosome, the particular genes in the reduct will be the points selected to crossover whereas in the event of no reduct is found, the single-point crossover is applied. Results obtained from their studies show that the likelihood of the proposed GA to reach a maximal solution is relatively higher than conventional GA. Other than that, Bennell et al. [10] proposed a multi-crossover genetic algorithm (MXGA) to solve

a new variant of the two-dimensional bin-packing problem where each rectangle is assigned a due date and each bin has a fixed processing time. Experiments conducted in their research shows that the MXGA performance better than the single crossover genetic algorithm in solving the bin-packing problem. One crucial similarity between the works reviewed above is such that GA is always a part of the proposed solution or method. Therefore, the motivation of our work is to implement GA in the proposed hierarchical heuristic approach in order to tackle the hostel room allocation problem. The hierarchical heuristics approach proposed in this work is different than the hierarchical heuristics approach proposed by Adewumi and Ali [2], in which instead of greedy-like heuristic, the rank based selection method is used in the selection stage. Moreover, the representation, as well as crossover and mutation mechanism of the GA, is different as well.

3 Hierarchical Heuristics Approach

This section explains the framework of the hierarchical heuristics approach. The main components of the approach, which are rank-based selection method and GA will be explained. The proposed hierarchical heuristics approach framework is shown in Fig. 1.

With respect to Fig. 1, necessary information of applicants, rooms and specified constraints will initially be provided to the framework. This information will be used during the selection stage; with the implementation of rank-based selection method, applicants will be ranked based on the weight they hold and higher ranker will be selected first. Weight hold by the applicant is determined via the degree in which the applicant satisfies the specified constraints and the order of priority of constraints will be shown in Sect. 4. A list of selected students or applicants will be produced at the end of the selection stage. During the allocation stage, the selected students will be allocated to the most suitable hostel room possible using GA-based on specified constraints. The constraints will be shown in Sect. 4.

3.1 Genetic Algorithm

A GA is a stochastic search algorithm which can be used to a variety of combinatorial optimization problems [11]. GA is also known as nature inspired algorithm for the reason that the steps in GA are based on the evolutionary process of biological organisms. GA evolves according to the principle “survival of the fittest”; this implies that offspring that carries best attributes genes will have higher chance to survive to the next generation. Generally, GA simulates the evolutionary process of biological organisms by generating individuals randomly that form up an initial population. Each individual is encoded into a string (chromosome) which represents a possible solution to a given problem. Fitness function is performed on

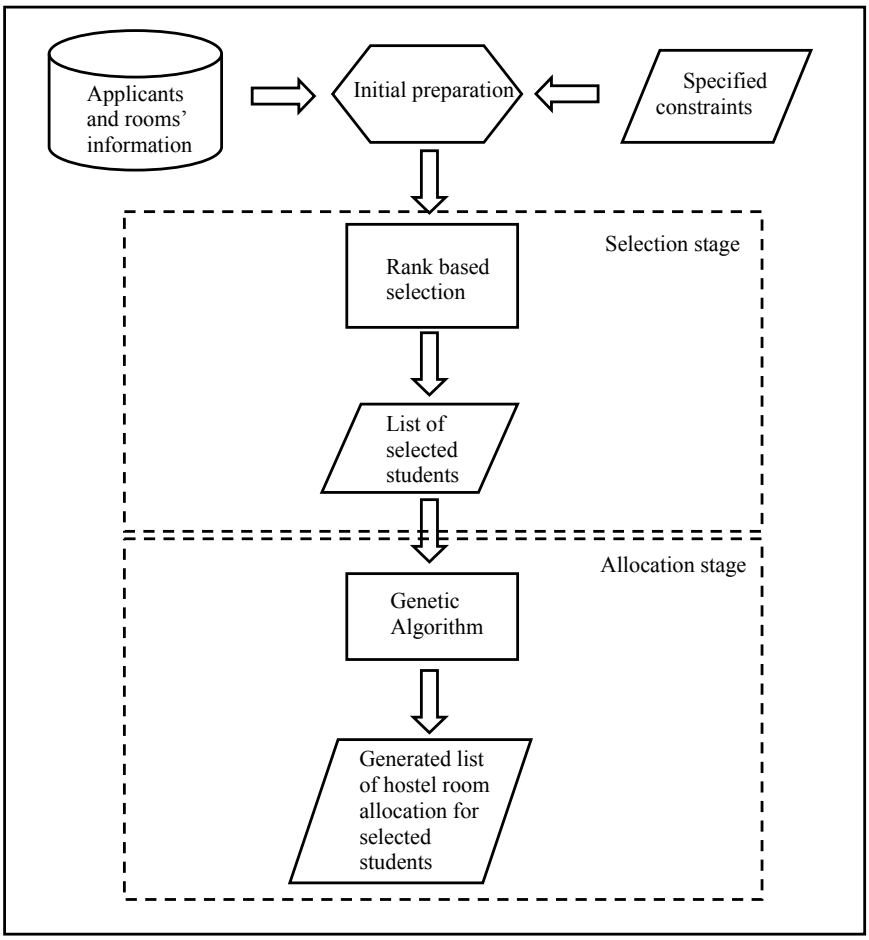


Fig. 1 Framework of hierarchical heuristics approach

each and every individual and each individual has its own fitness value. A probability of being selected during the selection process is assigned to each individual based on the individual's fitness value so as to ensure the characteristics of highly fit individuals or chromosomes are passed down to the next generation through a crossover procedure. The most general crossover method is called a one-point crossover; in which a single crossover point on both parents' chromosomes is selected first and then all information located beyond that single crossover point in either chromosome are swapped between these two parent chromosomes. The mutation is often applied to the offspring's chromosomes after crossover by altering some genes in the chromosomes, but with very low occurrence rate. In most of the case, less fit individuals will be replaced by the highly fit offspring and the new population will be reevaluated by the fitness function again. This sequence of processes will be repeated until a termination criterion is met.

3.2 Solution Encoding and Generic Operators

The encoding applied to represent candidate solutions for the GA used in the allocation stage as shown in Fig. 1 will be described in this section. The generic operators, which are crossover and mutation of the GA, will be described as well. Figure 2 illustrates how the candidate solutions are encoded. Based on Fig. 2, the integer values are the unique ID of selected students. The number of genes owned by each room varies with the types of room and as for this case, single room owns two genes while double room owns four genes. The length of the chromosome varies with the total number of rooms available. Figure 3 illustrates the process of crossover.

In Fig. 3, a self-crossover is used in order to ensure there is no repetition of integer values in the chromosome as the integer values are the unique ID of students. The crossover is done by randomly selects two portions with the same number of genes and swaps the genes of the selected portions in a random manner. The genes swapping phenomenon is controlled by a crossover rate of 0.45. On the other hand, the mutation process is almost the same as crossover process, in which the difference is that all the genes of the selected portions will be swapped completely in a random manner instead of controlled by a crossover rate of 0.45. The likelihood of the mutation process to happen is controlled by a mutation rate of 0.05.

3.3 Fitness Function

Fitness function, which is also known as evaluation function is used to evaluate individuals after initialization, crossover, and mutation steps to determine how fit the individuals are. In this work, the fitness of GA's individual will be evaluated based on the degree in which the constraints are satisfied. Each individual of the GA will have their own fitness value which is a set of real number in [0, 1]. A value of 0 indicates a complete violation of all given constraints while a value of 1 indicates no constraint is violated. The fitness of an individual, by referencing the fitness function in [2], is computed as:

$$f = \sum_{i=1}^{n} w_i u_i \in [0,\ 1] \tag{1}$$

where u_i in Eq. (1) is the utilization factor given as:

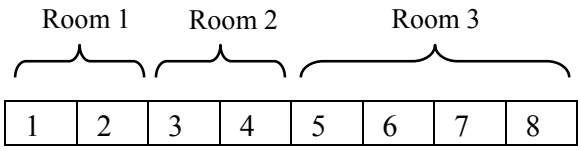


Fig. 2 Chromosome encoding of the GA in allocation stage

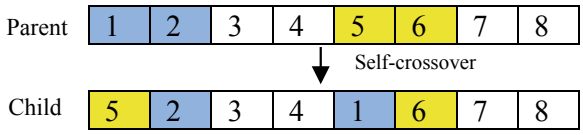


Fig. 3 Crossover process

$$u_i = \frac{1}{n} \tag{2}$$

w_i is the weight hold by gene i in a chromosome, n in Eqs. (1) and (2) is the total number of genes a chromosome has.

4 Experimental Setup

This section describes the experiments that are carried out in this work. The sample datasets for students (applicants) and rooms are shown in Tables 1, 2 and 3. The initial population of the GA will be randomly generated.

The main parameters for GA used are as follow 30 population size, 100 generations, crossover rate at 45% and mutation rate at 5%. The constraints applied to the selection and allocation process are as follow:

1. Selection:
   - Hard constraints
     - All first year students must be selected (FY).
     - Student representative must be selected (SR).
     - Uniformed units must be selected (SU).
     - Handicapped students must be selected (HS).
   - Soft constraints
     - Students with higher student development points are more likely to be selected (SD).
     - Students with higher CGPA are more likely to be selected (SC).

Table 1 Statistics of applicants per category

Category	Male	Female
First year	703	2635
Senior	297	1115
Student representative	4	8
Uniformed units	197	353
Handicapped	2	5
Food science and forestry	192	328
Others	605	3056

Table 2 Availability of rooms for male students

	Hostel				
Type	A/B	C/D	E	Angkasa	Usia
Single room capacity	52	52	52	52	
Double room capacity	148	148	148	148	0
Ground floor capacity	50	50	50	50	0

Table 3 Availability of rooms for female students

	Hostel				
Type	A/B	C/D	E	Angkasa	Usia
Single room capacity	140	140	140	140	140
Double room capacity	560	560	560	560	560
Ground floor capacity	150	150	150	150	150

2. Allocation:

- Hard constraints
  - Uniformed units must be allocated to hostel E (UH).
  - Handicapped students must be allocated to rooms situated on the ground floor (HG).
  - Single rooms are for senior students (SS).
- Soft constraints
  - Food Science (FS) and Forestry (FT) student should be allocated to hostel E (FE).
  - Students should be allocated to the hostel of their choices (SH).
  - First year students should be allocated to hostels inside the campus (SI).

Weights associated with selection and allocation constraints are shown in Tables 4 and 5. With respect to Table 4, weight associated to constraint FY is 0.51 and this is to make sure the sum of weights of other constraints is less than 0.51 as FY holds the highest priority in hard constraints. As for SR, SU, and HS, they have the same level of priority, in which fulfilling one or two or all of these constraints will give the student the same amount of weight, which is 0.25. This is to ensure the sum of weights for remaining soft selection constraints will not exceed 0.25. On the other hand, the four sets of weights' distribution for allocation constraints shown in Table 5 are used to examine the effect of weight associated constraints on the performance of the GA; during allocation stage, the results obtained from the rank

based selection are processed by the GA using these four sets of weights' distribution and the results obtained are used as a basis for optimizing the weights' distribution for allocation constraints. The optimization of weights' distribution for allocation constraints is done by simulating the GA 30 times using the same set of results obtained from rank-based selection but with 30 different sets of weights' distribution for allocation constraints which are randomly generated.

The results obtained from the rank-based selection during selection stage are shown in Table 6 where it shows the fulfillment of hard constraints as soft constraints of selection, which are SD and SC cannot be quantified. However, with the implementation of rank-based selection and normalization of student development point and CGPA, SD, and SC will always be optimized. The optimized weights' distribution for allocation constraints is shown in Table 7 while the allocation results obtained from simulating the GA using the four sets of weights' distribution and optimized weights' distribution are shown in Table 7. Based on Table 8, the performance of the GA is measured in term of constraints satisfaction in percentage

Table 4 Weights' distribution for selection constraints

Type	Constraints	Weight
Hard	FY	0.51
	SR, SU or HS	0.25
Soft	SD	0.12
	SC	0.12

Table 5 Weights' distribution for allocation constraints

Type	Constraints	Set 1	Set 2	Set 3	Set 4
Hard	UH	0.22	0.25	0.27	0.30
	HG	0.22	0.25	0.27	0.30
	SS	0.22	0.25	0.27	0.30
Soft	FE	0.12	0.09	0.07	0.04
	SH	0.11	0.08	0.06	0.03
	SI	0.11	0.08	0.06	0.03

Table 6 Selection results obtained from rank-based selection method

Category	Selected		Not selected	
	Male	Female	Male	Female
First year	703	2635	0	0
Senior	97	865	200	250
Student representative	4	8	0	0
Uniformed units	197	353	0	0
Handicapped	2	5	0	0
Food science and forestry	103	207	89	121
Others	494	2927	111	129

Table 7 Optimized weights' distribution for allocation constraints

Type	Constraints	Weight
Hard	UH	0.231
	HG	0.243
	SS	0.216
Soft	FE	0.098
	SH	0.109
	SI	0.105

Table 8 Allocation results obtained Using GA with different weights' distribution

Type	Constraints	No of students	Results				
			Set 1	Set 2	Set 3	Set 4	Optimized weight
Hard	UH	550	462	505	509	512	506
	HG	7	4	5	5	6	6
	SS	962	714	801	832	825	822
Total		1519	1180 (77.68%)	1311 (86.31%)	1346 (88.61%)	1343 (88.41%)	1334 (87.82%)
Soft	FE	310	255	202	181	172	251
	SH	3338	2890	2458	2231	2091	2718
	SI	4300	3457	3321	2989	2786	3362
Total		7948	6602 (83.06%)	5981 (75.25%)	5401 (67.95%)	5049 (63.52%)	6331 (79.65%)

instead of average fitness value of GA's population. This is because fitness value is based on weights' distribution for allocation constraints and, therefore, it is not suitable for indicating the performance of the GA. The results obtained from both selection and allocation stages show the feasibility of the suggested approach in solving the hostel room allocation problem.

5 Conclusion

This paper proposed a hierarchical heuristics approach that includes the implementation of GA to cope with hostel room allocation problem. The effect of weights associated with constraints is examined by simulating GA with multiple sets of weights' distribution for allocation constraints. Optimized weights' distribution for allocation constraints is obtained during the simulation process. Results obtained from the experiments conducted show the feasibility of the proposed approach in dealing with the problem at hand. However, improvements can still be made in the

future to increase the overall performance of the proposed approach by tuning the GA components and parameters such as encoding method, crossover method, mutation method, crossover rate, mutation rate, population size, etc.

References

1. Zhang, X-F., Koshimura, M., Fujita, H., Hasegawa, R.: Combining PSO and local search to solve scheduling problems, pp. 347–354 (2011)
2. Adewumi, A.O., Ali, M.M.: A multi-level genetic algorithm for a multi-stage space allocation problem. Math. Comput. Model. **51**(1), 109–126 (2010)
3. Zhang, Q., Manier, H., Marie-Ange, M.: A hybrid metaheuristic algorithm for flexible job-shop scheduling problems with transportation constraints, pp. 441–448 (2012)
4. Deisemara, F., Reinaldo, M., Socorro, R.: Relax and fix heuristics to solve one-stage one-machine lot-scheduling models for small-scale soft drink plants. Comput. Oper. Res. **37**, 684–691 (2009)
5. Seyedali, M., Andrew, L., Sanaz, M.: Confidence measure: a novel metric for robust meta-heuristic optimization algorithms. Inf. Sci. **317**, 114–142 (2015)
6. Yaqin, Z., Beizhi, L., Lv, W.: Study on job-shop scheduling with multi-objectives based on genetic algorithms, vol. 10, pp. 10–294 (2010)
7. Castelli, M., Vanneschi, L.: Genetic algorithm with variable neighborhood search for the optimal allocation of goods in shop shelves. Oper. Res. Lett. **42**(5), 355–360 (2014)
8. Soria-Alcaraz, J., Carpio, M., Puga, H.: A new approach of design for the academic timetabling problem through genetic algorithms, pp. 96–101 (2010)
9. Yang, H., Wang, M., Chen, Y., Huang, Y., Kao, C.: Crossover based on rough sets-a case of multidimensional knapsack problem, pp. 2411–2415 (2010)
10. Bennell, J., Soon Lee, L., Potts, C.: A genetic algorithm for two-dimensional bin packing with due dates. Int. J. Prod. Econ. **145**(2), 547–560 (2013)
11. Reeves, C.R.: Modern heuristic techniques for combinatorial problems. Blackwell Scientific, Hoboken (1993)

Evaluation of ASTER TIR Data-Based Lithological Indices in Parts of Madhya Pradesh and Chhattisgarh State, India

Himanshu Govil, Subhanil Guha, Prabhat Diwan, Neetu Gill and Anindita Dey

Abstract The present study was performed in some parts of Madhya Pradesh and Chhattisgarh State, India to compare the different quartz indices, feldspar indices and mafic indices according to Ninomiya (2005) and Guha (2016) using thermal infrared (TIR) bands (band 10, band 11, band 12, band 13, and band 14) of Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data for detecting quartz, feldspar and mafic minerals. Results showed that these indices are equally useful for delineating quartz, feldspar or mafic minerals. It was noticed from the correlation coefficients that Guha's mafic index (GMI) and Ninomiya's mafic index (NMI) presented almost the same result. Guha's quartz index (GQI) was more powerful than Ninomiya's quartz index (NQI) in identifying quartz content in alkali granites and this GQI was also comparable with the Rockwall and Hofstra's quartz index (RHQI) in identifying quartz content in alkali granite.

Keywords ASTER · Feldspar · Mafic · Quartz · Thermal infrared

1 Introduction

The visible and near-infrared (VNIR) bands and the short wave infrared (SWIR) bands of satellite data are not so useful for detecting quartz and feldspar minerals [13] whereas thermal infrared (TIR) bands are useful for the delineation of the aforesaid minerals because of the vibration of Si–O bonds [7]. Lithological mapping is an important task of geological exploration. Earth observation techniques play an important role in lithological and mineralogical mapping. TIR region are

H. Govil · S. Guha (✉) · P. Diwan
Department of Applied Geology, National Institute of Technology Raipur, Raipur, India
e-mail: subhanilguha@gmail.com

N. Gill
Chhattisgarh Council of Science and Technology, Raipur, India

A. Dey
Department of Geography, Nazrul Balika Vidyalaya, Guma, West Bengal, India

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_12

less studied than VNIR or SWIR regions for mineral identification due to the less availability of satellite sensors in the TIR domain.

Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) satellite sensor has three VNIR bands (band 1–3), six SWIR bands (band 4–9), and five TIR bands (band 10–14). ASTER VNIR and SWIR bands were used extensively in the mapping of clay, carbonate, aluminium hydroxide, and iron hydroxide [15].

In later studies, a different combination of ASTER TIR bands were frequently used in differentiating various types of rock [4, 8–10].

Ninomiya et al. [9] applied some selected lithological parameters to propose a series of ASTER TIR band based lithological indices for quartz, mafic, feldspar, and carbonate rock detection [1, 4, 9, 12]. High ratio of alkali feldspar is present in alkali feldspar rich granite while in granodiorite and tonalite this ratio becomes lower [2]. Granitoids are also characterized by a low amount of mafic mineral which is found in high amount in tonalite.

The present study tried to detect quartz, feldspar, and mafic mineral in various types of rocks using ASTER TIR bands. A number of geoscientists examined the utility of ASTER TIR bands in the formulation of rocks and mineral indices to identify the basic mineral content in the lithological variation [3–5, 9]. The published geological map was used as the reference map to get a reliable interpretation of the final results. Different lithological indices were compared to each other for enhancing the sub-variants of granitoids.

2 Study Area and Geology

The parts of Madhya Pradesh and Chhattisgarh State of east-central India were selected as the study area (Fig. 1) of the present research work. This entire region is composed of felsic and mafic rocks in significant amount. Granitoid or quartz enriched felsic rock has a combination of quartz, mafic and feldspar minerals. Geological map published from the Geological Survey of India with 1:250,000 scale was used as the reference map for analyzing the derived lithological indices.

3 Materials and Methods

Thermal bands specification of ASTER data was provided in Table 1. Emissivity values of quartz and feldspar minerals were analyzed and it was observed that the band 12 reflects the lowest emissivity value and band 10 and band 13 reflect the higher emissivity values for quartz mineral. ASTER band 11 has a lower emissivity than band 10 for feldspar mineral (Fig. 2). ASTER band 13 indicates the lowest emissivity and band 12 and band 14 have higher emissivity for mafic rocks and

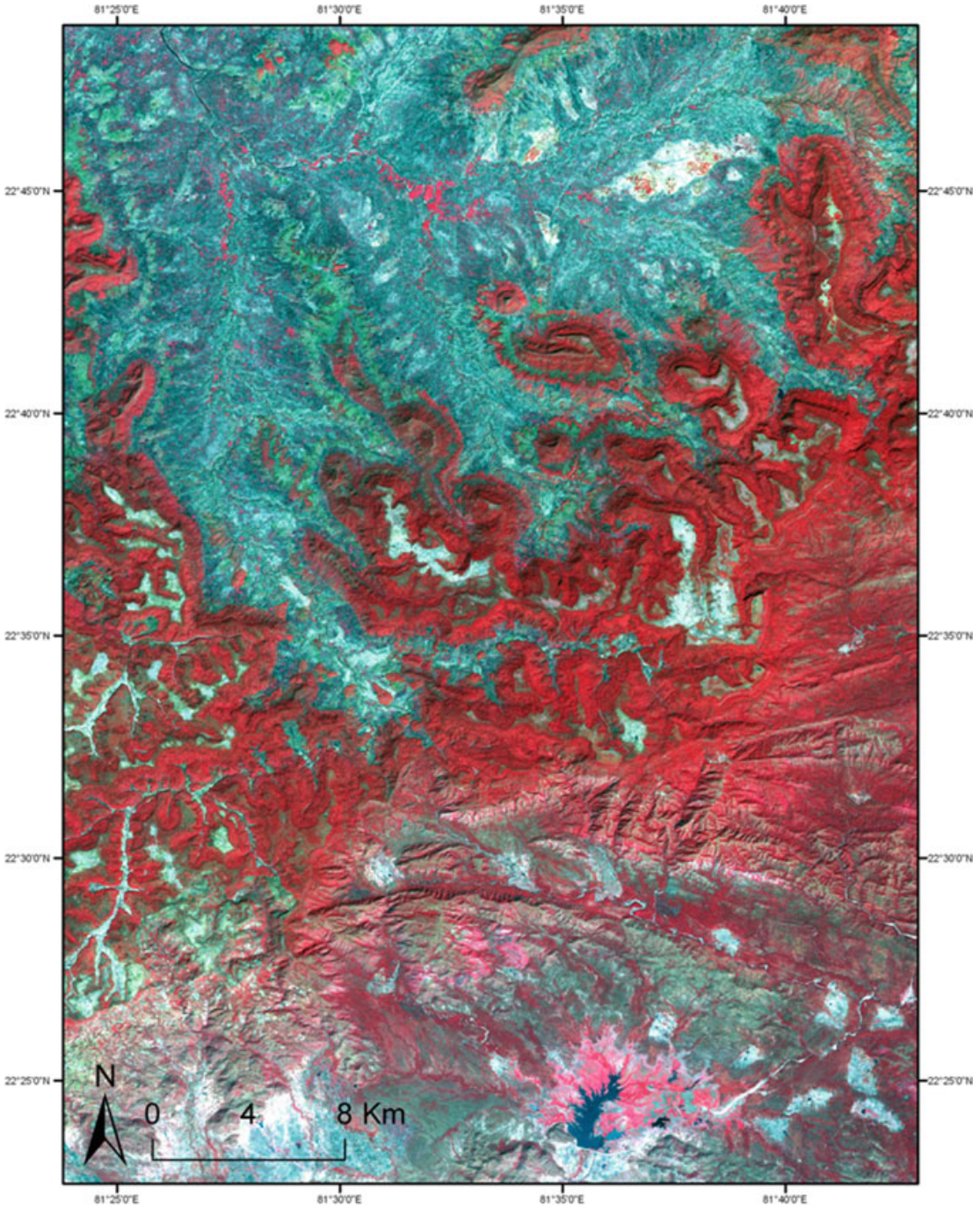


Fig. 1 Location of the study area

Table 1 General information of ASTER TIR bands

Data	Sensor	TIR bands	Wavelength (μm)	Spatial resolution (m)
ASTB070414051954	TIR	10	8.125–8.475	90
Date: 14 April 2007	TIR	11	8.475–8.825	90
	TIR	12	8.925–9.275	90
	TIR	13	10.25–10.95	90
	TIR	14	10.95–11.65	90

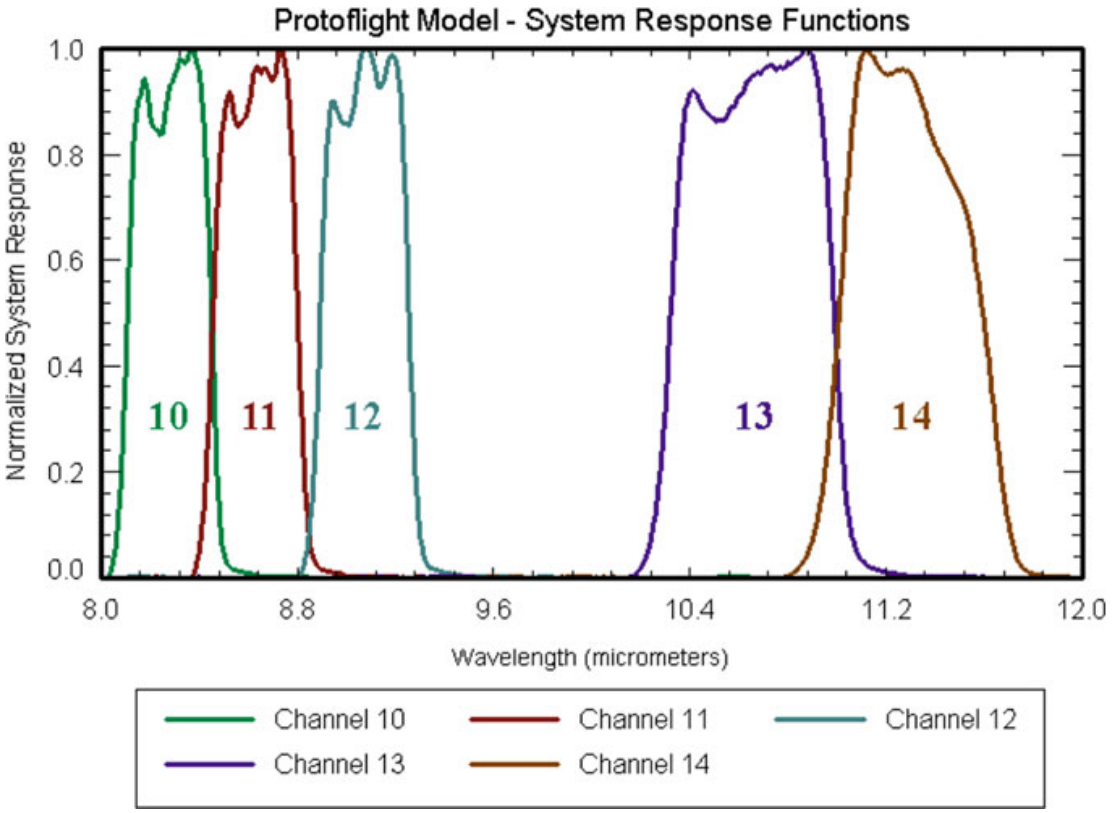


Fig. 2 Spectral ASTER TIR bands in the wavelength of the electromagnetic spectrum (*Source* NASA Jet Propulsion Laboratory)

minerals [14]. A significant reverse trend of emissivity values for TIR band 11 and band 12 was observed for quartz and feldspar minerals [14].

The lithological indices applied for detecting quartz, mafic and feldspar minerals were retrieved using TIR bands of ASTER sensor.

Guha's indices are described as follows [6]:

$$\text{Guha's quartz index (GQI)} = \frac{\text{band 10}}{\text{band 12}} \times \frac{\text{band 13}}{\text{band 12}} \tag{1}$$

$$\text{Guha's mafic index (GMI)} = \frac{\text{band 12}}{\text{band 13}} \times \frac{\text{band 14}}{\text{band 13}} \tag{2}$$

$$\text{Guha's feldspar index (GFI)} = \frac{\text{band 10}}{\text{band 11}} \times \frac{\text{band 12}}{\text{band 11}} \tag{3}$$

GQI, GMI and GFI were generally used to detect the rocks with high quartz content, high mafic minerals content, and high feldspar content, respectively. NQI and NMI were the two established lithological indices to detect quartz and mafic minerals, respectively [9]. Besides, Rockwall and Hofstra [11] proposed another index (RHQI) to detect quartz-rich rocks.

These indices are described as follows:

$$\text{Ninomiya's quartz index (NQI)} = \frac{\text{band 11} \times \text{band 11}}{\text{band 10} \times \text{band 12}} \tag{4}$$

$$\text{Ninomiya's mafic index (NMI)} = \frac{\text{band 12}}{\text{band 13}} \tag{5}$$

$$\text{Rockwall and Hofstra's quartz index (RHQI)} = \frac{\text{band 11}}{\text{band 10} + \text{band 12}} \times \frac{\text{band 13}}{\text{band 12}} \tag{6}$$

The present study compared GQI with respect to NQI in detecting quartz and GMI with respect to NMI in detecting mafic minerals. Applicability of RHQI and GQI was also examined.

4 Results

The study attempted to compare the lithological indices for identifying mafic, quartz and feldspar minerals. It was noticed that granite and quartz-rich granite were differentiated from mafic rich gneiss as yellow colour produced in the composite image of band 14 as red band, band 13 as green band and band 12 as blue band (Fig. 3a). Quartz particles were characterized with a red tint in the composite image of band 14 as red band, band 12 as green band and band 10 as blue band (Fig. 3c). NQI and NMI were the established rock bearing indices to distinguish quartz and mafic minerals in various types of lithological configurations. GQI provided a better result than NQI in the delineation of quartz content (Fig. 4a and b). Quartz enriched rocks were found as dark colour and bright colour in NQI and GQI images, respectively. This is because of the different emissivity values for quartz and feldspar in these TIR bands. GQI image was also compared to RHQI image. In RHQI image, the emittance value for quartz should be higher in band 11 [9]. Again, the emittance value for quartz was also high in band 13 compared to band 12 [6]. Thus, RHQI was used to delineate quartz content in granite and alkali granite (Fig. 4a and c). In GFI image, feldspar particles appeared as brighter in alkali granite (Fig. 4d). GMI and NMI built a strong correlation and thus these images were complementary to each other (Fig. 4e and f). It was found a relatively weak correlation (0.092) between GQI and NQI while GMI and NMI showed a strong (0.813) correlation (Table 2). However, GQI was comparable with RHQI in order to identify quartz enriched granitoids. GQI and RHQI built a very high correlation (0.808).

GFI and GQI had a very weak negative correlation (−0.090) because generally feldspar particle is increased slowly with the gradual decrease of quartz particle.

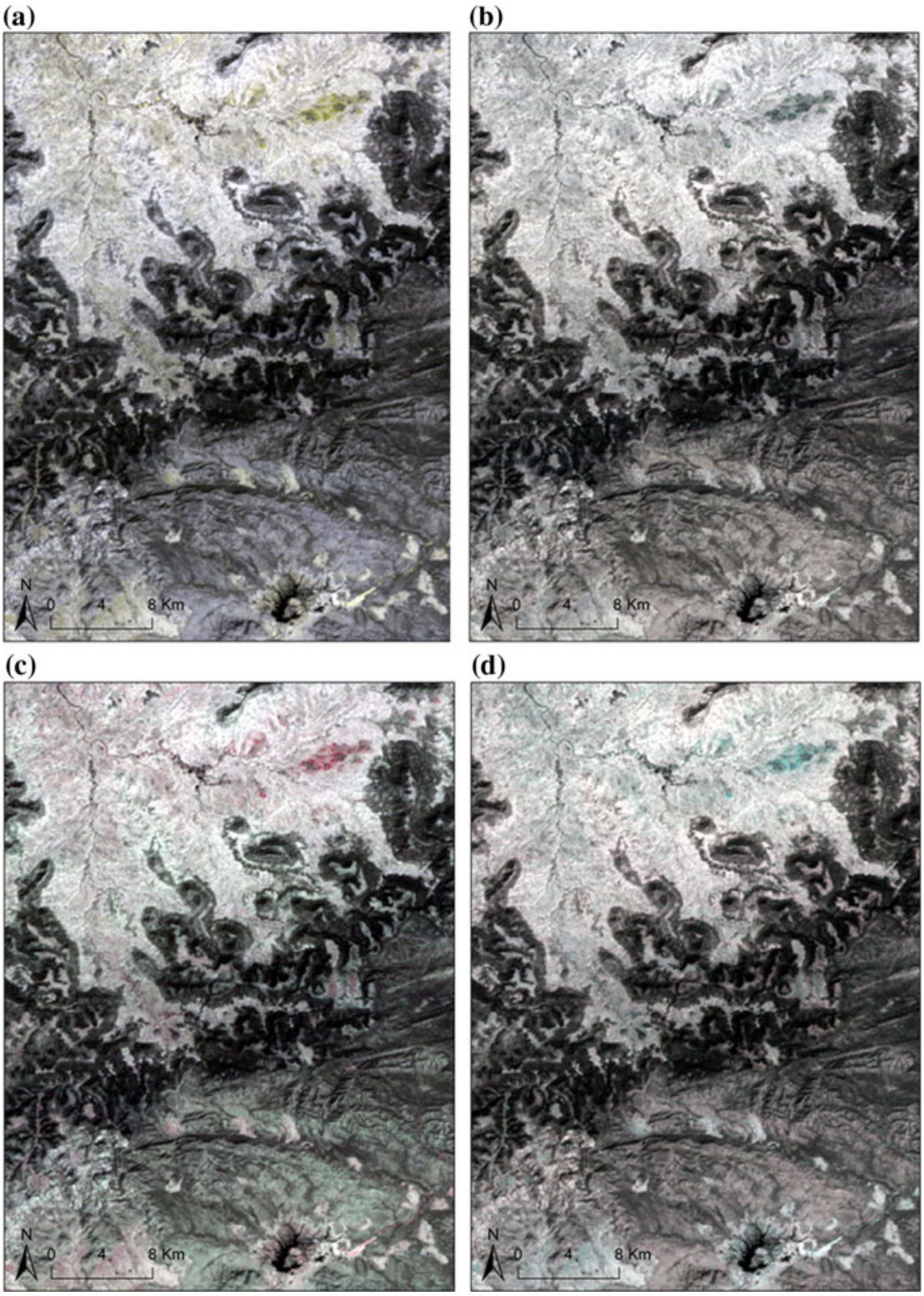


Fig. 3 ASTER TIR radiance composite images: **a** band 14 as red band, band 13 as green band, band 12 as blue band; **b** band 12 as red band, band 11 as green band, band 10 as blue band; **c** band 14 as red band, band 12 as green band, band 10 as blue band; **d** band 11 as red band, band 14 as green band, band 13 as blue band

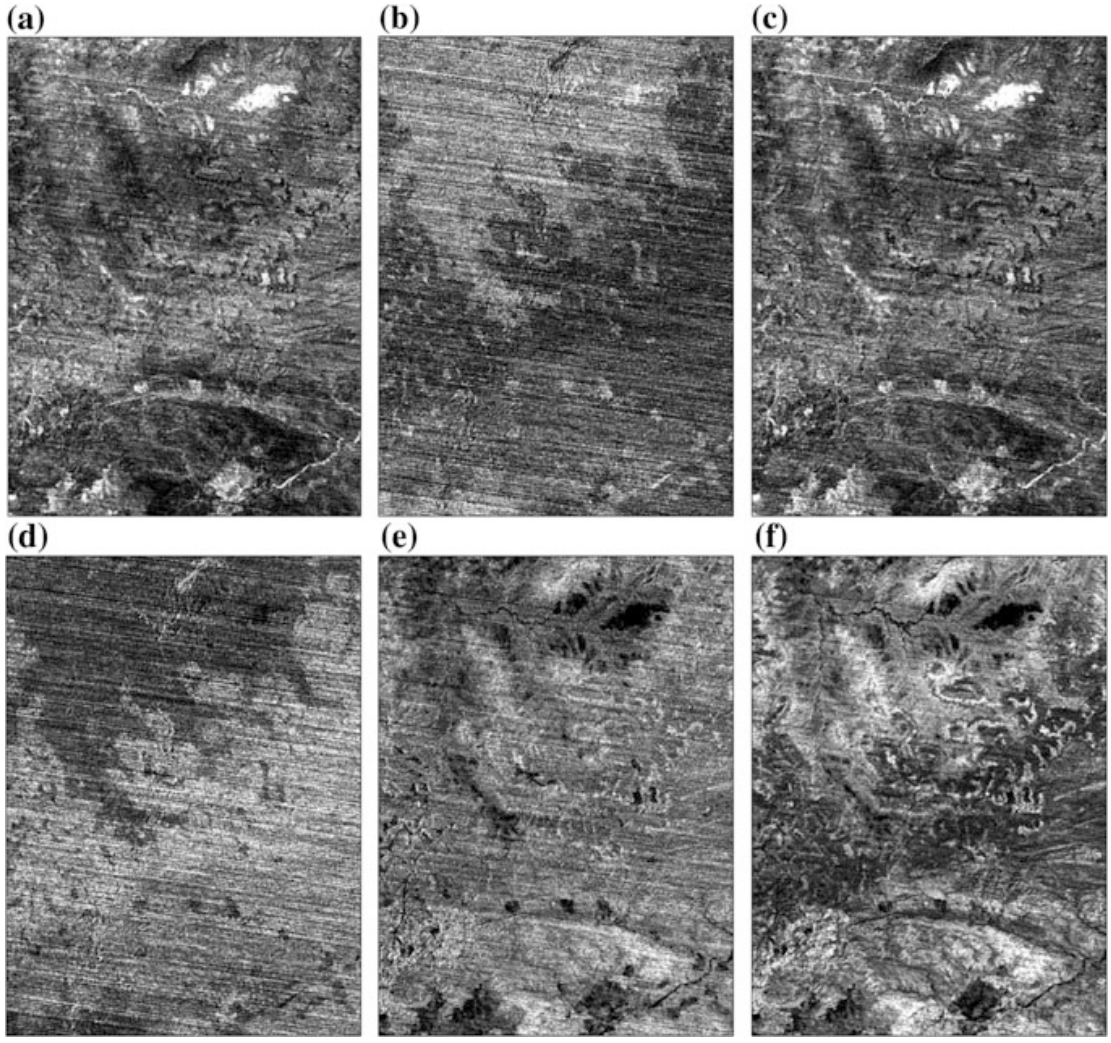


Fig. 4 **a** GQI image; **b** NQI image; **c** RHQI image; **d** GFI image; **e** GMI image; **f** NMI image

Table 2 Correlation Matrix of various lithological indices

	GQI	NQI	RHQI	GFI	GMI	NMI
GQI	1.00000	0.09179	0.80787	−0.09011	−0.80419	−0.88018
NQI	0.09179	1.00000	0.51516	−0.99992	−0.32162	−0.07198
RHQI	0.80787	0.51516	1.00000	−0.51341	−0.84916	−0.89145
GFI	−0.09011	−0.99992	−0.51341	1.00000	0.31993	0.06999
GMI	−0.80419	−0.32162	−0.84916	0.31993	1.00000	0.81306
NMI	−0.88018	−0.07198	−0.89145	0.06999	0.81306	1.00000

GMI and GFI were positively correlated with a regression value of 0.319 because feldspar particle is increased with the increase of mafic particle. GQI and GMI were negatively correlated with a very high regression value of −0.804 because quartz particle is increased rapidly with the decrease of mafic particle.

5 Discussions and Conclusions

The use of the lithological indices can be treated as a significant method for the delineation of quartz, mafic and feldspar minerals in granitoids. GQI was established as a more effective method compared to NQI (Fig. 4a and b). RHQI can retrieve quartz content almost as much as in GQI. GMI also provided a reliable comparison with NMI for delineation of mafic minerals (Fig. 4e and f). GMI and NMI were highly correlated (0.813) but the relationship became weak (0.091) between NQI and GQI (Table 2). GQI and QIRH also correlated with a high regression value (0.807) and hence can be used as important indices to determine quartz content in granite and alkali granite rocks. The results were significantly comparable to the published geological map. These indices can be evaluated in the different climatic and geologic environment to establish a more accurate conclusion.

Acknowledgements The authors are thankful to the United States Geological Survey (USGS) and the National Aeronautics and Space Administration (NASA) Jet Propulsion Laboratory (JPL).

References

1. Bertoldi, L., Massironi, M., Visona, D., Carosi, R., Montomoli, C., Gubert, F., et al.: Mapping the Buraburi granite in the Himalaya of Western Nepal: remote sensing analysis in a collisional belt with vegetation cover and extreme variation of topography. Remote Sens. Environ. **115**(5), 1129–1144 (2011)
2. Bose, M.K.: Igneous Petrology. World Press, Kolkata (1997)
3. Chen, J., Wang, A.J.: The pilot study on petrochemistry components mapping with ASTER thermal infrared remote sensing data. J. Remote Sens. **11**, 601–608 (2007)
4. Ding, C., Liu, X., Liu, W., Liu, M., Li, Y.: Mafic and ultramafic and quartz-rich rock indices deduced from ASTER thermal infrared data using a linear approximation to the Planck function. Ore Geol. Rev. **60**, 161–173 (2014)
5. Ding, C., Xuqing, L., Xiangnan, L., Liting, Z.: Quartzose–mafic spectral feature space model: a methodology for extracting felsic rocks with ASTER thermal infrared radiance data. Ore Geol. Rev. **66**, 283–292 (2015)
6. Guha, A., Kumar, V.K.: New ASTER derived thermal indices to delineate mineralogy of different granitoids of an Archaean Craton and analysis of their potentials with reference to Ninomiya's indices for delineating quartz and mafic minerals of granitoids—an analysis in Dharwar Craton, India. Ore Geol. Rev. **74**, 76–87 (2016)
7. Kahle, A.B.: Thermal inertia imaging: a new geologic mapping tool. Geophys. Res. Lett. **3**, 419–421 (1976)
8. Kalinowski, A., Oliver, S.A.: ASTER Mineral Index Processing Manual. http://www.ga.gov.au/image_cache/GA7833.pdf (2004)
9. Ninomiya, Y., Fu, B., Cudahy, T.J.: Detecting lithology with advanced spaceborne thermal emission and reflection radiometer (ASTER) multispectral thermal infrared radiance-at-sensor data. Remote Sens. Environ. **99**, 127–139 (2005)
10. Rajendran, S., Nasir, S.: ASTER spectral sensitivity of carbonate rocks—study in Sultanate of Oman. Adv. Space Res. **53**, 656–673 (2014)

11. Rockwall, B.W., Hofstra, A.H.: Identification of quartz and carbonate minerals across northern Nevada using ASTER thermal infrared emissivity data—implications for geologic mapping and mineral resource investigations in well-studied and frontier areas. Geosphere **4**, 218–246 (2008)
12. Rowan, L.C., Mars, J.C., Simpson, C.J.: Lithologic mapping of the Mordor, NT, Australia ultramaficcomplex by using the advanced spaceborne thermal emission and reflection radiometer (ASTER). Remote Sens. Environ. **99**(1–2), 105–126 (2005)
13. Salisbury, J.W., Walter, L.S.: Thermal infrared (2.5–13.5 μm) spectroscopic remote sensing of igneous rock types on particulate planetary surfaces. J. Geophys. Res. **94**(B7), 9192–9202 (1989)
14. Son, Y.-S., Kang, M.-K., Yoon, W.-J.: Lithological and mineralogical survey of the Oyu Tolgoi region, South-eastern Gobi, Mongolia using ASTER reflectance and emissivity data. Int. J. Appl. Earth Obs. Geoinf. **26**, 205–216 (2014)
15. Zhang, X., Pazner, M., Duke, N.: Lithologic and mineral information extraction for gold exploration using ASTER data in the south Chocolate Mountains (California). ISPRS J. Photogramm. **62**, 271–282 (2007)

Analyzing Linear Relationships of LST with NDVI and MNDISI Using Various Resolution Levels of Landsat 8 OLI and TIRS Data

Himanshu Govil, Subhanil Guha, Prabhat Diwan, Neetu Gill and Anindita Dey

Abstract The present study used the Normalized Difference Vegetation Index (NDVI) and the Modified Normalized Difference Impervious Surface Index (MNDISI) to determine the linear relationship between Land Surface Temperature (LST) distribution and these remote sensing indices under various spatial resolutions. Four multi-date Landsat 8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) images of parts of Chhattisgarh State of India were used from four different seasons (spring, summer, autumn and winter). The results indicate that LST established moderate to strong negative correlations with NDVI and weak negative to moderate positive correlations with MNDISI at various spatial resolutions (30–960 m). Generally, the coarser resolutions (840–960 m) possess stronger correlation coefficient values due to more homogeneity. The autumn or post-monsoon image represents the strongest correlation for LST–NDVI and LST–MNDISI at any resolution levels. The image of winter season reveals the best predictability of LST distribution with the known NDVI and MNDISI values.

Keywords Land surface temperature (LST) · Normalized difference vegetation index (NDVI) · Modified normalized difference impervious surface index (MNDISI) · Landsat · Spatial resolution

H. Govil · S. Guha (✉) · P. Diwan
Department of Applied Geology, National Institute of Technology Raipur, Raipur, India
e-mail: subhanilguha@gmail.com

N. Gill
Chhattisgarh Council of Science and Technology, Raipur, India

A. Dey
Department of Geography, Nazrul Balika Vidyalaya, Guma, West Bengal, India

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_13

1 Introduction

Generally, the spatial and temporal resolution of satellite sensor should follow the nature of surface configurations [24]. In any urban area, the high value of Land Surface Temperature (LST) is generated mostly in the Impervious Surface Area (ISA) with lower vegetation intensity [18]. Therefore, most of the thermal remote sensing methods used the Normalized Difference Vegetation Index (NDVI) [5] and ISA [1] as the significant indicators of LST, and a number of studies based on the relationships of LST with NDVI [6–8, 19], and impervious surface fraction [22], were used to explore high LST. Some scholars applied fractal techniques for analyzing LST–NDVI relationship [19]. An adjusted stratified stepwise regression method was proposed to sharpen the LST within a high-density urban land and the adjacent areas [25]. An improved error estimation method was developed considering the scale difference [4]. The relationships of LST with NDVI and Modified Normalized Difference Impervious Surface Index (MNDISI) were recently investigated for a single date image at different resolutions [11] in which the relationships became strongest at 30 m resolution. In the present paper, a detailed examination was performed by using four cloudless Landsat 8 OLI (Operational Land Imager) and TIRS (Thermal Infrared Sensor) data taken from the four different seasons (spring, summer, autumn and winter) to evaluate the relationships of LST with NDVI and MNDISI at 30–960 m spatial resolutions. The main aim of the present study was to assess the LST–NDVI and LST–MNDISI relationships generated in the satellite images of different seasons at various spatial resolution using Landsat 8 data.

2 Study Area and Data

Part of Chhattisgarh State including the capital city of Raipur and its surroundings was selected for the entire research work. The total study area extends between 20°59′24″ N to 21°35′24″ N and 81°22′48″ E to 82°01′12″ E covers a total geographical area of 4356 km^2 (Fig. 1). The study area has an elevation ranging from 147 m to 370 m. Mahanadi River flows along the east of Raipur city and the elevated southern part is covered by forest. The study area is considered as tropical wet (Savannah) type of climate. Spring season exists for only one and half months (February–March) characterized by pleasant weather. Summer or pre-monsoon months (March–June) are hot and dry and dust storms occur frequently. July–September months are considered as monsoon or rainy season. October and November months are often considered as the autumn or post-monsoon season, characterized by an excellent climatic condition with comparatively low temperature and moderate moisture content in the air. The presence of high density of green vegetation really adds an extra flavour in Chhattisgarh during autumn. Winter months (December–January) experience a cool and dry climate. The average annual

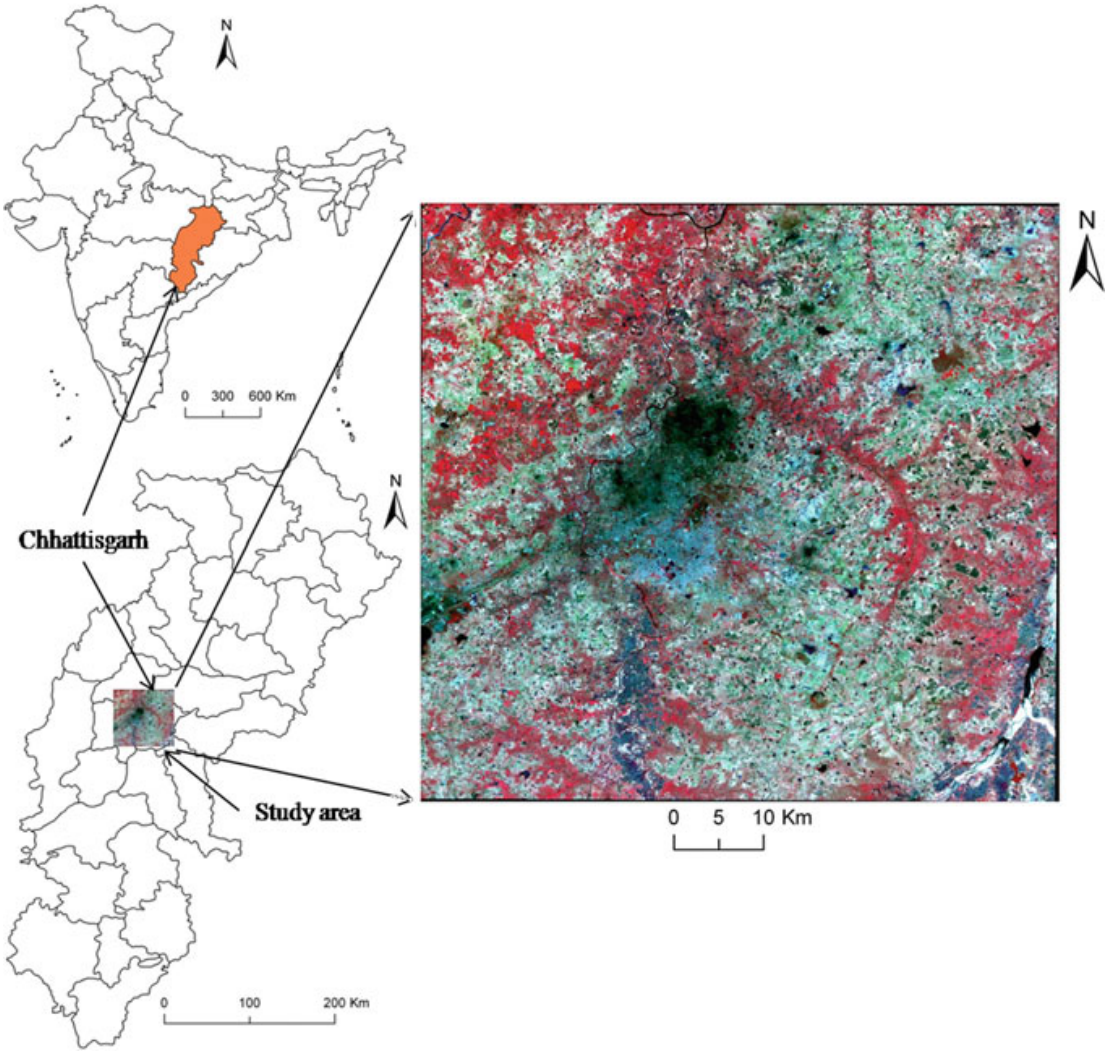


Fig. 1 Location of the study area

range of temperature is 20 °C–34 °C and average annual precipitation is 130 cm. An extensive bare land was observed throughout the study area including the capital city in the central part. The study area was also characterized by tropical mixed deciduous vegetation and red soil.

Four Landsat 8 OLI and TIRS data satellite images (Path/Row: 142/45) dated 3 February 2016, 23 April 2016, 16 October 2016 and 19 December 2016 were downloaded from the United States Geological Survey (USGS) Data Centre (Table 1) which were used as the representatives of the spring, summer or pre-monsoon, autumn or post-monsoon and winter seasons, prevails over the Chhattisgarh State of India, respectively. Landsat 8 TIRS dataset has two TIR bands (bands 10 and 11) in which band 11 has a larger calibration uncertainty. Thus, only TIR band 10 data was recommended in the present study. The TIR band 10 data was resampled to 30 m × 30 m pixel size of for all the satellite images.

Table 1 Specification of Landsat 8 OLI and TIRS satellite images

Data	Wavelength of Band 10 (μm)	Resolution of Band 10 (m)	Sun elevation (°)	Sun azimuth (°)	Cloud Cover (%)	Earth–Sun distance (astronomical unit-AU)	Path/ row	Coordinated Universal Time (UTC)	Date of acquisition
Landsat 8	10.60–11.19	100 × 100	43.37	143.17	1.74	0.98	142/045	04:56:09	3 Feb 2016
Landsat 8	10.60–11.19	100 × 100	65.99	108.18	0.01	1.00	142/045	04:55:38	23 Apr 2016
Landsat 8	10.60–11.19	100 × 100	53.59	145.43	0.01	0.99	142/045	04:56:27	16 Oct 2016
Landsat 8	10.60–11.19	100 × 100	39.72	152.38	0.21	0.98	142/045	04:56:22	19 Dec 2016

Table 2 Various indices used for the extraction of Land Use/Land Cover (LU-LC) features

Indices	Formula	Significance in LU-LC Identification	References
NDVI	(NIR–Red)/(NIR + Red)	Green vegetation extraction	[13]
MNDISI	[TIR–(MNDWI + NIR + SWIR1)/3]/ [TIR + (MNDWI + NIR + SWIR1)/3]	Impervious surface area extraction	[21]
MNDWI	(Green-SWIR1)/(Green + SWIR1)	Waterbody extraction	[23]

3 Methodology

3.1 Retrieving LST from Landsat 8 OLI Data

In this study, the mono-window algorithm was applied to retrieve LST from multi-temporal Landsat satellite image [6, 7, 14]. Ground emissivity, atmospheric transmittance and effective mean atmospheric temperature—these three parameters are needed to derive the LST using mono-window algorithm.

3.2 Extraction of NDVI and MNDISI

Various land surface biophysical parameters were applied to specify different types of land surface features [6, 7]. In this study, special emphasis was given on NDVI [9, 13] and MNDISI [10, 21] for determining the relationships with LST. NDVI was used as a vegetation index while MNDISI, based on modified normalized difference water index or MNDWI [23], acts as an index of impervious surface. These remote sensing indices were extracted by using the following formulas (Table 2).

4 Results and Discussion

4.1 Characteristics of LST, NDVI and MNDISI at Different Spatial Resolutions

There was a prominent variation of different time periods occurred in mean and Standard Deviation (STD) values of LST, NDVI and MNDISI under the different resolutions ranging from 30 to 960 m. The summer or pre-monsoon image had the maximum values of mean LST (44.26 °C) followed by spring image (27.37 °C), autumn or post-monsoon image (26.91 °C) and winter image (24.14 °C); and these

mean values were almost unaffected by the size of the pixel. But, Standard Deviation (STD) values of LST had been gradually decreasing with the increase of pixel size for each and every image. Again, the summer image had the maximum STD of LST (2.92 °C at 30 m resolution) followed by autumn image (2.09 °C at 30 m resolution), spring image (1.75 °C at 30 m resolution) and winter image (1.46 °C at 30 m resolution) and these STD values maintained a steady decreasing trend with the increase of pixel size, e.g., 2.24 °C, 1.64 °C, 1.34 °C and 1.02 °C, respectively for summer image, autumn image, spring image and winter image at 960 m resolution.

In the case of NDVI, the autumn image had the maximum value of mean NDVI (0.34) due to the abundance of green vegetation followed by winter image (0.14), summer image (0.14) and spring image (0.14) and these mean values remain almost unchanged at different spatial resolutions. But, STD values of NDVI was continuously decreasing as the spatial resolution becomes low and this phenomenon was noticed for the four multi-date images. The autumn image represented the maximum STD value of NDVI, both at 30 m resolution (0.10) and at 960 m resolution (0.07).

There was a very negligible variation in the multi-date images for mean MNDISI value at any particular spatial resolution. At 30 m resolution, the spring image (0.55) had the maximum values of mean MNDISI followed by winter image (0.54) because of the increase in imperviousness in soil due to lack of rainfall and vegetation. The summer image (0.53) had less value of mean MNDISI than spring image and winter image while it was least in the case of autumn image (0.52) because of the high percentage of soil moisture. STD values of MNDISI had a steady descending trend with the decrease of resolution or increase of pixel size and it was a common phenomenon for all the images. The winter image had the highest STD values of MNDISI at any resolution (0.05 at 30 m to 0.03 at 960 m resolution). But, for the rest of the three images, a dynamic effect was observed with the change of resolution. Up to 300 m resolution, the autumn image had greater STD values of MNDISI than summer and spring image while 360 m resolution onwards these STD values of MNDISI become lesser.

Figure 2 represented the spatial distribution pattern of LST, NDVI and MNDISI in the highest (30 m) resolution level for all the multi-date images. It was clear that in the spring image, the higher LST values generated over the western, central, northern and southern portions of the study area while these specific zones were simultaneously characterized by low NDVI and high MNDISI value. The figure also indicated that in the summer image, the lower LST values were found in the southern and the southeastern parts and the higher LST values were found in the northern portion of the study area. The NDVI values were normally inversely correlated with the LST values while MNDISI values were not very much significantly correlated to the LST values. In the autumn image, the central and western parts are characterized by high LST, low NDVI and high MNDISI values. The

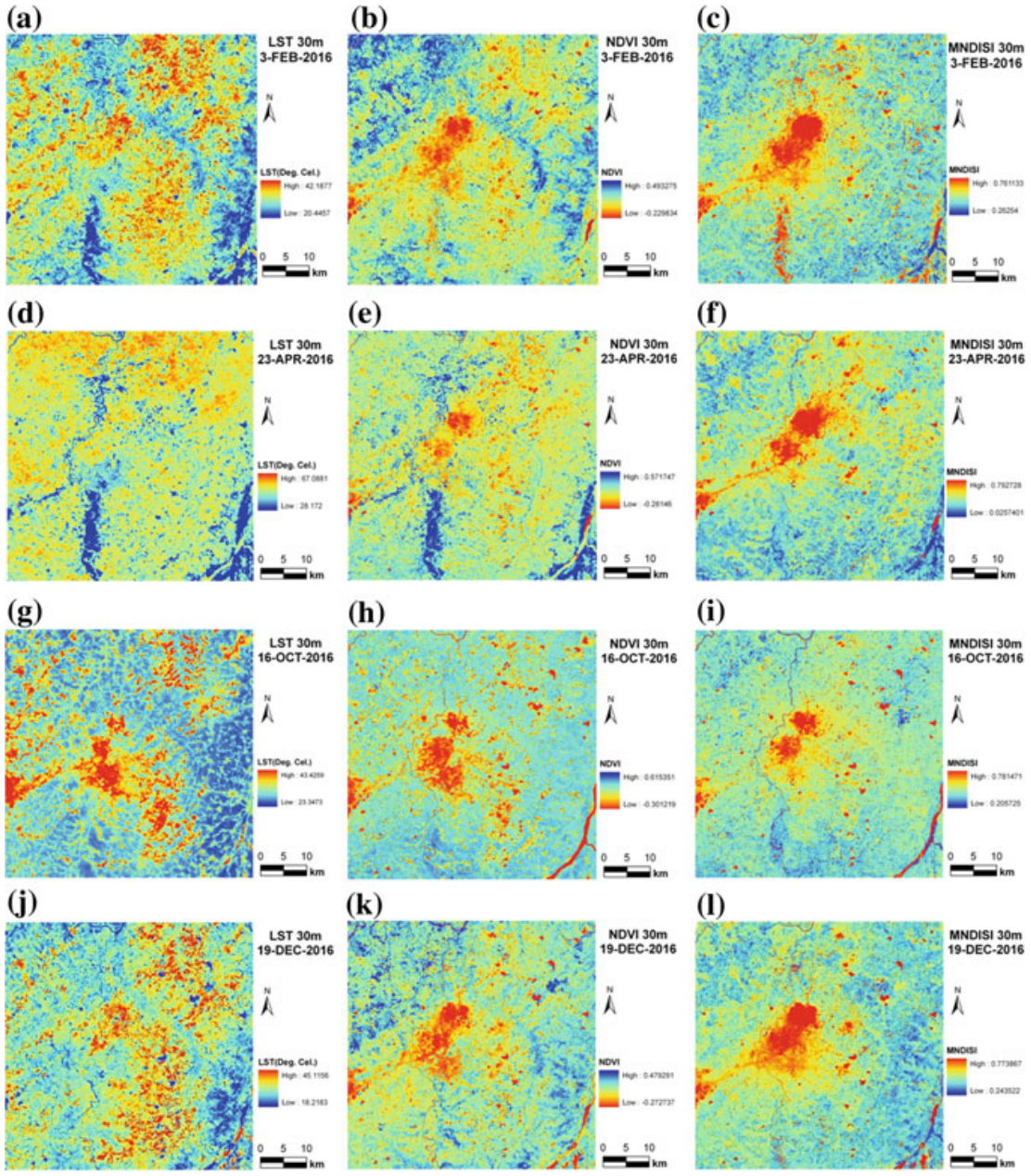


Fig. 2 Spatial distribution of LST, NDVI and MNDISI at 30 m resolution for spring image (**a**–**c**), summer image (**d**–**f**), autumn image (**g**–**i**) and winter image (**j**–**l**)

results remain almost similar at 30 m and 960 m resolution levels. In the winter image, no such evidence of significant changes was found due to the pixel size. The central part was characterized by high LST, low NDVI and high MNDISI values. Conversely, the northwest and the southeast portions had low LST, high NDVI and low MNDISI values.

4.2 Validation of Derived LST from Landsat 8 Data with Respect to MODIS Data

In the present study, MODIS data (MOD11A1 of 1 km resolution) was applied for the validation of LST retrieved from the aggregated Landsat 8 data of 960 m. Here, due to some missing pixels, MODIS data of exact corresponding date were not matched for LST validation process. MOD11A1 data (daily level 3 LST product) of the just preceding and the just following dates of Landsat 8 data were used in this study. The average MODIS LST values of the aforementioned two dates were taken for the final validation. No rainfall or strong winds occurred in between the acquisition of the Landsat 8 and MODIS imageries and almost similar weather conditions were noticed for the used Landsat 8 data and the corresponding MODIS data. Table 3 presented a strong correlation of retrieved LST from aggregated Landsat 8 data (960 m) and the corresponding MOD11A1 data (1 km) for multi-date satellite images.

4.3 Relationships Between LST–NDVI and LST–MNDISI Under Different Spatial Resolutions

The correlation coefficients between LST–NDVI and LST–MNDISI for various levels of spatial resolution (30–960 m) in four multi-date imageries were determined through pixel-by-pixel linear regression analysis. The 30 m LST, NDV, and MNDISI were aggregated to various resolution levels (60–960 m with 60 m intervals). The Pearson's correlation coefficients (significant at 0.001 level using one-tailed Student's *t* test) between LST–NDVI and between LST–MNDISI under different resolution levels and different dates were clearly understood from Table 4. The LST values had developed a negative correlation with the NDVI values at all resolution levels for all images. The three highest negative correlations were found in between 840 and 960 m resolution levels and the three lowest negative correlations were found in between 30 and 120 m resolution levels for all multi-date images. A steady and gradual increasing trend of negativity was observed in LST–NDVI correlation analysis with the decrease in resolution levels, i.e., weaker correlations developed at the higher or finer resolutions while stronger correlations developed at the lower or coarser resolutions and it appeared for all images. The

Table 3 Validation of LST (°C) retrieved from Landsat 8 data (960 m resolution) with corresponding MODIS data

	3 February 2016	23 April 2016	16 October 2016	19 December 2016
Correlation coefficient	0.74	0.79	0.85	0.73

Table 4 Correlation coefficients between LST–NDVI and between LST–MNDISI

Pixel size (m)	Number of pixels	3-FEB-2016		23-APR-2016		16-OCT-2016		19-DEC-2016	
		Correlation coefficients		Correlation coefficients		Correlation coefficients		Correlation coefficients	
		LST and NDVI	LST and MNDISI	LST and NDVI	LST and MNDISI	LST and NDVI	LST and MNDISI	LST and NDVI	LST and MNDISI
30	4840000	−0.36	−0.12	−0.50	−0.02	−0.59	0.12	−0.16	−0.14
60	121000	−0.38	−0.10	−0.53	−0.01	−0.60	0.13	−0.18	−0.12
120	303520	−0.41	−0.09	−0.58	0.01	−0.64	0.16	−0.20	−0.11
180	134658	−0.43	−0.08	−0.59	0.03	−0.66	0.18	−0.22	−0.09
240	76158	−0.44	−0.07	−0.60	0.05	−0.68	0.20	−0.24	−0.08
300	48830	−0.45	−0.06	−0.61	0.07	−0.69	0.22	−0.26	−0.06
360	33850	−0.46	−0.05	−0.62	0.09	−0.70	0.24	−0.27	−0.04
420	24960	−0.46	−0.04	−0.63	0.10	−0.71	0.25	−0.28	−0.03
480	19041	−0.46	−0.03	−0.64	0.11	−0.72	0.27	−0.29	−0.02
540	15127	−0.47	−0.03	−0.64	0.12	−0.72	0.28	−0.30	−0.01
600	12319	−0.48	−0.02	−0.65	0.14	−0.73	0.29	−0.31	0.01
660	10200	−0.48	−0.02	−0.65	0.14	−0.73	0.30	−0.32	0.01
720	8464	−0.48	−0.01	−0.66	0.15	−0.74	0.31	−0.32	0.02
780	7225	−0.49	−0.01	−0.66	0.16	−0.74	0.32	−0.34	0.04
840	6241	−0.49	0.01	−0.67	0.16	−0.75	0.34	−0.35	0.05
900	5476	−0.49	0.01	−0.67	0.17	−0.75	0.34	−0.35	0.06
960	4761	−0.49	0.01	−0.67	0.17	−0.75	0.35	−0.35	0.06

strongest negative correlation (correlation coefficient values ranging from −0.59 to −0.75) was observed in the autumn image. The summer image also had a strong negative correlation (−0.50 to −0.67). A moderate negative correlation (−0.36 to −0.49) appeared in spring image while winter image experienced a low to moderate negative correlation (−0.16 to −0.35). The scenario became slightly different in LST–MNDISI correlation analysis (Table 4). The spring image (−0.12 to −0.01) revealed a very weak negative correlation between LST and MNDISI from 30 m to 780 m resolution while only the three coarser resolutions (840–960 m) had a very weak positive correlation. In the winter image (−0.14 to 0.06), up to 540 m spatial resolution, the correlation was very weak negative and it became very weak positive from 600 to 960 m resolutions. The summer image (−0.02 to 0.17) indicated very weak to weak positive correlation between LST and MNDISI except for 30 m and 60 m resolutions (very weak negative correlation). But, only the autumn image (0.12–0.35) holds a weak to moderate positive correlation between LST–MNDISI throughout the entire range of the resolution levels (30–960 m). The three highest positive correlations between LST and MNDISI were found between 840 m and 960 m resolution levels while the three lowest correlation values were found in between 30 and 120 m resolution levels for all the images. A constant increasing trend of positive correlation is observed between LST and MNDISI with the

decrease in resolution level and it occurs for each and every image (Table 4). Therefore, the correlations between LST–NDVI and between LST–MNDISI were best observed at 840–960 m resolution levels or at the coarser resolution levels due to the presence of more homogeneity in landscape features.

The stronger correlation between LST–NDVI and between LST–MNDISI at coarser resolution levels supports the relevance of linear correlation in determining LST with the known values of NDVI and MNDISI. The low density of vegetation and high density of impervious surface may together produce a high range of LST. The correlation graphs (significant at 0.05 levels) between LST–NDVI and between LST–MNDISI at various resolution levels and for all the images were presented in Fig. 3. Table 5 revealed the regression statistics of the correlation coefficients of LST–NDVI and LST–MNDISI relationship with respect to various resolution levels. In the correlation analysis between the correlation coefficients of LST–NDVI relationship and the various resolution levels, the values of coefficient of multiple

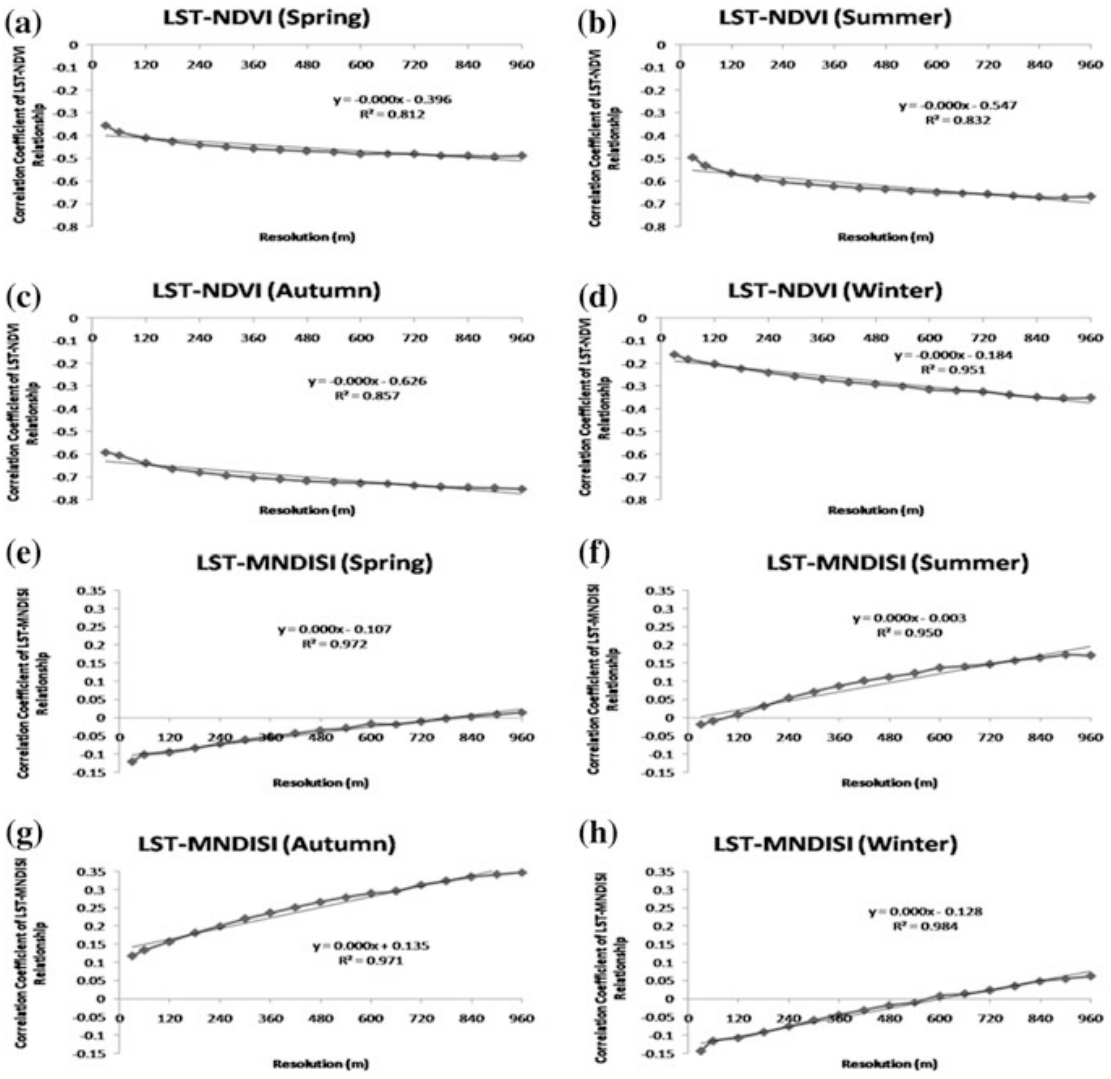


Fig. 3 Seasonal variation on LST–NDVI (**a**–**d**) and LST–MNDISI (**e**–**h**) correlation coefficients for various spatial resolutions

Table 5 Regression statistics of correlation coefficients of LST–NDVI and LST–MNDISI relationships with respect to different resolution levels for all images

	LST–NDVI correlation coefficients with various resolution levels (30–960 m)				LST–MNDISI correlation coefficients with various resolution levels (30–960 m)			
	Spring image	Summer image	Autumn image	Winter image	Spring image	Summer image	Autumn image	Winter image
Intercept	−0.40	−0.55	−0.63	−0.18	−0.11	−0.01	0.14	−0.13
Slope	−0.01	−0.01	−0.01	−0.01	0.01	0.01	0.01	0.01
Multiple R	−0.90	−0.91	−0.93	−0.98	0.99	0.98	0.99	0.99
R-squared	0.81	0.83	0.86	0.95	0.97	0.95	0.97	0.98
Adjusted R-squared	0.79	0.81	0.84	0.94	0.97	0.94	0.97	0.98
Standard error	0.02	0.02	0.02	0.01	0.01	0.01	0.01	0.01

determination or R-squared values are 0.81, 0.83, 0.86 and 0.95 for the spring image, summer image, autumn image and winter image, respectively while in the correlation analysis between the correlation coefficients of LST–MNDISI relationship and the various resolution levels, these R-squared values became 0.97, 0.95, 0.97 and 0.98 for the spring image, summer image, autumn image and winter image, respectively (Table 5). These R-squared values implicit a very strong correlation between the correlation coefficient values and the resolution levels for all the four multi-date satellite data. Hence, it can be concluded that 81.21%, 83.26%, 85.78% and 95.14% of the variance in the correlation coefficient values of LST–NDVI were predictable from the resolution levels for the images of the spring, summer, autumn and winter, respectively and 97.22%, 95.07%, 97.11% and 98.40% of the variance in the correlation coefficient values of LST–MNDISI were predictable from the resolution levels for the images of the spring, summer, autumn and winter, respectively. On the basis of multiple R, R-squared and adjusted R-squared values (Table 5), the winter image estimated the strongest correlation and may be considered as the best image under any possible resolution levels for the predictability of LST distribution with the known NDVI and MNDISI values. The overall results indicated that the NDVI and the MNDISI may be accepted as influential factors for LST distribution.

5 Conclusion

In this study, LST, NDVI and MNDISI were computed with 30–960 m resolutions for four multi-date imageries by using Landsat 8 OLI and TIRS data in parts of Chhattisgarh State, India. The linear relationships of LST–NDVI and LST–MNDISI varied with various resolution levels but yielded the higher correlation coefficient values around the coarser resolution (840–960 m) levels for all the four

images due to more homogeneity. For each data, the LST values tended to be negatively (weak to strong) correlated with the NDVI values and showed weak negative to moderate positive correlation with the MNDISI values. Autumn image (16 October 2016) reflected the strongest correlation followed by summer image while spring and winter images experienced weaker correlation for both the relationships. Moreover, a very strong negative correlation was also estimated between LST–NDVI correlation coefficients and various resolution levels. But, LST–MNDISI correlation coefficients and various resolution levels generated a very strong positive correlation. These strong relationships were sustained for all four different imageries. Amongst the four imageries, the winter image possessed the strongest correlation and ensured the best prediction level of LST distribution when NDVI and MNDISI values were known. NDVI and MNDISI may also be considered as significant variables for the prediction of LST distribution.

The study further recommends that these results may be investigated by the inclusion of other satellite imageries of low or high resolution and other LU-LC indices. Any nonlinear or nonparametric regression models may also be evaluated. A new study area may be selected with more homogeneous or heterogeneous landscape features. An important task of the present study was to observe the variation of correlation analysis for LST–NDVI and LST–MNDISI relationships in multi-date imageries. Hence, areas having extreme climatic conditions (hot, cold, wet or dry) may also be recommended for the future examination.

Acknowledgements The authors are indebted to the United States Geological Survey (USGS).

Disclosure Statement No potential conflict of interest was reported by the authors.

References

1. Arnold, C.L., Gibbons, C.J.: Impervious surface coverage—the emergence of a key environmental indicator. J. Am. Plan. Assoc. **62**(2), 243–258 (1996). https://doi.org/10.1080/01944369608975688
2. Carlson, T.N., Ripley, D.A.: On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sens. Environ. **62**, 241–252 (1997)
3. Chen, X.L., Zhao, H.M., Li, P.X., Yi, Z.Y.: Remote sensing image-based analysis of the relationship between urban heat island and land use/cover changes. Remote Sens. Environ. **104**(2), 133–146 (2006)
4. Chen, X., Li, W., Chen, J., Zhan, W., Rao, Y.: A simple error estimation method for linear-regression-based thermal sharpening techniques with the consideration of scale difference. Geo-spat. Inf. Sci. **17**(1), 54–59 (2014). https://doi.org/10.1080/10095020.2014.889546
5. Goward, S.N., Xue, Y.K., Czajkowski, K.P.: Evaluating land surface moisture conditions from the remotely sensed temperature/vegetation index measurements: an exploration with the simplified simple biosphere model. Remote Sens. Environ. **79**, 225–242 (2002). https://doi.org/10.1016/S0034-4257(01)00275-9

6. Guha, S., Govil, H., Dey, A., Gill, N.: Analytical study of land surface temperature with NDVI and NDBI using Landsat 8 OLI and TIRS data in Florence and Naples city, Italy. Eur. J. Remote Sens. **51**(1), 667–678 (2018). https://doi.org/10.1080/22797254.2018.1474494
7. Guha, S., Govil, H., Mukherjee, S.: Dynamic analysis and ecological evaluation of urban heat islands in Raipur city, India. J. Appl. Remote Sens. **11**(3), 036020 (2017). https://doi.org/10.1117/1.JRS.11.036020
8. Gutman, G., Ignatov, A.: The derivation of the green vegetation fraction from NOAA/AVHRR data for use in numerical weather prediction models. Int. J. Remote Sens. **19**(8), 1533–1543 (1998). https://doi.org/10.1080/014311698215333
9. Ke, Y.H., Im, J., Lee, J., Gong, H.L., Ryu, Y.: Characteristics of landsat 8 oli-derived NDVI by comparison with multiple satellite sensors and in-situ observations. Remote Sens. Environ. **164**, 298–313 (2015). https://doi.org/10.1016/j.rse.2015.04.004
10. Liu, C., Shao, Z., Chen, M., Luo, H.: MNDISI: a multi-source composition index for impervious surface area estimation at the individual city scale. Remote Sens. Lett. **4**(8), 803–812 (2013). https://doi.org/10.1080/2150704X.2013.798710
11. Mao, W., Wang, X., Cai, J., Zhu, M.: Multidimensional histogram-based information capacity analysis of urban heat island effect using Landsat 8 data. Remote Sens. Lett. **7**(10), 925–934 (2016). https://doi.org/10.1080/2150704X.2016.1182656
12. Markham, B.L., Barker, J.K.: Spectral characteristics of the LANDSAT thematic mapper sensors. Int. J. Remote Sens. **6**(5), 697–716 (1985)
13. Purevdorj, T.S., Tateishi, R., Ishiyama, T., Honda, Y.: Relationships between percent vegetation cover and vegetation indices. Int. J. Remote Sens. **19**, 3519–3535 (1998)
14. Qin, Z., Karnieli, A., Barliner, P.: A mono-window algorithm for retrieving land surface temperature from landsat TM data and its application to the Israel-Egypt border region. Int. J. Remote Sens. **22**(18), 3719–3746 (2001). https://doi.org/10.1080/01431160010006971
15. Sobrino, J.A., Raissouni, N., Li, Z.: A comparative study of land surface emissivity retrieval from NOAA data. Remote Sens. Environ. **75**(2), 256–266 (2001)
16. Sobrino, J.A., Jimenez-Munoz, J.C., Paolini, L.: Land surface temperature retrieval from Landsat TM5. Remote Sens. Environ. **9**, 434–440 (2004). https://doi.org/10.1016/j.rse.2004.02.003
17. Sun, Q., Tan, J., Xu, Y.: An ERDAS image processing method for retrieving LST and describing urban heat evolution: a case study in the Pearl River Delta Region in South China. Environ. Earth Sci. **59**, 1047–1055 (2010)
18. Voogt, J.A., Oke, T.R.: Thermal remote sensing of urban climates. Remote Sens. Environ. **86**, 370–384 (2003). https://doi.org/10.1016/S0034-4257(03)00079-8
19. Weng, Q.H., Lu, D.S., Schubring, J.: Estimation of land surface temperature-vegetation abundance relationship for urban heat island studies. Remote Sens. Environ. **89**, 467–483 (2004). https://doi.org/10.1016/j.rse.2003.11.005
20. Wukelic, G.E., Gibbons, D.E., Martucci, L.M., Foote, H.P.: Radiometric calibration of Landsat Thematic Mapper thermal band. Remote Sens. Environ. **28**, 339–347 (1989)
21. Xu, H.Q.: A new remote sensing index for fastly extracting impervious surface information. Geomat. Inf. Sci. Wuhan Univ. **11**, 1150–1153 (2008). https://doi.org/10.13203/j.whugis2008.11.024
22. Xu, H.Q., Lin, D.F., Tang, F.: The impact of impervious surface development on land surface temperature in a subtropical city: Xiamen, China. Int. J. Climatol. **33**(11), 1873–1883 (2013). https://doi.org/10.1002/joc.3554
23. Xu, H.X.: A study on information extraction of water body with the modified normalized difference water index (MNDWI). J. Remote Sens. **9**, 589–595 (2005)

24. Zhang, H.K., Huang, B., Zhang, M., Cao, K., Yu, L.: A generalization of spatial and temporal fusion methods for remotely sensed surface parameters. Int. J. Remote Sens. **36**(17), 4411–4445 (2015). https://doi.org/10.1080/01431161.2015.1083633
25. Zhu, S., Guan, H., Millington, A.C., Zhang, G.: Disaggregation of land surface temperature over a heterogeneous urban and surrounding suburban area: a case study in Shanghai, China. Int. J. Remote Sens. **34**(5), 1707–1723 (2013). https://doi.org/10.1080/01431161.2012.725957

Automatic Robot Processing Using Speech Recognition System

S. Elavarasi and G. Suseendran

Abstract Nowadays, speech recognition is becoming a more useful technology in computer applications. Many interactive speech-aware applications exist in the field. In order to use this kind of easy way of communication technique into the computer field, speech recognition technique has to be evolved. The computer has to be programmed to accept the voice input and then process it to provide the required output, using various speech recognition software. Speech recognition is the process of converting speech signal to a sequence of words using appropriate algorithm. This provides an alternative and efficient way for the people who are not well educated or not having sufficient computer knowledge to access the systems and where typing becomes difficult. This speech recognition technique also reduces the manpower to accept and process the commands. In our research work, we have to implement this speech recognition technique in customer care center, where many queries have to be processed every day. Some of the queries are repeated often and the responses also seem to be the same. In such cases, we have to propose a methodology to automate the query-processing activities using this speech recognition technique. The ways of how to automate the system and how to process the queries automatically are explained in our methodology with suitable algorithm.

Keywords Feature extraction · Speech modeling · Speech automation · Robot processing · Automate query processing · Interactive speech-aware applications

S. Elavarasi (✉) · G. Suseendran
Department of Information and Technology, School of Computing Sciences, Vels Institute of Science Technology & Advanced Studies (VISTAS), Chennai, India
e-mail: elavarasi_msc@yahoo.co.in

G. Suseendran
e-mail: suseendar_1234@yahoo.co.in

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_14

1 Introduction

In this decade, people want to make all the actions to be carried out in a much easier and faster way. In order to accomplish this, most of the things have to be modernized and computerized to perform the action efficiently. As the new trends emerge in each and every field, things have to be changed to accommodate the trend. In such a way, many new technologies have emerged in the computer field also. One such technology is *speech recognition* or *voice recognition.*

Speech recognition is a technology that has emerged in order to make easier the process of providing the requirements to the system by the user through speech rather than text. That is, speech recognition accepts the user's spoken words as input and then translates the words to appropriate text or commands to accomplish the task. It is also called as a software to accept and understand dictation to undertake the commands. While accepting the voice input, the system also gathers the information such as gender, expression or emotion and in some cases, the identity of the speaker. Thus, speech recognition converts the speech to text, and so it is also termed as *automatic speech recognition* or *ASR* or *speech to text.* This is also a potential for human–computer interaction.

Speech recognition is mostly used in many areas, since it is more useful for those who are not well educated or not having sufficient computer knowledge. The essential requirement for the user is to have sufficient language skills and proper pronunciation efficiency such that the user is able to process the commands. The major research challenges in this area are noise, speaker variation, language variation and vocabulary (Fig. 1).

While designing the speech recognition system, the researcher requires providing special attention toward the challenges:

- Speech module and illustration
- Speech pre-processing
- Feature extraction
- Database maintenance
- Performance evaluation.

In recent years, the enhancement and effective progress of speech recognition technology have been implemented in many areas to simplify the process of user requirements and to improve the performance. Some of the applications are:

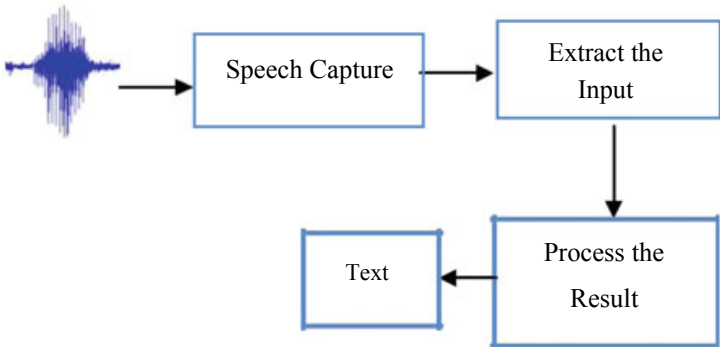


Fig. 1 Speech to text conversion

- Aerospace
- Automatic subtitling
- Hands-free computing
- Robotics
- Video games and so on.

In our research work, we have to implement the speech recognition technology in *customer care service* area to improve the customer satisfaction and experience some pre-defined operations with reduced manual efficiency. The major premise is to promote better customer service and to improve customer relationships.

Customer care service is the act of taking care of customer requirements by offering and delivering high-quality service in a short period of time. In order to provide such kind of response, more manual power has to be empowered. To avoid this dilemma, we have to provide suitable methodology to accomplish the process of responding the customer requirement through speech recognition technology automatically. This paper describes the development of an efficient speech recognition system over customer care service using various technologies.

2 Related Works

Gupta et al. discussed that with development in the requirements for installed figuring and the interest for rising implanted stages, it is necessary that the discourse acknowledgment frameworks (SRS) are accessible on them as well. PDAs and other handheld gadgets are ending up increasingly ground-breaking and reasonable also. It has turned out to be conceivable to run interactive media on these gadgets. In that paper, different methods about discourse acknowledgment framework were discussed. Likewise, it displayed the rundown of procedure with their properties of feature extraction and feature coordinating. Through this audit paper, it is discovered that MFCC is generally used to include extraction and VQ is better over DTW [1].

Singh et al. expressed that using fake neural systems (ANNs), numerical models of the low-level circuits in the human mind, to enhance discourse acknowledgment execution, through a model known as the ANN-hidden Markov model (ANNHMM) have indicated guarantee for expansive vocabulary discourse acknowledgment frameworks. Accomplishing higher recognition precision, low word mistake rate, creating discourse corpus relying on the idea of dialect and tending to the issues of wellsprings of fluctuation through methodologies like missing data techniques and convolutive non-negative matrix factorization are the significant contemplations for building up an effective ASR. In this paper, an exertion has been gained to feature the ground made so far for ASRs of various dialects and the mechanical viewpoint of programmed discourse acknowledgment in nations like China, Russian, Portuguese, Spain, Saudi Arab, Vietnam, Japan, UK, Sri Lanka, Philippines, Algeria and India [2].

Arora et al. endeavored to portray a writing audit of automatic speech recognition. It talked about past years' propels gained in order to give ground that has been

refined here of research. One of the vital difficulties for scientists is ASR precision. The speech acknowledgment system centers around troubles with ASR, essential building squares of discourse preparing, highlight extraction, discourse acknowledgment and execution assessment. The primary target of the survey paper is to expose the advancement made in ASRs of various dialects and the mechanical perspective of ASR in various nations and to thoroughly analyze the procedures used in different phases of speech acknowledgment and recognize the examined subject in this testing field. They are not displaying comprehensive depictions of frameworks or numerical definitions, yet rather, they are introducing unmistakable and novel highlights of chosen frameworks and their relative benefits and demerits [3].

Rashmi conveyed a review of various calculations that can be used in discourse acknowledgment in light of the points of interest and disservices. Likewise, it helps in picking the better calculation in light of the examination done [4].

Gamit et al. depicted that speech acknowledgment is the procedure of consequently perceiving the talked expressions of individual in view of the data content in discourse flag. This paper presents a concise overview on automatic speech recognition and examines the different characterization procedures that have been refined in this wide region of discourse handling. The target of this survey paper is to outline a portion of the notable strategies that are broadly used in a few phases of discourse acknowledgment system [5].

Shaikh Naziya et al. examined about speech advances that are limitlessly used and have boundless employments. These advances empower machines to react accurately depending on human voices, and give helpful and profitable administrations. The paper gave a diagram of the discourse acknowledgment process, its fundamental model, its application and approaches, and furthermore, examined near investigation of various methodologies that are used for discourse acknowledgment framework. The paper additionally gives a review of various strategies of discourse acknowledgment framework and furthermore demonstrates the rundown portion of the notable techniques used in different phases of discourse acknowledgment system [6].

Navneet et al. introduced the programmed discourse acknowledgment framework and examined the significant subjects and advances made in the previous 60 long stretches of research, in order to give a mechanical point of view and energy about the principal advance that has been proficient in this vital region of discourse correspondence. Following quite a while of innovative work, the exactness of programmed discourse acknowledgment stays as one of the imperative research challenges. The outline of speech recognition framework requires watchful considerations to the accompanying: definition of different sorts of discourse classes, discourse acknowledgment process, ASR configuration issues and discourse acknowledgment strategies. The target of this audit paper is to outline and look at a portion of the notable techniques used in diverse periods of talk affirmation system and perceive an investigation on subject and applications which are at the front line of this stimulating and testing field [7].

Lawrence et al. looked into some of the significant features in the innovative work of programmed discourse acknowledgment amid the most recent couple of decades to give a mechanical point of view and a valuation for the principal advance that has been made in this imperative region of data and correspondence technology [8].

Prabhakar et al. gave a description of major inventive perspective and valuation for the vital progression of talk affirmation; gave graph technique made in each period of talk affirmation, besides condense; took a gander at changed talk affirmation systems; and perceived an investigation on subjects and applications which are at the front line of this stimulating and testing field [9].

Bhavneet Kaur gave a concise presentation of SRS portraying how the innovation functions, and after that, talks about the general engineering of SRS, strategies, benefits of using this framework, major mechanical point of view and valuation for the principal advancement of discourse acknowledgment. It gives a way to deal with the acknowledgment of discourse flag using recurrence ghastly data with Mel recurrence for the change of discourse that includes portrayal in a HMM-based acknowledgment approach, and furthermore, gives a review of strategies created in each phase of discourse acknowledgment alongside the momentum and explores future options on the same. This paper depicts the real difficulties for SRS framework which have been went over by clients' criticism and different examinations, which must be settled as quickly as time permits for better execution outcome [10].

Swati Atame et al. discussed that one of the technologies used in these fields is automatic singer identification, which is used to recognize from features of the audio signal, the singer of the song or who is the one who is singing, and the genuine singer [5]. This same area of singer identification and recognition can be used in the bioinformatics whereby the voice of the singer is used to gain access to a particular singer. The system would be very much useful to singers of the song who are actual singers and also to the common man who can store their speech and can later use this input signal to access the system. For the professional singers, there are many possibilities that the original singers' voice is get mimicked which may in this situation lead to pirated copies of the voice of the singer which are then sold into the market [11].

Hori et al. exhibited strategies for discourse-to-content and speech-to-discourse programmed synopsis in light of discourse unit extraction and connection. For the previous case, a two organized outline strategies comprising essential sentence extraction and word-based sentence compaction are examined. Sentence and word units which augment the weighted total of etymological probability, measure of data, certainty measure and linguistic probability of linked units are separated from the discourse acknowledgment results and connected for star ducing outlines. For the last case, sentences, words and between-filler units are researched as units to be removed from unique discourse.

Renals et al. portrayed the advancement of a framework to translate and outline voice messages. The consequences of the exploration introduced in this paper are two overlays. Initially, a cross-breed connectionist way to deal with the voicemail interpretation undertaking demonstrates that focused execution can be accomplished using a setting autonomous framework with less parameters than those in view of blends of Gaussian probabilities. Second, a successful and vigorous blend of factual with earlier learning hotspots for term weighting is used to remove data from the decoder's yield keeping in mind the end goal to convey rundowns to the message beneficiaries by means of a GSM short message service (SMS) entryway [12].

3 Proposed Methodology

3.1 Proposed Method

Speech recognition has been the most explored point since mid 1960s and is a standout among the most famous and dynamic territory of research. In speech recognition, there exists lot of fields for research such as:

- Speech recognition
- Speaker recognition
- Speech conversion
- Feature extraction
- Noise reduction.

Speech recognition is used in many applications since it does not need any syntax, procedure or coding to access the commands. It just simply accepts the spoken words as input and then translates it into appropriate text to perform the task. Thus speech recognition technology has unlimited users with exciting range of tasks. The main theme of the speech recognition technology is to "*listen*", "*identify*", "*understand*" and "*respond*" to the spoken information. It has the potential to act like an interface between the humans and computer, that is, human–computer interface.

In our research work, we choose this speech recognition technique to be implemented in a customer care organization in order to provide better performance in a short period of time. In customer care organization, they have the task of providing response for the queries submitted by the user. The queries may be of different types such as account details, server problem, server restart, password reset, account lock, account balance, account transaction and last transaction.

In some situations, the customer may submit the queries continuously, for which the executive provides response for it. Some of the queries may be repeated often that the executive must be processed every time. In this case, it requires huge manual power with 24 $\times$ 7 services and high cost.

In order to overcome these circumstances, we have to propose a methodology to automate the process of providing responses for the queries with some pre-defined activities. This has to be carried out by programming the customer care server with some set of instructions with appropriate keywords in the database. The methodology can be executed as follows:

When the customer calls for sending query, the server accepts the query as voice input and records it. Upon receiving the input, the server converts the voice input into text and then starts to search in the database for the keyword found in the text. As a result of this database search, the appropriate instructions for the query have been found from the database, and the response to it is sent to the customer. Thus, the query has been automatically processed by the speech recognition technology like ROBOT processing, and hence our methodology can also be termed as automatic ROBOT processing (ARP) (Fig. 2).

In case the customer needs to change his/her password means, it comes under the category "*Password Reset*". The customer first calls the customer care service and asks for password reset through voice command as: "*I Need Password Reset*". This command has been recorded by the server and then processed to convert into text. Then the server starts to search from the database for the keyword "*Password Reset*". In database, there exist several instructions for the password reset category that are to be programmed by the programmer to process the request in an appropriate manner.

In that database, if there exists as "*Please tell your Username*" and "*Please tell your Date of Birth*" for "password reset" category, then the server gets the instructions and converts it into speech and then transfers to the customer as to submit their username and date of birth through speech. Upon receiving the details from the customer, the server starts converting the details into text from speech and then validates the details. Only if the information provided by the customer matches with the database, the server allows the customer to reset their password. Otherwise, he/she has to be denied (Fig. 3).

Thus, the small pre-determined activities in the customer care service have been automated with the speech recognition technique to simplify the process and to reduce the manual power. It also reduces the cost of processing the pre-determined

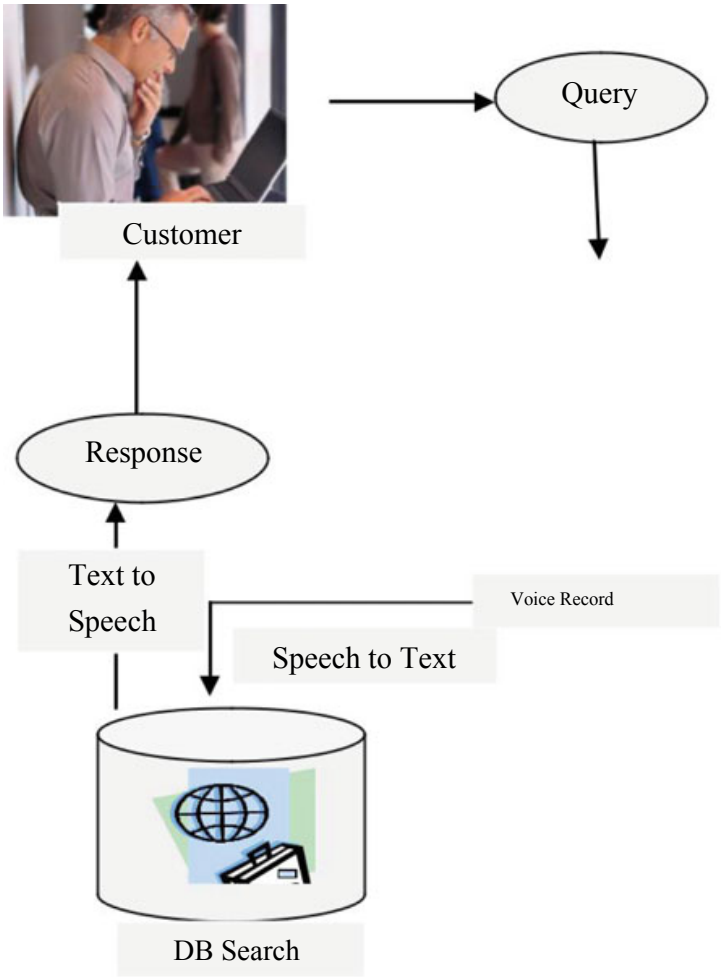


Fig. 2 Processing the customer query

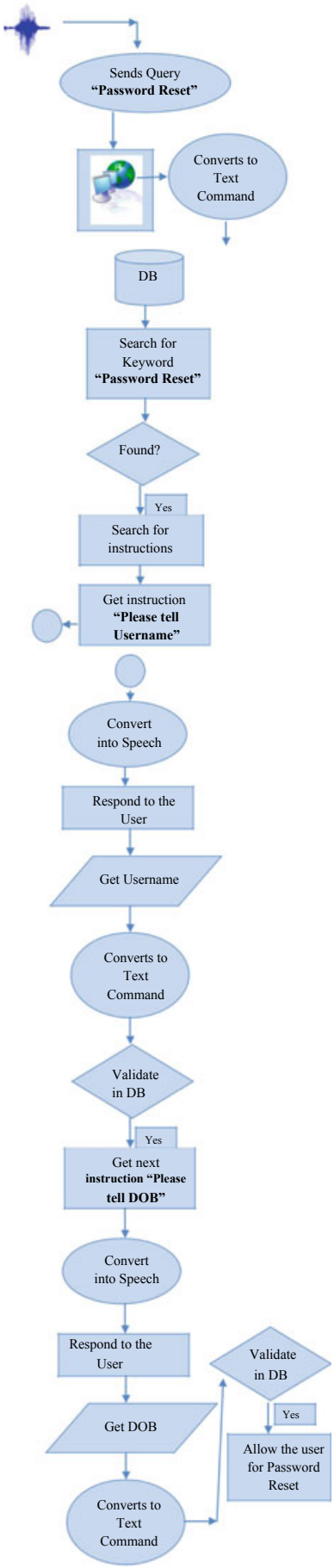


Fig. 3 Processing query for password reset

repetitive query processing. This kind of automatic processing the customer queries resembles ROBOT-type processing and thus we termed our methodology as ***automatic robot processing (ARP)***.

Along these lines, our proposed strategy acknowledges the voice information and follows the procedures to process the client inquiries effectively. The proposed strategy comprises a calculation to depict the stream of the procedure.

3.2 Algorithm-The Flow of Automatic Robot Processing (ARP)

B. Algorithm- The Flow of Automatic Robot Processing (ARP)

```
Begin
Load the server with Voice Recognition Software
 Load Database with instructions for Query Processing
If customer enters then
do
  Get the voice Command
  Translate the Voice Command to Text Command
  Search for relevant keywords
  Get appropriate instructions from the database
  Translate the instructions into speech commands
  Responds the customer with the speech command
until the query processing ends
End If
 End
```

4 Experimental Setup and Performance Metrics

We have to evaluate the performance of our proposed methodology in order to prove that our ARP query processing performs better than the existing techniques. This has to be carried out by taking some set of queries, and those queries have to be processed both manually and through our proposed ARP algorithm. The performance ratio is analyzed and the result has to be tabulated in order to prove that our ARP algorithm performs much better than the existing technique. While the

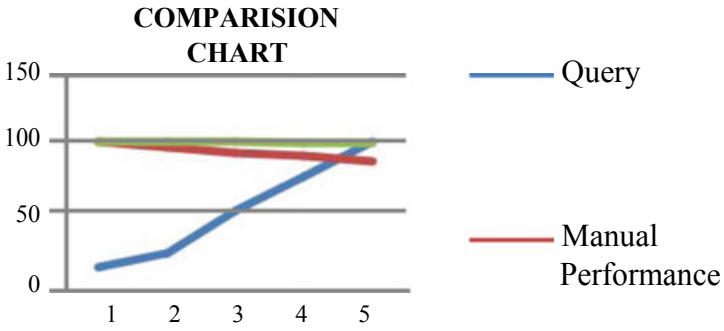


Fig. 4 Comparison chart

Table 1 Comparison results

Query	Manual performance	ARP performance
10	100	100
20	96	100
50	92	99.5
75	90	99
100	86	98

queries have to be processed manually, as the number of queries becomes increased, the ratio of performance seems low. However, when the queries have to be processed through our proposed algorithm, the performance becomes invariable. The results are tabulated and the comparison chart is shown below. With these results, it becomes clear that our ARP algorithm performs better and the queries are processed successfully (Fig. 4; Table 1).

5 Conclusions

Speech recognition is one of the most assimilating areas of artificial intelligence. This speech recognition technique helps the people who were uneducated or not having sufficient English knowledge or computer knowledge. After undertaking deep research, it has been proved that speech recognition has made a role for handling the speech-oriented activities and has been seen in many walks of life. In this computer world, various techniques are discussed about speech recognition system. This speech recognition is one of the most challenging problems to deal with. We have attempted in this paper to provide a solution to automate some pre-defined activities in the customer care organization. We also provide suitable algorithm to process this robot processing. In this case, it seems easy to access the queries in customer care center without any manual resources.

In future, our approach must be improved with much propelled route in order to handle more operations automatically with reduced manpower.

References

1. Gupta, S., Pathak, A., Saraf, A.: A study on speech recognition system: a literature review. Int. J. Sci. Eng. Technol. Res. IJSETR **3**(8), 2193–2194 (2014)
2. Ghai, W., Singh, N.: Literature review on automatic speech recognition. Int. J. Comput. Appl. **41**(8), 44–46 (2012)
3. Arora, S.J., Singh, R.P.: Automatic speech recognition: a review. Int. J. Comput. Appl. **60**(9), 37–41 (2012)
4. Rashmi, C.R.: Review of algorithms and applications in speech recognition system. Int. J. Comput. Sci. Inf. Technol. **5**(5), 5258–5261 (2014)
5. Gamit, M.R., Dhameliya, K., Bhatt, N.S.: Classification techniques for speech recognition: a review. Int. J. Emerg. Technol. Adv. Eng. **5**(2), 59–61 (2015)
6. Shaikh Naziya, S., Deshmukh, R.R.: Speech recognition system—a review. IOSR J. Comput. Eng. **18**(4), 3–8 (2016)
7. Kaur, I., Kaur, N., Ummat, A., Kaur, J., Kaur, N.: Automatic speech recognition: a review. Int. J. Comput. Sci. Technol. **7**(4), 44–46 (2016)
8. Juang, B.H., Rabiner, L.R.: Automatic speech recognition—a brief history of the technology development. Int. J. Biomed. Eng. Technol. **3**(2), 4–17 (2005)
9. Prabhakar, O.P., Sahu, N.K.: A survey on: voice command recognition technique. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **3**(5), 578–584 (2013)
10. Kaur, B.: Major challenges of voice command recognition technique. Int. J. Sci. Eng. Res. **5**(8), 216–224 (2014)
11. Atame, S., Shanthi Therese, S., Gedam, M.: A survey on: continuous voice recognition techniques. Int. J. Emerg. Trends Technol. Comput. Sci. **4**(3), 38–40 (2015)
12. Koumpis, K., Renals, S.: Transcription and summarization of voicemail speech. In: Proceedings of ICSLP (2000)

Banking and FinTech (Financial Technology) Embraced with IoT Device

G. Suseendran, E. Chandrasekaran, D. Akila and A. Sasi Kumar

Abstract In recent years the traditional financial industries have motivated for a new technology of financial technology (FinTech) clinch embraced with internet of things (IoT). The requirements of FinTech and IoT need to be integrated into new business environment. Several companies are affected because of the financial-level investments. So, there is a need to improve the next level of the business. FinTech can introduce a new service of tools and products for the emergent businesses through the internet of services which provide ideas linked in internet. Nowadays, increasing number of companies uses the IoT and creates new added values. The administrators of existing money-related organization in the direct society are dreadful by means of budgetary innovation. The social innovation is accomplished by new innovation. To make a powerful business plan and action, the FinTech and IoT are combined in order to create new innovative ideas based on the requirements.

Keywords FinTech · IoT · Innovation · Business model · Cyber security · Banking

G. Suseendran (✉) · D. Akila · A. Sasi Kumar
Department of Information Technology, School of Computing Sciences,
Vels Institute of Science Technology & Advanced Studies (VISTAS), Chennai, India
e-mail: suseendar_1234@yahoo.co.in

D. Akila
e-mail: akiindia@yahoo.com

A. Sasi Kumar
e-mail: askmca@yahoo.com

E. Chandrasekaran
Department of Mathematics, Veltech Dr. RR & Dr. SR University, Chennai, India
e-mail: e_chandrasekaran@yahoo.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_15

1 Introduction

The current generation is surrounded by smart phones, which connect devices working from home, public places and office, colleges, schools and everywhere. The consumers use this opportunity to work smarter with the help of internet of things (IoT). Banking, finance and insurance companies are easily embraced with IoT devices. The banking and financial sectors create a new way of collecting the valuable information about the customer through IoT sensor devices using smart phones. FinTech is similar to electric vehicle innovation which works with capital valuation, trading, investment and asset valuation, and provides new improving accounting systems when compared to existing financial technology. Financial engineers handle hundreds of millions of datasets. It is not easy for them to maintain all the data. So FinTech is adapting to a larger field of the IoT and will transform all the customer information to cloud using IoT devices. Most of the banking and financial sectors can rapidly improve in e-commerce by connecting cloud business through the profitable offerings of business tactics, where the cloud is stored with a lot of customers' data. In recent years, all the banks are communicating with the customer through smart phones, social networks and any new sensors, in financial technology to create some new industries [1].

1.1 IoT-Aided Banking Services

The internet of things is a big approach in financial services. The IoT connected to the concerned bank through the internet will send and receive data that are stored in cloud. Figure 1 explains the details of the customer, bank and IoT transactions [2].

Around the world, billions of devices are connected with each other. These devices share the information on the cloud with the permission of the bank and allow the entire customers to view the account details and provide access power while using smart devices. It is the easiest way of communicating with the customer and also conveying the personal information through messages and alerts awaiting works.

1.2 Benefits of IoT in Banking Services

The most important benefit of IoT in banking services is providing the credit and debit cards for easy access of the services of the banks. Also, the bank can analyze the usage of ATMs in the specific area to increase and decrease the installations of ATMs. While using IoT device, all the customer information are stored in the devices. So, the bank uses this opportunity to help in identifying the customer's business needs, like supplier, retailer and distributors [3]. Figure 2 shows the details of the bank providing valuable services to a customer.

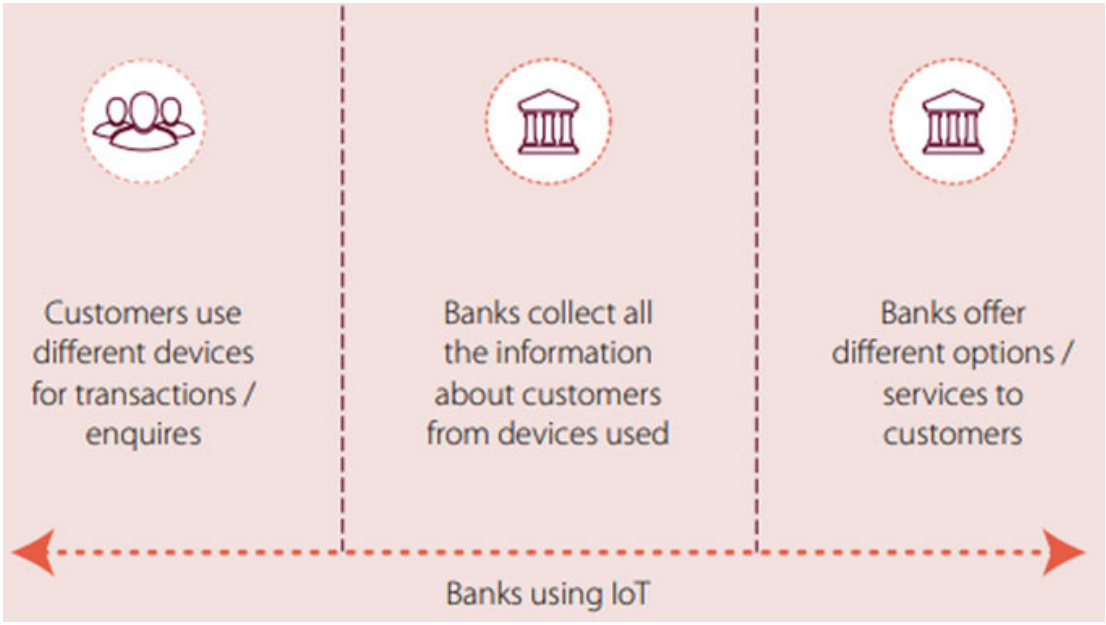


Fig. 1 How to connect customer, bank and IoT devices

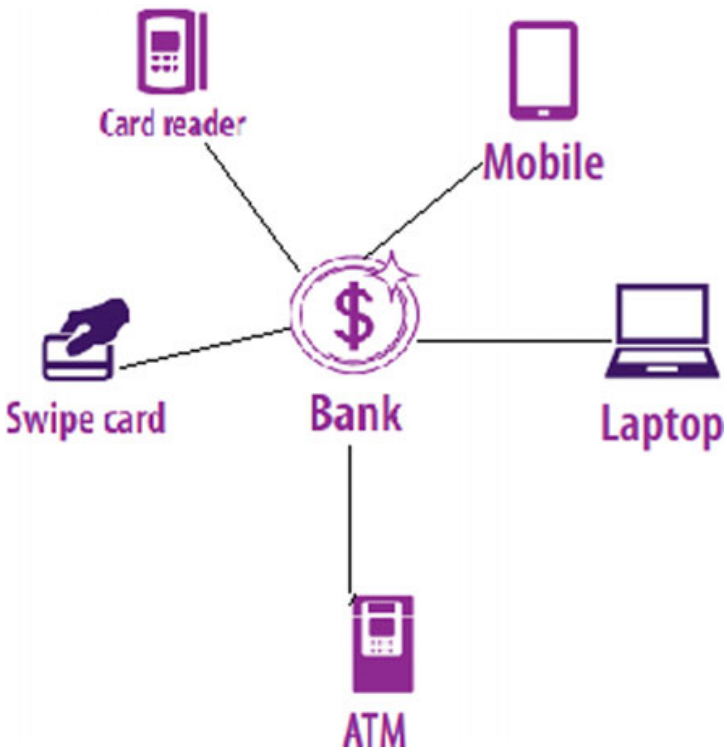


Fig. 2 Bank providing valuable services

2 Cyber Security in the Banking and FinTech on IoT

The banking and financial industry is the main target of cyber attacks. In this network world, personal data are easily accessed by the attacker. So, cybercrime is feared all around the world today. We will protect the cyber security in the banking and FinTech services industries. In Fig 3 IoT connects device through the internet and then connects to the FinTech, cloud, machine learning and industrial productivity.

Cyber Attack

Cyber attack is a careful mistreatment of a computer system, technology and networks. The hacker uses malicious code or software to alter the system and secrete code that can compromise the data to cybercrimes, such as health care documents, banking accounts details and hacking lock of the system.

As we are increasing the use of IoT connecting devices in the banking sector, the risk of cyber attacks also increases. The IoT connecting device communicates, analyzes and presents some new ways for technology. It is not only the data but other kinds of sensitive information are also shared through the IoT. Hence the risks are exponentially high [4].

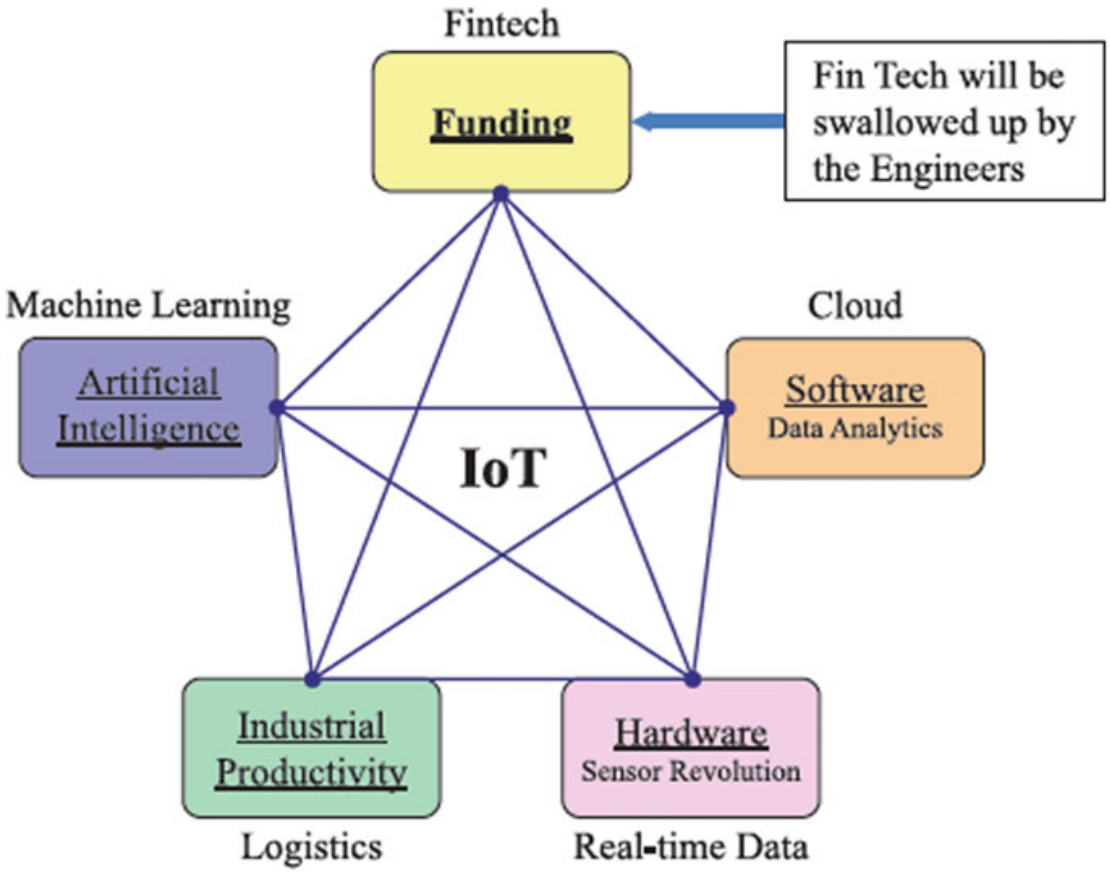


Fig. 3 Internet of things

Unencrypted Data
The data breach happens due to improper encryption, and the stolen data have immediate access.

Unprotected Third-Party Services
The internet services are an extremely worldwide connector; therefore, the cyber attacker can easily access the data of the targeted user, because third-party services are unprotected.

Unsecured Mobile Banking
In recent years, mobile banking users have increased. Using this opportunity, the mobile hacker accesses the data due to small computation time. For securing the data in mobile banking, a cryptography method of encryption and decryption is used.

A Constantly Changing Treat Landscape
In the past few years, cyber threat landscape has changed the financial services. Cybercriminals must change the low-value payment into high-value payment. So, a number of breaches affect the financial sector.

False-Positives
Anti-money laundering (AML) monitoring system is false-positive. The issue will point out the fake activity accessing time and calculated with the help of an analyst.

The Big Breach
The huge volume of financial data is increasing the risk on customers' security from hackers and cybercrime as it occurs at the night time. So beware of the breach in banking sectors.

These are major cyber security threats in banking and FinTech [5].

3 Challenges of Banking and Financial Institutions

Mainly four challenges are faced by banks and financial industries which respond consumer expectations, heavy competition on financial companies, regulatory pressure and not making enough money [6, 7].

i. **Not making enough money**
Many of the banks run on unprofitability because the financial industry is still not making enough return on investment.
ii. **Consumer expectation**
Many bank employees are feeling at most pressure because the expectations of the consumer are high. To maintain the standard they need to work hard which leads to the pressure.

iii. **Increased competition from financial technology companies**
FinTech companies is using the start up business to provide financial services to technologies. FinTech is a big challenge because of using a traditional banking system. They can change the modification and technical operation quickly, and another backup process is handled.

iv. **Regulatory pressure**
The bank requirements are continuously changing and because of that banks invest a lot of amount in some other business. So, the system processes to keep up with the higher goal requirements.

4 New Tools and Product for FinTech Sector

Traditional financial sectors are lacking because of some redundant files and were forced to move with the new tools and new products. With the advantages from new technologies, which are requirements, design and modeling, investments and delivery models are merged with IoT through the internet facilities. Table 1 lists out the new tools and products that are developed and developing in the financial sector. In 2018, 15 top Indian financial markets had radical transformation by technology and innovation in India. The FinTech sector in India with USD 1.2 billion in the year 2016 is expected to touch USD 2.4 billion in the year 2020. **MobiKwik**—Indian digital wallet company, **Capital Float**—digital financial company serving businesses, **Bank Bazaar**—online marketplace for bank loans, credit cards and insurance policies, **Incred**—web-based financial services, **PolicyBazaar**—online insurance aggregator, **Fino Payments Bank**—providing a

Table 1 New tools and product of FinTech company [9]

S. No.	FinTech sector	New tool	New products
1	Banking	Improved loan risk Monitoring Emp Debt Profit Analysis of SMS lending Financing failing applications	Record keeping with sensors Factoring and leasing Trade finance and energy finance Security for accounts Goolglization of accounts
2	Wealth management	Real-time IoT data for stocks IoT as main source for ideas IoT replaces bond derivatives	Telematics as a metric for start ups Link health monitor to wealth management Data to profit customer
3	Insurance	Insuring in high-risk areas Sensors data for smart payload Unbiased vehicle data	Pricing assets in risk-prone areas Accurate pricing product liability Weather detection reduces claims
4	Capital market	Leveraging IoT for crowd investing Block chain IPOs	New commodity data streams New banks for public access to capital

technology solutions for institutions like banks, governments and insurance companies, **CCAvenue**—popular payment gateway, **Razorpay**—a product suite that manages the entire payments lifecycle for all business are the topmost FinTech companies in India [8].

5 Use Cases of IoT—Digital Future

In banking, IoT is interconnected with connecting devices; the system that provides services does machine-to-machine (M2M) communications and is connected to lot of protocols, domains and applications. The IoT has impacted the traditional financial process such as trade financial, payments, personal financial management (PFM) and insurance [10]. There are many use cases that can be implemented in banking in the period ranging from small to long term (Fig. 4).

5.1 Account Management on Things

Biometrics (voice/touch) can make the accounts' access anywhere simpler through the digital channel. Using the new technology called Wet Ink, the customer can sign in remotely through any touch screen gadget and can be marked promptly onto physical paper with Wet Ink.

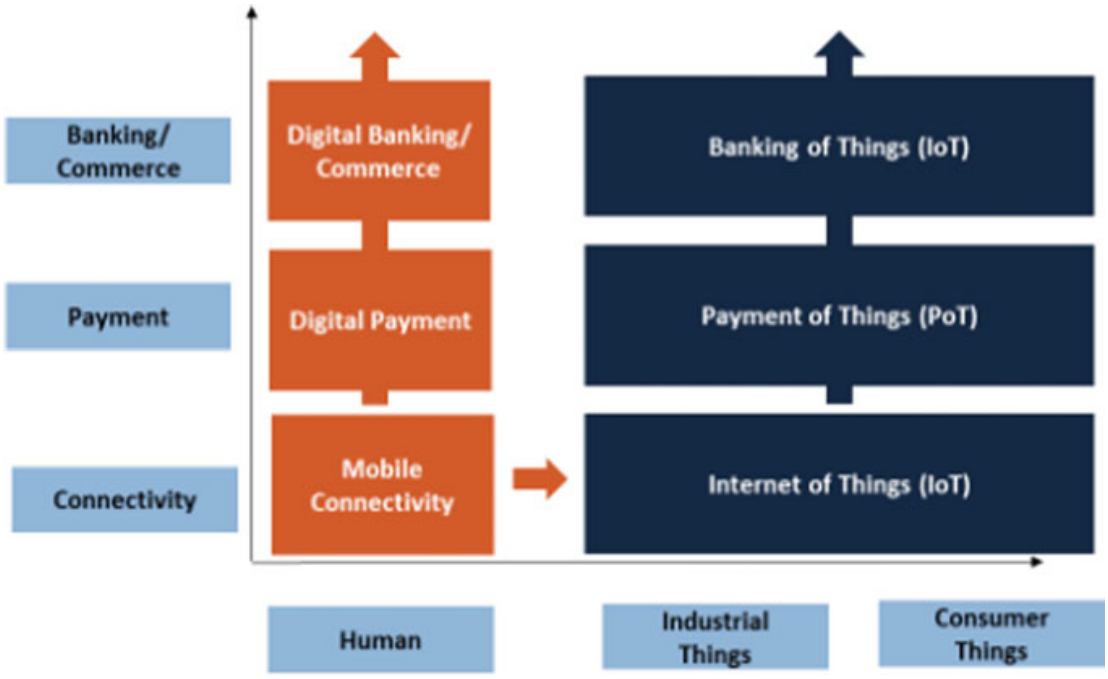


Fig. 4 IoT enabling banking of things in digital future

5.2 Leasing Finance Automation

New daily leasing models have enabled the digital assets worldwide and are effectively turning the traditional products and its services. For delineations, the rented resources could be bolted or debilitated remotely by the bank.

5.3 Smart Collaterals

An IoT device can empower the money-related banks to all the more likely command over client hypothecated resources, for example, observing their wellbeing, autos and home. Financial offerings such as manufacturing machinery, cars, building home loans as collaterals are provided in short term and can be done in a digital way automatically. For representations, in the event of EMIs are not paid, the motor can be impaired and the nature of security can be observed continuously.

5.4 Automated Payment Through Internet of Things

The traditional banking of payment transaction is automated and integrated of services. IoT can raise some conditions on security concerns and digital security in payments. So, the customer can do the payment transaction automatically like one bank to other bank or one bank to other companies.

5.5 Risk Mitigation in Trade Finance

High estimation of products are utilize the RFID monetary space. With the help of IoT the shipments including the delicate merchandise are monitored for example, restorative atoms. These executions of the hazard are relief and more educated choices at banks for including exchange back.

5.6 Wallet of Things

Wallet is associated with each device, where more devices have become digital and smart and that all banks have automated payments through IoT. For example, upkeep administrations utilizing wallet automatically can be stopping the payment transactions.

6 International Global Fintech Benchmark Report 2017

An international global FinTech benchmark report led an online review on FinTech chiefs from monetary establishments around the globe. An in-depth interview was conducted among senior executives from leading FinTech companies (Fig. 5).

FinTech technologies will come to emerging with Amazon, Google and business-based platforms by the next three years. From overview reacting investigation and huge information be arranged anticipated that would a large portion of consideration. 76% of back up plans, 65% of banks and 58% of advantage administration organizations are positioned the information examination as one of the best most FinTech advancements. Moreover, application programming interface advancements and mechanical autonomy and robo-guides are positioned in the high range of the world wide. Figure 6 shows the rising of FinTech with more enthusiasm in the upcoming next three years [7].

Insurance companies are more intent towards IoT because all valuable information and pricing are provided to the customer. The IoT connecting with the insurer person and his information from risk partners. On the off chance that any

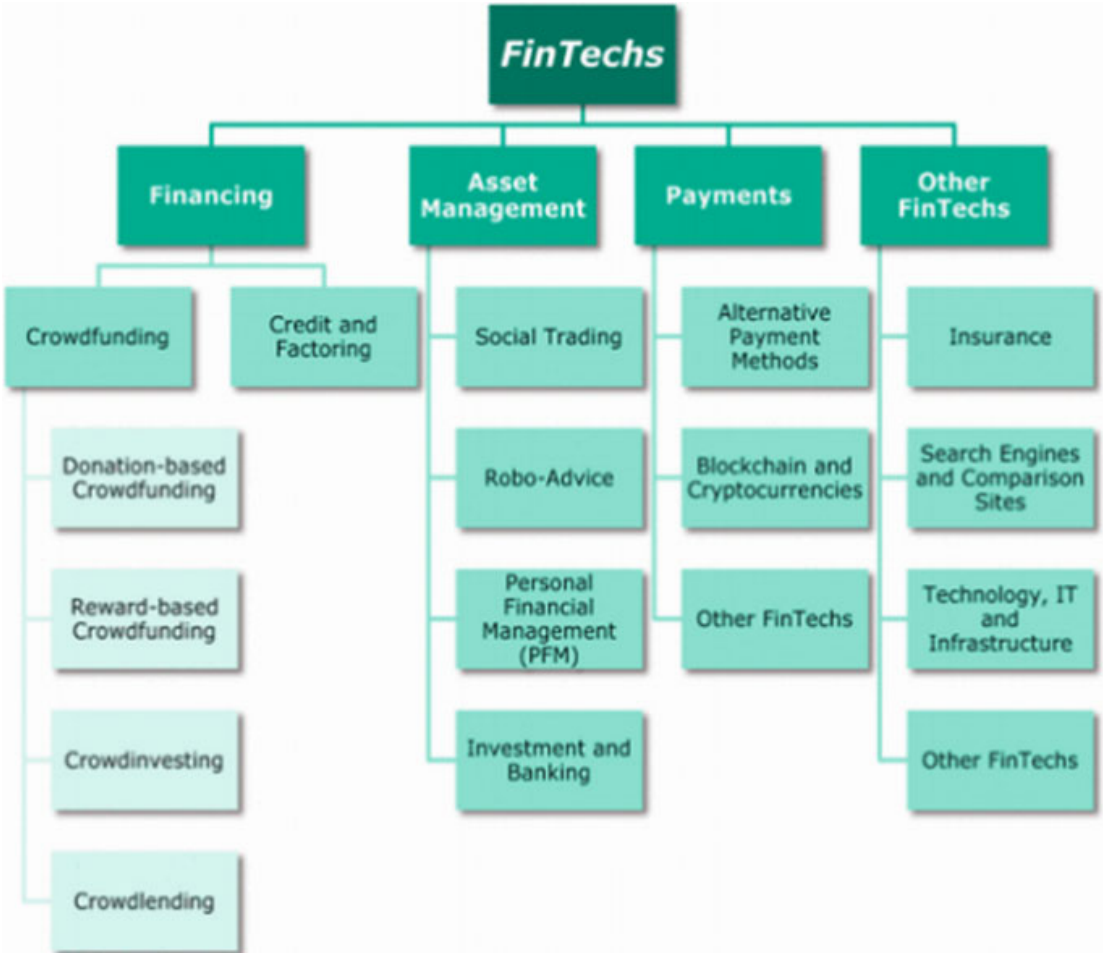


Fig. 5 Segments and elements of FinTech

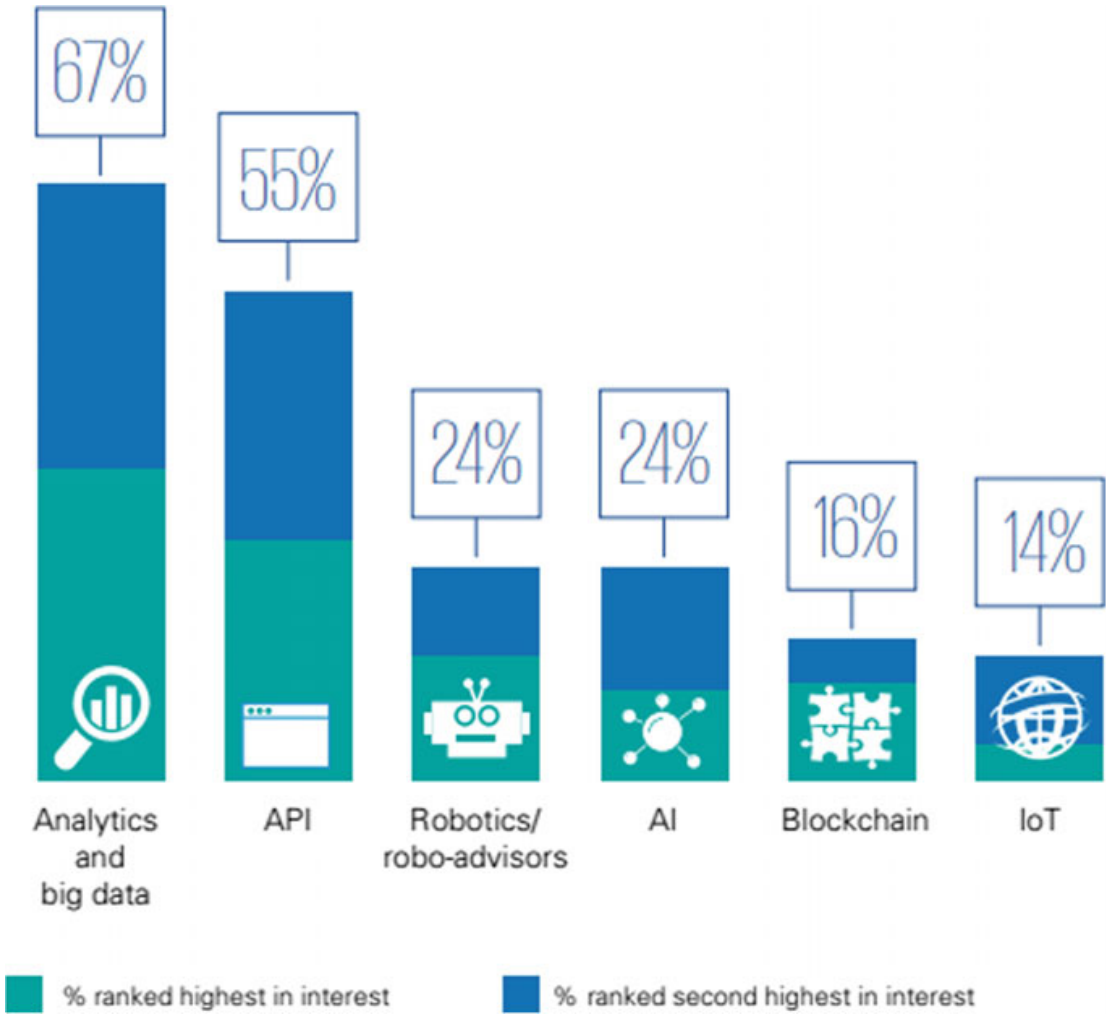


Fig. 6 Emerging FinTech in next three years in the field

mishaps to giving security after the occasion as a case. 36% resource administration organizations additionally positioned man-made consciousness as a best of advancements of intrigue. Safety net providers have more grounded enthusiasm for advances identified with IoT at that point banks or resource administration organizations.

In the course of recent months, the square chain has essential measurement of consideration. It was anticipating the high positioned organizations in the benefit administration. But, it was just 35% positioned zone intrigue.

7 Analysis and Discussion

For this part of work all information and data are collected from various research articles and magazines, and these data were discussed and investigated in this section. We divided the analyses into two categories: the first part is the status of

FinTech block chain and crypto currencies. EU, USA and India have how alternative payment methods and investment are to be demonstrated. Banking is providing positive characteristic as well as negative characteristic using FinTech global financial service sector elements [11].

7.1 FinTech Current Status and Positive Characteristics

FinTech is an imperative piece of open monetary administration segment. It gives information on monetary and keeping money of conventional individuals. The prospect getting diverse sorts of money related administrations. The fiasco and the wretchedness in all the three locales, for instance, the EU, USA and India, did negative effects and impacts on the advancement and interruption in the monetary administration sector [12].

The new FinTech companies are using block chain innovation that has given the opportunity and benefits for more money transaction related administrations. The term digital revolution was reported by Accenture in 2015.

According to Table 2, the Luxembourg is in the top position in digitalization index of 1.00 and USA has 0.92 in digitalization index, which means this is closing to reach the full potential value, and EU also has the average range digitalization value of 0.62, which means the EU member countries are positioned with different values. But India is lagging in digitalization world as it has the value 0.29 and positioned 83rd in the digitalization index.

The FinTech companies bid these services at a lower level price compared to traditional financial service sectors because fully automated technology accessing the operation and process work completely or partially. The motivation or main concept of FinTech service is to reduce the human errors and increase banking business transaction and accessing. In this system automatically applicable by India. India will reached the highest digitalization index ranking position when compared to EU and USA ranking level.

7.2 Fintech Current Status and Negative Characteristics

They are many positive characteristics identified by the people, which are block chain and crypto currencies, an alternative payment system and FinTech technology and banking solutions. Nevertheless, the treats related to the FinTech essentials are really negative effects in the FinTech financial service sectors. The negative elements are due to affecting the FinTech operations and that related process failed incompletely. That's why India is not able to develop it, while the EU and USA quickly increased the work regularly.

Table 2 BBVA digitalization index 2015 [13]

S. No.	Country	Index	S. No.	Country	Index
1	Luxembourg	1.00	44	Kazakhstan	0.47
2	United Kingdom	0.97	45	South Africa	0.47
3	Hong Kong SAR	0.95	46	Slovakia	0.46
4	United States	0.92	47	Mauritius	0.46
5	Netherlands	0.90	48	Colombia	0.45
6	Japan	0.88	49	Russian Federation	0.45
7	Singapore	0.87	50	Italy	0.44
8	Norway	0.86	51	Azerbaijan	0.44
9	Finland	0.85	52	Poland	0.43
10	Sweden	0.84	53	Romania	0.43
11	Switzerland	0.82	54	Croatia	0.43
12	Iceland	0.82	55	Montenegro	0.42
13	Canada	0.81	56	Kuwait	0.41
14	New Zeeland	0.80	57	Mexico	0.41
15	Australia	0.79	58	Greece	0.40
16	Germany	0.78	59	Armenia	0.40
17	Denmark	0.77	60	Georgia	0.40
18	Korea, Rep.	0.76	61	Panama	0.40
19	Estonia	0.76	62	Macedonia FYR	0.39
20	France	0.76	63	China	0.38
21	Austria	0.73	64	Thailand	0.38
22	United Arab Emirates	0.71	65	Morocco	0.37
23	Belgium	0.69	66	Philippines	0.35
24	Ireland	0.68	67	Sri Lanka	0.34
25	Island	0.68	68	Egypt	0.33
26	Bahrain	0.65	69	Indonesia	0.33
27	Lithuania	0.65	70	Bulgaria	0.33
28	Maita	0.64	71	Moldova	0.33
29	Malaysia	0.63	72	Tunisia	0.33
30	Spain	0.62	73	Argentina	0.32
31	Qatar	0.61	74	Kenya	0.32
32	Saudi Arabia	0.59	75	Peru	0.32
33	Portugal	0.59	76	EI Salvador	0.31
34	Chile	0.58	77	Serbia	0.31
35	Latvia	0.55	78	Dominican Rep	0.31
36	Czech Republic	0.52	79	Vietnam	0.31
37	Oman	0.51	80	Honduras	0.30
38	Turkey	0.50	81	India	0.29
39	Costa Rica	0.49	82	Albania	0.28
40	Jordan	0.48	83	Albania	0.24

(continued)

Table 2 (continued)

S. No.	Country	Index	S. No.	Country	Index
41	Cyprus	0.48	84	Senegal	0.24
42	Hungary	0.48	85	Guatemala	0.22
43	Uruguay	0.47	86	Ukraine	0.21

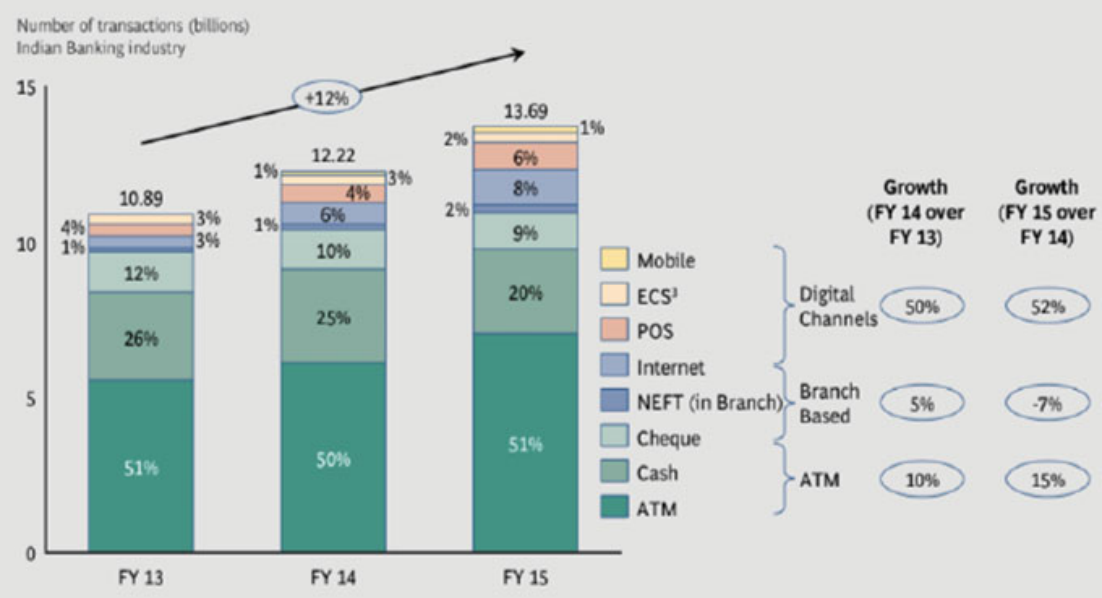


Fig. 7 Digitalization transaction of India in 2013–2015

According to the data, Fig. 7 displays the digitalization transaction in India's growth; however, the Billon Eco System (BCG) research in India that surveyed the ATM mobile transaction and withdrawal transaction at ATM, NEFT-based transaction in mobile, mobile internet and traditional transaction of cash and cheque and all our day-to-day our life usage showed 12% increase in the year 2013–2014 and also in 2014–2015; so, year by year as we increase FinTech operations, banking financial services also increased [14].

This is very a good example of block chain and crypto currencies method. The customer will need to safety for payment transaction because the hackers follow to your regular work. This cyber security and data privacy also comes from the USA and EU. The regularly EU, USA and India using the FinTech methods and block chain Methods for money transaction [15].

8 Conclusions

In this study the overall functions of financial advanced technologies, challenges of banking and financial industries, cyber security of banking while connecting IoT through the internet to store the all information's on cloud and benefits of the FinTech while connecting IoT device have been described. In addition, the use cases of FinTech on IoT in digital society in futures are mentioned in detailed. All banks, financial industry, insurance companies are moved to automated technology that reduces the workloads and easy to mingle the customer activities. From authors' point of view, the FinTech embraced with IoT in future will boom the society into the next generation.

References

1. Rega, F.G.: The bank of the future, the future of banking—an empirical analysis of European banks. SSRN Electron. J. (2017). https://doi.org/10.2139/ssrn.3071742
2. Infosys Limited: IOT-Enabled Banking Services. https://www.infosys.com/industries/financial-services/white-papers/Documents/IoT-enabled-banking.pdf
3. Nakashima, T.: Creating credit by making use of mobility with FinTech and IoT. IATSS Res. **42**(2), 61–66 (2018)
4. Us, A., Us, C., Account, M.Y.: Cybersecurity in the Banking and Financial Services Sector—IoT Threats, Potential Solutions and Blockchain (2018). https://www.stoodnt.com/blog/cybersecurity-in-banking-financial-services/
5. Cuipa, E., Ramani, S., Shetty, N., Smart, C.: Financing the Internet of Things: An Early Glimpse of the Potential M-RCBG Associate Working Paper Series, No. 85 (2018)
6. Schubert, J.: 4 Top Challenges Facing The Banking Industry Right Now (2015). https://www.digitalistmag.com/industries/banking/2015/08/27/4-top-challenges-facing-banking-industry-right-now-03352186
7. Foes, O.R.: Fintech and banks, regulation, and the real sector. Eur. Econ. Banks Regul. Real Sect. **2**, 1–162 (2017)
8. S. in Join: 15 Top Fintech Companies in India 2018 (2018). https://www.whizsky.com/2018/04/20-top-fintech-companies-in-india-2018/
9. Schulte, P., Liu, G., Alerts, E., Journals, I.I.: FinTech is merging with IoT and AI to challenge banks: how entrenched interests can prepare. Inst. Investig. J. **20**(3), 41–57 (2018)
10. Story, C., Talk, I., Study, C.: Accelerating your journey to #truly digital banking (2017). file:///C:/Users/Nathiya/Downloads/FinacleConnect-Apr-Sep-2017-Vol09-Issue34.pdf
11. Pejkovska, M.: Potential negative effects of FinTech on the financial services sector Examples from the European Union, India and the United States of America (2018). https://www.theseus.fi/handle/10024/148416
12. Fenwick, M., Mccahery, J.A., Vermeulen, E.P.M., Fenwick, M., Mccahery, J.A.: FinTech and the Financing of Entrepreneurs : From Crowdfunding to Marketplace Lending FinTech and the Financing of Entrepreneurs (2017). https://ecgi.global/sites/default/files/working_papers/documents/fenwick-mccahery-vermeulen.pdf
13. Accenture 2015: The Future of FinTech and Banking (2015). https://www.accenture.com/t20150707T195228Z__w__/us-en/_acnmedia/Accenture/Conversion-Assets/DotCom/Documents/Global/PDF/Dualpub_11/Accenture-Future-Fintech-Banking.pdf#zoom=50

14. Shah, A., Vibha, K., Prateek, R., Abhishek, A.: Digital Payments 2020: The making of a $500 billion ecosystem in India. [ebook] Boston Consulting Group and Google. http://image-src.bcg.com/BCG-GoogleDigitalPayments2020-July2016_tcm21-39221.pdf
15. Pulse, K., Share, F.R.: Global FinTech market sees quiet Q1'17 as M&A slows, VC funding holds steady: KPMG Pulse of FinTech Report (2018). https://home.kpmg.com/in/en/home/media/press-releases/2017/05/kpmg-pulse-fintech-report.html

Big Data Management

GRNN++: A Parallel and Distributed Version of GRNN Under Apache Spark for Big Data Regression

Sk. Kamaruddin and Vadlamani Ravi

Abstract Among the neural network architectures for prediction, multi-layer perceptron (MLP), radial basis function (RBF), wavelet neural network (WNN), general regression neural network (GRNN), and group method of data handling (GMDH) are popular. Out of these architectures, GRNN is preferable because it involves single-pass learning and produces reasonably good results. Although GRNN involves single-pass learning, it cannot handle big datasets because a pattern layer is required to store all the cluster centers after clustering all the samples. Therefore, this paper proposes a hybrid architecture, GRNN++, which makes GRNN scalable for big data by invoking a parallel distributed version of K-means++, namely, K-means‖, in the pattern layer of GRNN. The whole architecture is implemented in the distributed parallel computational architecture of Apache Spark with HDFS. The performance of the GRNN++ was measured on gas sensor dataset which has 613 MB of data under a ten-fold cross-validation setup. The proposed GRNN++ produces very low mean squared error (MSE). It is worthwhile to mention that the primary motivation of this article is to present a distributed and parallel version of the traditional GRNN.

Keywords Apache Spark · Big data regression · GRNN · HDFS · K-means++ · K-means‖

Sk. Kamaruddin · V. Ravi (✉)
Centre of Excellence in Analytics, Institute for Development and Research in Banking Technology, Castle Hills Road No. 1, Masab Tank, Hyderabad 500057, India
e-mail: padmarav@gmail.com

Sk. Kamaruddin
e-mail: skkamaruddin@gmail.com

Sk. Kamaruddin
SCIS, University of Hyderabad, Hyderabad 500046, India

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_16

1 Introduction

The term prediction is used for regression as well as classification. When the target variable is discrete the associated task belongs to classification and when the target variable is continuous the associated task is a regression. The regression or forecasting model takes the past data to predict the target variable. There are different neural network architectures used for regression, namely, MLP, RBF, WNN, GRNN, and GMDH.

MLP has been used for regression in several research works, namely, Kusakunniran et al. [1] presented a gait recognition system based on motion regression using MLP; Agirre-Basurko et al. [2] suggested MLP to forecast gas levels; Gaudart et al. [3] performed a regression for epidemiological data using MLP.

Similarly, RBF also has been used for regression, namely, Mignon and Jurie [4] proposed a face reconstruction algorithm, based on RBF regression in eigenspace; Hannan et al. [5] presented heart disease diagnosis using RBF; Taki et al. [6] used RBF for energy consumption prediction for wheat production.

WNN is used for regression models for forecasting reservoir inflow [7]. Vinaykumar et al. [8] implemented WNN for the estimation of cost of software development; Chauhan et al. [9] used DE-trained WNN for bankruptcy prediction in banks. Rajkiran and Ravi [10] used WNN for predicting reliability of software.

In addition, GMDH is used for regression, namely, Astakhov and Galitsky [11] presented tool life prediction in gundrilling; Elattar et al. [12] used GMDH for short-term load forecasting; Srinivasan [13] used GMDH for energy demand prediction; Ravisankar and Ravi [14] implemented GMDH for bankruptcy prediction in banks; Mohanty et al. [15] used GMDH for prediction of software reliability; Reddy and Ravi [16] proposed kernel GMDH for regression.

The ANN with its vast variants of architectures have been implemented in different application domains, namely, wireless networks [17], robot manipulator control [18], wind energy systems [19], cancer prediction in healthcare [20], crop production [21], and business [22].

Among these different architectural variants of ANN, the GRNN architecture is preferable as it uses a single pass for learning, and being non-parametric estimation produces a good result. In this article, we have proposed a hybrid version of GRNN for prediction. The motivation and contribution for the proposed method are introduced in Sect. 2.

The paper is structured in the following way: Motivation and contribution are presented in Sect. 2. A literature review is presented in Sect. 3. The proposed methodology is described in Sect. 4. The system setup for conducting experiments is discussed in Sect. 5, and the information about the dataset used in the experiment is described in Sect. 6. The results and discussion are presented in Sect. 7. Section 8 concludes the paper with future directions.

2 Motivation and Contribution

We have presented in the introduction section the different ANN architectures for regression. Among them, GRNN is an optimal choice for exploration being a one-pass trainer and non-parametric estimator. According to Specht [23], if a large dataset is involved, we should have a process of clustering to have neurons in the pattern layer. But, clustering of massive data is not efficient enough through any standard clustering approach. Even if we opt for the parallel version of K-means, it is not efficient enough with respect to optimal clustering. Again, we find there is no published work involving prediction of large-scale dataset using the GRNN. These shortcomings and adversaries are the strong cause of the motivation for the current research work.

In this proposed work GRNN is used for prediction of the large-scale dataset. Here, we have used a clustering approach using K-means‖ [24] which is a parallel version of K-means++ [25] to reduce the pattern layer neurons. It is a variant of K-means which has optimal initial seed selection approach. The K-means‖ is embedded into the architecture of GRNN for carrying out clustering before the pattern layer and implemented in a parallel distributed computing framework of Apache Spark which is henceforth called as GRNN++. The GRNN++ has shown very high accuracy in the form of mean squared error. The hybrid architecture has proved the implementation of the best features of two worlds, that is, of clustering and regression.

3 Literature Review

There are several variants of K-means proposed in the literature by the research community. Also, a large effort has been dedicated toward parallelization of K-means algorithm. Zhao et al. [26] have propounded a MapReduce framework for parallel K-means. Liao et al. [27] have presented an improved version of parallel K-means using MapReduce and a method for selecting the initial centroids. The K-means++ [25] is optimized by the process of selection of initial cluster centers, but it has a major drawback as being sequential. This limitation restricts its application to the massive dataset. It has to perform k-passes over the dataset to figure out k initial cluster centers. This drawback is addressed by K-means‖ [24]. Apart from K-means, evolving clustering method is another clustering approach which has been parallelized by Kamaruddin and Ravi [28].

The regression has always been the fascinating area of machine learning, and different approaches have been presented by the researchers. One of the efficient neural network architecture for regression is GRNN [23] which has the advantage of non-parametric estimation along with one-pass of training. These advantages enable the GRNN to be faster with better accuracy. There are several contributions by the researchers in different applications areas using GRNN, namely, Leung et al.

[29] have forecasted exchange rate using GRNN; Kayaer and Yildirim [30] implemented GRNN for medical diagnosis of diabetes; Li et al. [31] implemented GRNN for image quality assessment; Li et al. [32] implemented GRNN with fruit fly optimization algorithm for power load distribution prediction.

But the primary disadvantage of GRNN is that all the training data should be stored in the pattern layer, which requires more computational time. Thus, results in an inefficient architecture for the large-scale dataset.

Some researchers have presented hybrid architectures, and those are generally found to be more efficient. Ravi and Krishna [33] implemented generalized regression with auto-associative neural network (GRAANN) for imputing missing values. Similarly, the hybrid architecture of auto-associative extreme learning machine and multiple linear regression proposed by Tejasviram et al. [34] has shown better prediction capability. Kamaruddin and Ravi [35] presented a hybridization of particle swarm optimization and AANN for one-class classification and had applied it to credit card fraud detection which had demonstrated good performance.

From the study of the related literature, we found that GRNN has not been implemented in a parallel distributed environment and is not able to handle large-scale dataset. So the current research work addresses the issue.

4 Introduction to Proposed Methodology: GRNN++

GRNN is not able to handle large-sized dataset without efficient clustering. It will raise memory issues and increase the execution time. These drawbacks are due to the incapability of the standard clustering approach to handle large-sized dataset. In the proposed GRNN++ method this drawback has been addressed. The GRNN++ is implemented with Apache Spark and HDFS. It has two components, namely, (i) K-means‖ clustering, and (ii) GRNN for prediction. The two components are introduced in the following two subsections.

4.1 K-Means++: An Overview

The K-means clustering algorithm has an objective of minimal intra-cluster variance; that is, the sum of the squared distance of the points belonging to a cluster from their cluster center should be minimum. The K-means chooses random initial "K" cluster centers and updates the cluster centers with iterations until there is no further change in the cluster center, and thus finds reasonable solutions quickly. However, it suffers from at least two major drawbacks: (1) the algorithm has the worst-case execution time which is super-polynomial of input size; (2) the formed clusters may not be optimal with respect to the objective function.

The K-means++ [25] has addressed the deficiency of optimal clustering by the introduction of a method for cluster center initialization before execution of the iterations of K-means algorithm. The K-means++ clustering algorithm using an innovative method for initial cluster center selection or seed selection was proposed by Arthur and Vassilvitskii [25]. The algorithm has the following steps:

K-Means++ algorithm:

Let D be the dataset of p points in R^d and $x \in D$. Let $d(x)$ be the smallest distance from any given data point to its nearest cluster center which has been already selected.

1. Select first center C_1 uniformly at random from D.
2. For each data point $x \in D$, compute $d(x)$ and take a new center C_i, with probability $\frac{d(x)^2}{\sum_{x \in D} d(x)^2}$.
3. Repeat Step 2 until k centers have been selected altogether.
4. Proceed with the iterations of the K-means.

4.2 K-Means‖: An Overview

The K-means‖ is the parallel implementation of K-means++. Say, there are n number of samples, and k initial cluster centers are needed. The K-means++ performs k number of passes to sample one initial cluster center in each pass. But, K-means‖ samples first center uniformly at random. Then, the next centers will be chosen non-uniformly with a given probability which is stochastically biased by the already chosen centers. This happens in a parallel way across all partitions. Thus, the so-obtained $O(k \log n)$ points are finally re-clustered into k initial centers for the standard K-means iterations.

4.3 GRNN: An Overview

GRNN [23] is a feed forward neural network having roots in statistics. GRNN represents an advanced architecture in the neural networks that implements non-parametric regression with one-pass training.

The topology of GRNN is depicted in Fig. 1 and involves four layers of neurons, namely, input, pattern, summation, and the output.

The pattern layer comprises training neurons. The test sample is fed to the input layer. The distance, d_i, between the training samples presents as a neuron in the pattern layer, and the data point from the test set used for regression is used to figure out how well each training neuron in pattern layer can represent the feature space of the test sample, X. This probability density is calculated with Gaussian activation function. Thus, the summation of the product of target value and the result of activation function

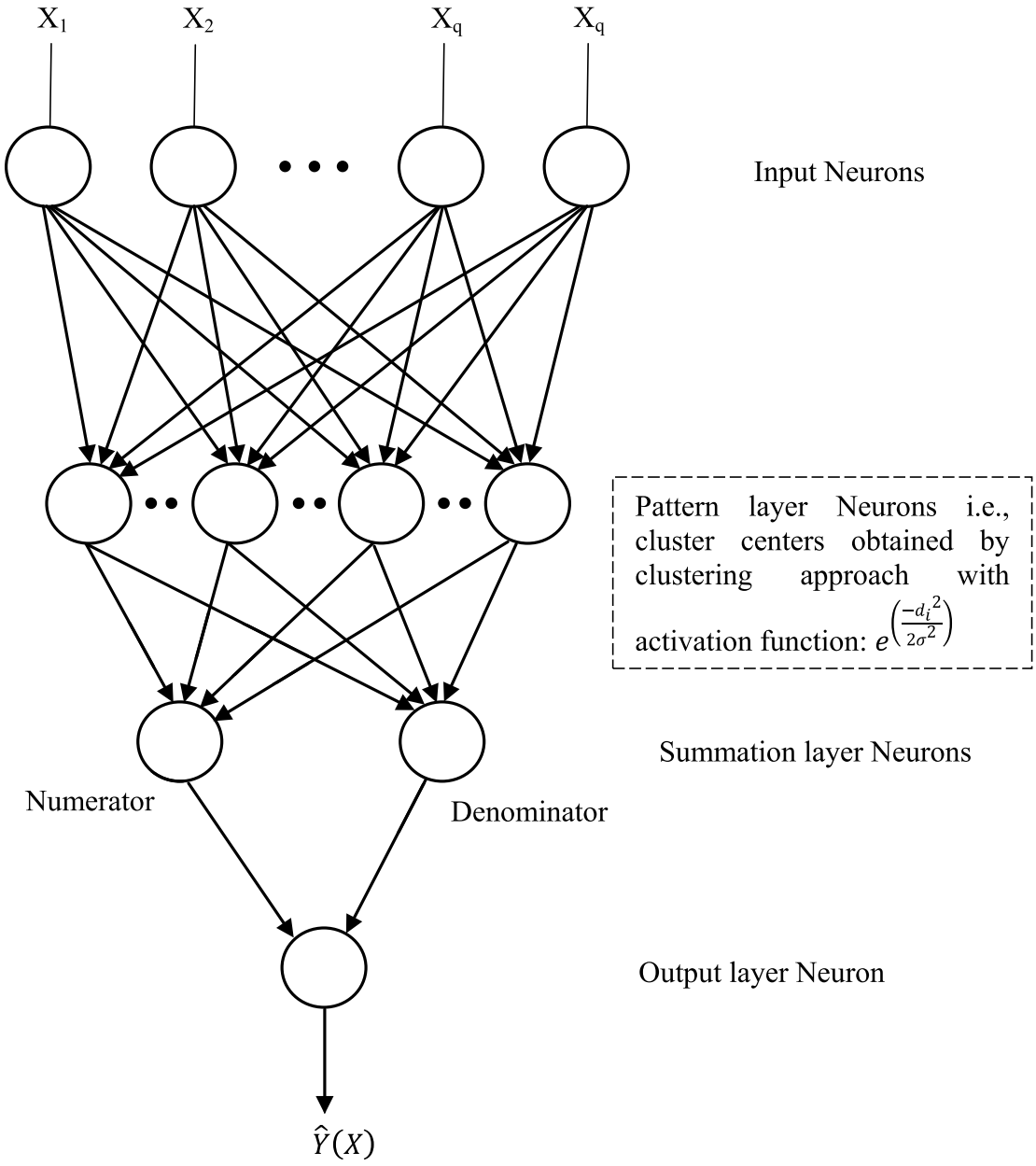


Fig. 1 Architecture of GRNN

for each neuron that is, $\sum_{i=1}^{n} Y_i * e^{\left(\frac{-d_i^2}{2\sigma^2}\right)}$ forms the numerator term in the summation layer and the summation of the term $e^{\left(\frac{-d_i^2}{2\sigma^2}\right)}$ that is, $\sum_{i=1}^{n} e^{\left(\frac{-d_i^2}{2\sigma^2}\right)}$ forms the denominator term in the summation layer. Then, in the output layer, the prediction is calculated as $\widehat{Y}(X) = \frac{\sum_{i=1}^{n} Y_i * e^{\left(\frac{-d_i^2}{2\sigma^2}\right)}}{\sum_{i=1}^{n} e^{\left(\frac{-d_i^2}{2\sigma^2}\right)}}$. Here $d_i^2 = (X - X_i)^T * (X - X_i)$. X is a test sample and X_i is the neuron present in the pattern layer. For big datasets, the equations involved in

summation and output layer change to accommodate the information about the cluster centers. For further details please see the training algorithm presented in Sect. 4.4.

4.4 GRNN++: The Proposed Method

The training data is clustered with K-means‖. Thus, the cluster centers represent the data samples belonging to the cluster most effectively. Then the Dunn-like Index (a modified Dunn Index), a cluster validity index, is used for measuring the cluster validity.

The Dunn Index, proposed by Dunn [36], presents the approach for measuring the cluster quality so that the best clusters were selected based on the high compactness of clusters and high separation among them, that is, compact and well-separated clusters. The compactness with good separation is found by the ratio of minimum intra-cluster distance to the maximum inter-cluster distance. The Dunn Index can be presented as follows:

$$Dunn\,Index(D) = \min_{1 \le i \le M} \left\{ \min_{i+1 \le j \le M} \left(\frac{dist(c_i, c_j)}{\max_{1 \le l \le M} (diam(c_l))} \right) \right\}. \tag{1}$$

where M is the total number of clusters under consideration, $dist(c_i, c_j)$ is the inter-cluster distance and $diam(c_l)$ is the maximum intra-cluster distance or the largest diameter among all the clusters.

As the Dunn Index performs a calculation of maximum intra-cluster distance, that is, the diameter of a cluster under consideration, it is not suitable for large-scale dataset as it will involve large computational overhead. Also, it is affected by noisy data and outliers. So, we have used a Dunn-like Index [37] for cluster validity measurement. In this approach, the diameter, that is, the denominator of (1) is calculated by finding the radius using the distance from each sample present in the cluster to its cluster center, that is, $d\left(x, \overline{c_j}\right)$, where $x \in c_j$ cluster and $\overline{c_j}$ is its center. This Dunn-like Index is not affected by noise or outlier present in the data. The Dunn-like Index can be presented as:

$$Dunn-like\,Index(DL) = \min_{1 \le i \le M} \left\{ \min_{i+1 \le j \le M} \left(\frac{dist\left(c_i, c_j\right)}{\max_{1 \le l \le M} \left(2 * \frac{1}{|c_l|} \sum_{x \in c_l} d(x, \overline{c_l}) \right)} \right) \right\}. \tag{2}$$

The cluster centers form the clusters generated by the K-means‖ algorithm and validated by Dunn-like Index. They now represent the pattern nodes of GRNN++.

Then, each test sample is passed on to the neurons present in the pattern layer with a Gaussian probability density function resulting in higher value if the neuron has a strong similarity to the test sample. Now, the summation layer represents the numerator, denominator and the output layer presents the prediction as described in the overview of the GRNN section (refer to Sect. 4.3).

Training Algorithm for the proposed GRNN++

Step 1: Perform data normalization with the range [0,1].
Step 2: Divide the data to perform 10-Fold cross-validation.
Step 3: Take the training set and cluster them with K-Means|| with the k = 2 to 10.
Step 4: Compute Dunn-like Index by varying the value of k and find the k which produces optimal clustering.
Step 5: Consider the cluster centers represented by the optimal k as the neurons in the pattern layer of GRNN++ and pass the training or test data, as the case may be, to the pattern layer nodes with a Gaussian transfer function.
Step 6: The product of A_i, i.e., the sum of the output values of the samples present in the i^{th} cluster and the result of activation function for i^{th} neuron in the pattern layer $e^{\left(\frac{-d_i^2}{2\sigma^2}\right)}$ i.e., $\sum_{i=1}^{n} A_i * e^{\left(\frac{-d_i^2}{2\sigma^2}\right)}$ forms the numerator term in the summation layer (say *Num*).
Step 7: The product of B_i, i.e., count of the samples present in the i^{th} cluster and the result of activation function for i^{th} neuron in the pattern layer $e^{\left(\frac{-d_i^2}{2\sigma^2}\right)}$ i.e., $\sum_{i=1}^{n} B_i * e^{\left(\frac{-d_i^2}{2\sigma^2}\right)}$ forms the denominator term in the summation layer (say *Denom*).
Step 8: In the output layer, the prediction is computed as $\hat{Y}(x) =$ *Num / Denom*.
Step 9: The accuracy of prediction is measured by computing MSE.
Step 10: Steps 3 to 9 are repeated for the 10 folds of data.
Step 11: Average MSE for 10-FCV is computed and reported.

The architecture of GRNN++ is the same as the traditional GRNN except for the presence of the clusters (obtained through K-means||) occupying the pattern layer in place of the patterns themselves (see Fig. 1). The third and fourth layer of GRNN++ are identical to that of GRNN.

5 Experimental Setup

The system configuration for carrying out experimental work consists of standalone Spark cluster 2.2.0 as the computational framework on top of HDFS which is the distributed storage system. The program development environment used is Apache

Table 1 System configuration description

Central processing unit	Intel® Core™ i7-6700 CPU @ 3.40 GHz
No. of cores	4 Physical cores or 8 logical cores
Primary memory	32 GB
OS	Ubuntu 16.04 LTS

Table 2 Details of resource allocation in the cluster and coding environment

Allocated resource details	Node type	
	Master	Worker
Memory allocated for driver process	14 GB	–
Count of worker daemons	2	4
Memory allocated to worker daemon	7 GB	7 GB
Number of executors per node	2	4
Memory allocated to each executor	7 GB	7 GB
Cores allocated per executor	2	2
Memory allocated to all executors	14 GB	28 GB
Total memory utilized (out of 32 GB)	28 GB	28 GB
Framework for computation	Apache Spark 2.2.0	
Distributed storage system	HDFS (Hadoop 2.7.3)	
Coding interface	Apache Zeppelin 0.7.3	
Programming language	Scala 2.11.8	

Zeppelin 0.7.3 with Scala 2.11.8 as the coding environment. The Spark cluster consists of a system as a master node and seven systems as worker nodes. The driver program resides in the master node along with two instances of worker daemons. Each of the worker nodes executes four instances of worker daemons. All the seven systems have a similar configuration (refer to Table 1). In a standalone cluster system, each worker daemon has one executor. The details of resources allocated to worker daemon or executor, driver process, and programming environment details are presented in Table 2.

The GRNN++ was implemented with the above cluster configuration and programming environment. We found the best execution time by adjustment of the amount of memory allocation to the executors along with the optimum data partitions.

Table 3 Attribute details of ethylene–CO gas sensor array dataset

Attribute name	Attribute details
Time (s)	Time measured in seconds for the sensor readings
CO conc. (ppm)	The concentration of CO measured in ppm
Ethylene conc. (ppm)	The concentration of ethylene measured in ppm
Sensor readings (16 channels)	Readings from 16 chemical sensors forming the data comprising 16 columns

6 Dataset Description

The current work has analyzed the data from an array of gas sensors involving different gas mixtures. The dataset [38] was collected from UCI ML repository [39]. The dataset contains the readings gathered from 16 sensors exposed to varying concentrations of gas mixtures. The original dataset contains readings of two different mixtures of gas in air, namely, ethylene with methane, and ethylene with CO. We have analyzed only the ethylene and CO mixture in air dataset for prediction of ethylene and CO concentration in air. The sensor readings were the data from 16-sensor array signals gathered uninterruptedly for about 12 h.

The ethylene and CO mixture in air dataset, which is analyzed in the current work, presents 19 features. Here, the first feature represents the time of sensor reading; the second feature represents CO concentration in ppm; and the third feature represents the ethylene concentration in ppm. The next 16 features represent the sensor readings from 16 chemical sensors (see Table 3). The dataset contains 4.2 million samples. The dataset is of 643 MB.

For our regression work, we have dropped the first feature being the time series values. Out of the second and third feature, one is selected as the dependent variable. We have selected the features from the fourth to the last as independent variables for our experimentation.

7 Results and Discussion

The research work conducted comprises the analysis of the dataset mentioned above with the proposed approach. The results of the carried out work cannot be compared with any other technique as we could not find any work carried out in the same domain and using such hybrid architectural techniques.

The experiment was carried out with ten-fold cross-validation (10-FCV) of the min–max normalized dataset. The K-means‖ was carried out with k value ranging from 2 to 10 and then the Dunn-like Index was taken into account for cluster validity for finding out the optimal cluster centers. Then, with the selected cluster centers GRNN++ was carried out with sigma (σ) value ranging from 0.001 to 1.0

Table 4 Mean MSE for 10-FCV

	Ethylene concentration prediction from ethylene to CO dataset	CO concentration prediction from ethylene to CO dataset
Gaussian activation function	0.075	0.076

with an increment of 0.001. The error is calculated as MSE. Table 4 presents the average MSE of 10-FCV.

The activation function used in GRNN++ is the Gaussian function which captures the features of a test data and correspondingly presents the prediction value.

8 Conclusion and Future Directions

The GRNN++, a parallel and distributed version of GRNN, is proposed in Apache Spark, where HDFS takes care of the distributed data storage. The GRNN++ is implemented in Scala programming language. The GRNN++ could achieve big data regression in a single pass and by yielding a very low mean squared error of 10-FCV. The hallmark of the proposed method is the invocation of the parallel distributed K-means++ aka K-means‖ into the architecture of the traditional GRNN. GRNN being the non-parametric estimation provides better accuracy.

The innovation of the proposed method resulted in a drastic reduction in the count of neurons in the pattern layer of GRNN. This feature has made the system overcome the major shortcoming of GRNN which is computational overhead.

The effectiveness of GRNN++ is tested on a static dataset, but it can be modified to suit online streaming data mining by invoking parallel evolving clustering method (PECM) [28] in place of the K-means‖. Thus, it can be extended to handle real-time or quasi-real-time prediction from streaming data. Further, the logistic activation function can also be considered in place of the Gaussian in the pattern layer.

References

1. Kusakunniran, W., Wu, Q., Zhang, J., Li, H.: Multi-view gait recognition based on motion regression using multilayer perceptron. In: 2010 20th International Conference on Pattern Recognition, pp 2186–2189. IEEE, Istanbul (2010)
2. Agirre-Basurko, E., Ibarra-Berastegi, G., Madariaga, I.: Regression and multilayer perceptron-based models to forecast hourly O_3 and NO_2 levels in the Bilbao area. Environ. Model Softw. **21**, 430–446 (2006)

3. Gaudart, J., Giusiano, B., Huiart, L.: Comparison of the performance of multi-layer perceptron and linear regression for epidemiological data. Comput. Stat. Data Anal. **44**, 547–570 (2004)
4. Mignon, A., Jurie, F.: Reconstructing faces from their signatures using RBF regression. In: Procedings of the British Machine Vision Conference 2013, pp 103.1–103.11. British Machine Vision Association, Bristol (2013)
5. Hannan, S.A., Manza, R.R., Ramteke, R.J.: Generalized regression neural network and radial basis function for heart disease diagnosis. Int. J. Comput. Appl. **7**, 7–13 (2010)
6. Taki, M., Rohani, A., Soheili-Fard, F., Abdeshahi, A.: Assessment of energy consumption and modeling of output energy for wheat production by neural network (MLP and RBF) and Gaussian process regression (GPR) models. J. Clean. Prod. **172**, 3028–3041 (2018)
7. Budu, K.: Comparison of wavelet-based ANN and regression models for reservoir inflow forecasting. J. Hydrol. Eng. **19**, 1385–1400 (2014)
8. Vinaykumar, K., Ravi, V., Carr, M., Rajkiran, N.: Software development cost estimation using wavelet neural networks. J. Syst. Softw. **81**, 1853–1867 (2008)
9. Chauhan, N., Ravi, V., Karthik Chandra, D.: Differential evolution trained wavelet neural networks: Application to bankruptcy prediction in banks. Expert Syst. Appl. **36**, 7659–7665 (2009)
10. Rajkiran, N., Ravi, V.: Software reliability prediction using wavelet neural networks. In: International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007), pp 195–199. IEEE, Sivakasi (2007)
11. Astakhov, V.P., Galitsky, V.V.: Tool life testing in gundrilling: an application of the group method of data handling (GMDH). Int. J. Mach. Tools Manuf **45**, 509–517 (2005)
12. Elattar, E.E., Goulermas, J.Y., Wu, Q.H.: Generalized locally weighted GMDH for short term load forecasting. IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.) **42**, 345–356 (2012)
13. Srinivasan, D.: Energy demand prediction using GMDH networks. Neurocomputing **72**, 625–629 (2008)
14. Ravisankar, P., Ravi, V.: Financial distress prediction in banks using group method of data handling neural network, counter propagation neural network and fuzzy ARTMAP. Knowl. Based Syst. **23**, 823–831 (2010)
15. Mohanty, R., Ravi, V., Patra, M.R.: Software reliability prediction using group method of data handling. In: Sakai, H., Chakraborty, M.K., Hassanien, A.E., Ślęzak, D., Zhu, W. (eds.) Rough Sets, Fuzzy Sets, Data Mining and Granular Computing. RSFDGrC 2009, pp 344–351. Springer, Berlin (2009)
16. Reddy, K.N., Ravi, V.: Kernel group method of data handling: application to regression problems. In: Panigrahi, B.K., Das, S., Suganthan, P.N., Nanda, P.K. (eds.) Swarm, Evolutionary, and Memetic Computing. SEMCCO 2012, pp 74–81. Springer, Berlin (2012)
17. Ahad, N., Qadir, J., Ahsan, N.: Neural networks in wireless networks: techniques, applications and guidelines. J. Netw. Comput. Appl. **68**, 1–27 (2016)
18. Jin, L., Li, S., Yu, J., He, J.: Robot manipulator control using neural networks: A survey. Neurocomputing **285**, 23–34 (2018)
19. Marugán, A.P., Márquez, F.P.G., Perez, J.M.P., Ruiz-Hernández, D.: A survey of artificial neural network in wind energy systems. Appl. Energy **228**, 1822–1836 (2018)
20. Agrawal, S., Agrawal, J.: Neural network techniques for cancer prediction: a survey. Proc. Comput. Sci. **60**, 769–774 (2015)
21. Khoshroo, A., Emrouznejad, A., Ghaffarizadeh, A., Kasraei, M., Omid, M.: Sensitivity analysis of energy inputs in crop production using artificial neural networks. J. Clean. Prod. **197**(Part 1), 992–998 (2018)
22. Tkáč, M., Verner, R.: Artificial neural networks in business: two decades of research. Appl. Soft Comput. **38**, 788–804 (2016)
23. Specht, D.F.: A general regression neural network. IEEE Trans. Neural Netw. **2**, 568–576 (1991)
24. Bahmani, B., Moseley, B., Vattani, A., Kumar, R., Vassilvitskii, S.: Scalable K-means++. Proc. VLDB Endow. **5**, 622–633 (2012)

25. Arthur, D., Vassilvitskii, S.: k-means ++: the advantages of careful seeding. In: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms, pp 1027–1035 (2007)
26. Zhao, W., Ma, H., He, Q.: Parallel K-means clustering based on MapReduce. In: Jaatun, M. G., Zhao, G., Rong, C. (eds.) Cloud Computing, pp. 674–679. Springer, Berlin (2009)
27. Liao, Q., Yang, F., Zhao, J.: An improved parallel K-means clustering algorithm with MapReduce. In: 2013 15th IEEE International Conference on Communication Technology, pp 764–768. IEEE (2013)
28. Kamaruddin, S., Ravi, V., Mayank, P.: Parallel evolving clustering method for big data analytics using apache spark: applications to banking and physics. In: Reddy, P., Sureka, A., Chakravarthy, S., Bhalla, S. (eds.) Lecture Notes in Computer Science. Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics, pp. 278–292. Springer, Cham (2017)
29. Leung, M.T., Chen, A.-S., Daouk, H.: Forecasting exchange rates using general regression neural networks. Comput. Oper. Res. **27**, 1093–1110 (2000)
30. Kayaer, K., Yildirim, T.: Medical diagnosis on Pima Indian diabetes using general regression neural networks. In: Proceedings of the International Conference on Artificial Neural Networks and Neural Information Processing (ICANN/ICONIP), pp 181–184 (2003)
31. Li, C., Bovik, A.C., Wu, X.: Blind image quality assessment using a general regression neural network. IEEE Trans. Neural Netw. **22**, 793–799 (2011)
32. Li, H., Guo, S., Li, C., Sun, J.: A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm. Knowl. Based Syst. **37**, 378–387 (2013)
33. Ravi, V., Krishna, M.: A new online data imputation method based on general regression auto associative neural network. Neurocomputing **138**, 106–113 (2014)
34. Tejasviram, V., Solanki, H., Ravi, V., Kamaruddin, S.: Auto associative extreme learning machine based non-linear principal component regression for big data applications. In: 2015 Tenth International Conference on Digital Information Management (ICDIM), pp 223–228. IEEE, Jeju (2015)
35. Kamaruddin, S., Ravi, V.: Credit card fraud detection using big data analytics: use of PSOAANN based one-class classification. In: Proceedings of the International Conference on Informatics and Analytics—ICIA-16, pp 1–8. ACM Press, Pondicherry (2016)
36. Dunn, J.C.: A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. J. Cybern. **3**, 32–57 (1973)
37. Bezdek, J.C., Pal, N.R.: Some new indices of cluster validity. IEEE Trans. Syst. Man Cybern. B Cybern. **28**, 301–315 (1998)
38. Fonollosa, J., Sheik, S., Huerta, R., Marco, S.: Reservoir computing compensates slow response of chemosensor arrays exposed to fast varying gas concentrations in continuous monitoring. Sensors Actuators B Chem. **215**, 618–629 (2015)
39. Gas sensor array under dynamic gas mixtures Data Set, https://archive.ics.uci.edu/ml/datasets/Gas+sensor+array+under+dynamic+gas+mixtures

An Entropy-Based Technique for Conferences Ranking

Fiaz Majeed and Rana Azhar Ul Haq

Abstract Ranking of conferences is carried out to guide the researchers so as to publish their work in top-level venues. Existing works are needed to be improved in regard to accuracy. In this paper, we have proposed a ranking technique of conference papers named as influence language model (InLM). It uses entropy to calculate score for the purpose of ranking. It identifies level of the paper as well as level of the conference based on entropy score. Experiments have been performed on bigger dataset. The results reflect better position in comparison to the existing systems.

Keywords Conference ranking · Language model · Influence language model · Entropy · Scientific research

1 Introduction

The ranking of the conferences is an important area of research to assess quality of research work. Such a ranking provides recommendations to the researchers to publish their valuable work in reputed conferences. Several measures have been developed to rank the conferences [3].

This paper solves the problem of ranking the conference papers so that relevant papers of top ranked conferences are placed at first places in the retrieved results. The quality of paper is measured based on its frequency of citations. The main purpose of this research is the semantic ranking of conference papers using influ-

F. Majeed (✉)
Department of Information Technology, Faculty of Computing and Information Technology, University of Gujrat, Gujrat, Pakistan
e-mail: fiaz.majeed@uog.edu.pk

R. A. U. Haq
Department of Computer Science, Faculty of Computing and Information Technology, University of Gujrat, Gujrat, Pakistan
e-mail: azhar4u@live.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_17

ence language model (InLM). In conference paper ranking, we utilize search and ranking technique that use query string to rank the conference papers. To evaluate the InLM technique, a dataset containing 32,000 papers and 50 user queries is built. We have performed comparative analysis of InLM with the existing methods.

The motivation of this research work is that quality researchers need to search the quality conference papers. As conferences have no standardized impact factor system similar to journals, this research work provides them quality conference papers (which is measured based on citations of the particular paper). The relevant documents are then ranked based on calculation of entropy that uses citations of the paper.

The key contributions of this paper are summarized as follows:

- We have proposed a conference papers ranking technique named as InLM.
- We have presented a ranking algorithm based on entropy to rank the conference papers.

The paper has been organized as follows: The review of related works is comprehensively described in Sect. 2. Later, the proposed solution is presented in Sect. 3, and experimental evaluation is provided in Sect. 4. Finally, conclusion and future work is given in Sect. 5.

2 Related Works

Existing methods on conferences ranking have achieved less accuracy in comparison to the journals ranking. A review of those methods is provided to justify further improvement in this domain: the ranking of authors and conferences based on citations is performed by adding time factor. It is the extension of PagRank for yearly conferences ranking. The DBLP dataset is used for experiments and top 10 authors and conferences have been provided [7]. In another work, conferences are ranked based on publication of their extended papers in top journals. For this purpose, acknowledgment information of finance journals is taken and analyzed. The dataset comprised footnotes of 3000 journal articles and 9000 conference articles. Based on this, 47 conferences have been ranked [4].

Currently, two citation analysis systems only exist for journals ranking, which are CiteSeer and Institute for Scientific Information (ISI). The ranking system for the conferences is equally important. Thus, Scientific Collection Evaluator by using Advanced Scoring (SCEAS) system is developed that considers age of the conference. Using DBLP dataset, a web-based application is built [6].

In computer science (CS) domain, few ranking methods exist. Thus, ranking of CS conferences is performed using RsIT technique. It uses a set of conferences, relatedness measure and baseline method to rank the top CS conferences. The Microsoft Academic dataset is used for experiments. The comparison is carried out with the existing CCF rating list. This method provides better results in comparison

to the CCF list [2]. An analysis of the criteria of existing measures for CS conferences is performed. The conference ranking is compared using bibliometric indicators and acceptance rate. The experiments are carried out on three datasets: RankX, Perfil-CC and ERA2010 [3]. The health of conferences related to the software engineering is checked by defining a suite of metrics. The comparison of conferences by parameters like author, PC candidates and scientific prestige is used. Based on the results, 11 top conferences are selected for the interval of 10 years [8]. A technique to rank the CS conferences is proposed with the extension of existing metrics. It generates self-organizing maps for the purpose of ranking. This method is evaluated in correlation to ERA using Microsoft Academic Search dataset including 3442 conferences [1].

The extension of h-index is presented for the journals and conferences ranking. For this purpose, yearly h-index and normalized yearly h-index have been proposed. The DBLP dataset is used with 2024 conferences [5].

In the existence of above conferences ranking methods, no system could be used as standard similar to the ISI. This is due to the fact that current methods do not achieve the required level of accuracy. It is still required to work on achieving the accuracy. This work is build based on this idea.

3 Proposed Solution

The architecture of the InLM is shown in Fig. 1. The input query is processed on the pre-processed dataset. Further, the retrieved results are ranked based on the proposed InLM-based ranking algorithm.

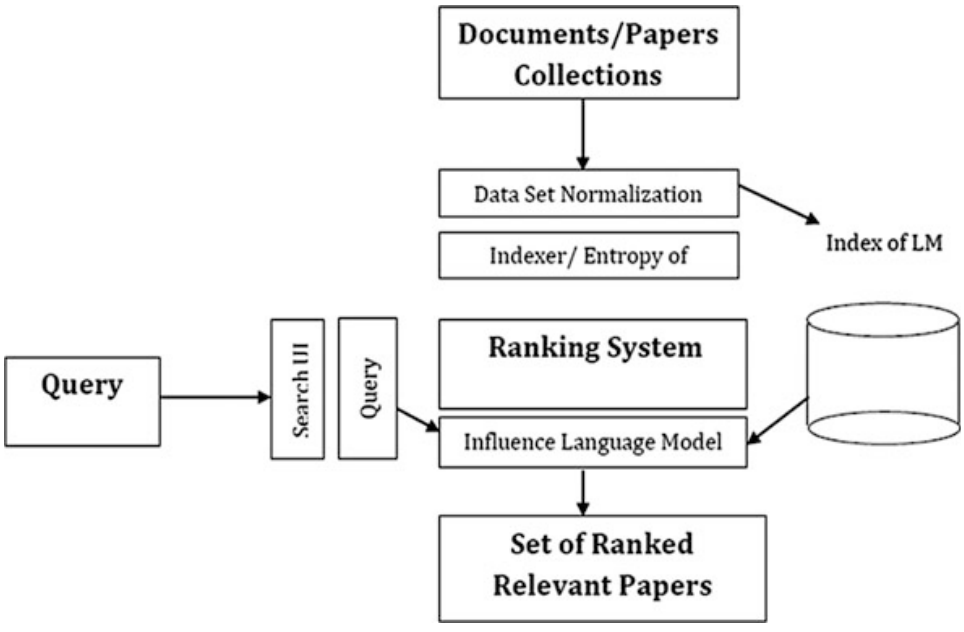


Fig. 1 InLM ranking system

3.1 Entropy Score

As shown in Eq. (1), the entropy of the conferences can be calculated as:

$$\mathrm{Ent}(v) = -\sum_{i=1}^{m} w_i \log_2(w_i) \tag{1}$$

where word_i has the probability w_i at venue v. Individually, all publications of each conference can be used for computation of the probability for input query.

The entropy of papers is calculated as given in Eq. (2):

$$\mathrm{Ent}(p) = \sum_{i=1}^{m} w_i \log_2(w_i) \tag{2}$$

In a paper (p), w_i provides the probability of word_i.

3.2 Calculation of Citations

In comparison to the DBLP, it was tricky to develop a crawler for the Google Scholar. The procedure was to input the query in the search box of Google Scholar and then find out the relevant paper from retrieved results after pressing the search button. Google Scholar provides the citation count for each paper. This process was required to be automated. Initially, the papers were searched manually for analysis. The URL of the relevant results was changed programmatically in such a way that half part of URL was updated to route the citations in our database. The search page of Google Scholar was explored by using the PHP's html document object model for required data, thus put all the citations in Wamp Server and updated the database. For the improved language model, citations of each ranked paper were measured.

3.3 Influence Language Model

This method measures the worth of a paper, whether it is published in low- or high-level conference. The entropy of the conference is calculated to check the level of a conference. The publication of a paper in certain level of conference is proportional to the number of citations. For instance, a paper published in high-level conference will have increased citations, whereas low citations will be received for the paper published in a low-level conference. The InLM increases the score for high-level conference having low entropy, whereas for the low-level conference with high entropy, the method decreases the score.

Two thresholds of entropy 1.8 and 1.85 have been defined and taken as limiting values. The conference or venue is considered as high level in case the entropy value remains below 1.8. Further, in case the entropy value reaches above 1.85, the venue is considered as low level. In influence language model, the probability of a query over document $P(q|d)$ and the resulted value with the entropy result is multiplied. The entropy of the high-level venue is multiplied to increase the score as high-level venue gain smaller value of entropy then multiply the resulted entropy value with $P(q|d)$ (Eq. 3). The entropy is first divided with 3 for decreasing its value for low-level venues and then $P(q|d)$ is multiplied with the resulted entropy value (Eq. 4). We simply multiply $P(q|d)$ with entropy for the venues whose value of entropy is between the range 1.8 and 1.85. The proposed composite model does not consider venues influence.

$$\begin{aligned} &\text{If ent } < 1.8 \\ &\text{ent_value} = \text{ent} \times 3 \\ &\quad \text{Elseif ent } > 1.85 \end{aligned} \tag{3}$$

$$\text{ent_value} = \text{ent}/3 \tag{4}$$

The values of language model are multiplied with entropy values. So the main Eq. (5) of InLM is as follows:

$$\text{InLM} = \text{LM} \times \text{ent_value} \tag{5}$$

In Eq. (5), the entropy value of each conference paper is multiplied with each paper of language model result.

It is necessary to prepare the data before input to the machine learning algorithm, so the pre-processing procedure is given in Algorithm 1.

Algorithm 1

1. Extract all the titles from html pages.
   - Read the html file in java.
   - Patternpt2 = Pattern.compile(“b>.*?</b”);
2. Remove stop words from extracted data.
   - Give input of the raw file.
   - By using the string tokenizer, string is matched. The stop words are removed if those are matched.
   - Then line-by-line matching of each token is performed.
   - StringTokenizer stnz = new StringTokenizer(“*+<>,&-:;.()[]$#@!/\\’\”` ~ ^? ”);
   - By using the above code, each token is matched one by one.

3. Conversion in lower case
   - The search becomes much easier by converting these letters in lowercase. Lowercase function is used for this purpose.
   - Token = token.toLowerCase ();.
4. Frequency Removal
   In the research paper titles, there are many words which are used frequently by many researchers. If a word is matched more than thrice, then it will be removed from the data.
5. Normalize the whole data.

3.4 Application of the InLM

For instance, three papers say P1, P2 and P3 are chosen from the dataset. According to the existing language models, all have the same score because of having equal p (q|d). Let according to previous models,

$$\begin{aligned} P(q|d) \text{ for } \quad & P1 = 0.85 \\ P(q|d) \text{ for } \quad & P2 = 0.85 \\ P(q|d) \text{ for } \quad & P3 = 0.85 \end{aligned}$$

These papers should not have the same scores as they contain the same length of query and published in different conferences, as well as have same number of query terms.

The venue's entropy can be calculated as:

Let

doc 1 entropy = 1.87 (P1 belongs to low-level conference)
doc 2 entropy = 1.69 (P2 belongs to high-level conference)
doc 3 entropy = 1.95 (P3 belongs to low-level conference)

The following result is achieved when the results of previous language models are divided with entropies of conference.

Doc 1 = 0.85 * (1.87/3) = 0.52 (P1 is also low-level conference)
Doc 2 = 0.85 * (1.69/3) = 4.30 (P2 is of high-level conference)
Doc 3 = 0.85 * (1.95/3) = 0.55 (P3 is of low-level conference)

According to the values, the level of conference is being decided.

4 Experimental Evaluations

The experimental evaluation is performed on big dataset. In this respect, the details of the dataset are given below.

4.1 Dataset

Conferences data:

- Data of 100 conferences
- Around 32,000 papers
- Collection of all the title of publications

Query set:

We have developed a query set for experimental purpose. Few of the queries are given below:

- Machine language
- Data mining
- Formal languages
- Group theory
- Information retrieval
- Computer networks
- Semantic web
- Analysis of algorithm
- Natural language processing
- Databases

These queries are used to find the best conference papers from the conferences data. In other words, quality papers of the authors in the relevant field are ranked by the proposed InLM. The dataset is made up of different CS conferences and more than 32,000 research papers of these conferences. The searching technique used for this purpose is content-based ranking. Specifically, there are 100 CS conferences that are extracted from the DBLP. The research papers are taken in the duration of 2003–2007 (5 years).

The support vector machine (SVM) algorithm has been used for the classification of the data and the results are measured using precision, recall and F-measure.

4.2 Pre-processing

Initially, the whole data are pre-processed. In Fig. 2, raw data in html pages are shown, which contain conference information, that is, author's name, title of the paper and name of the conference where the paper has been published.

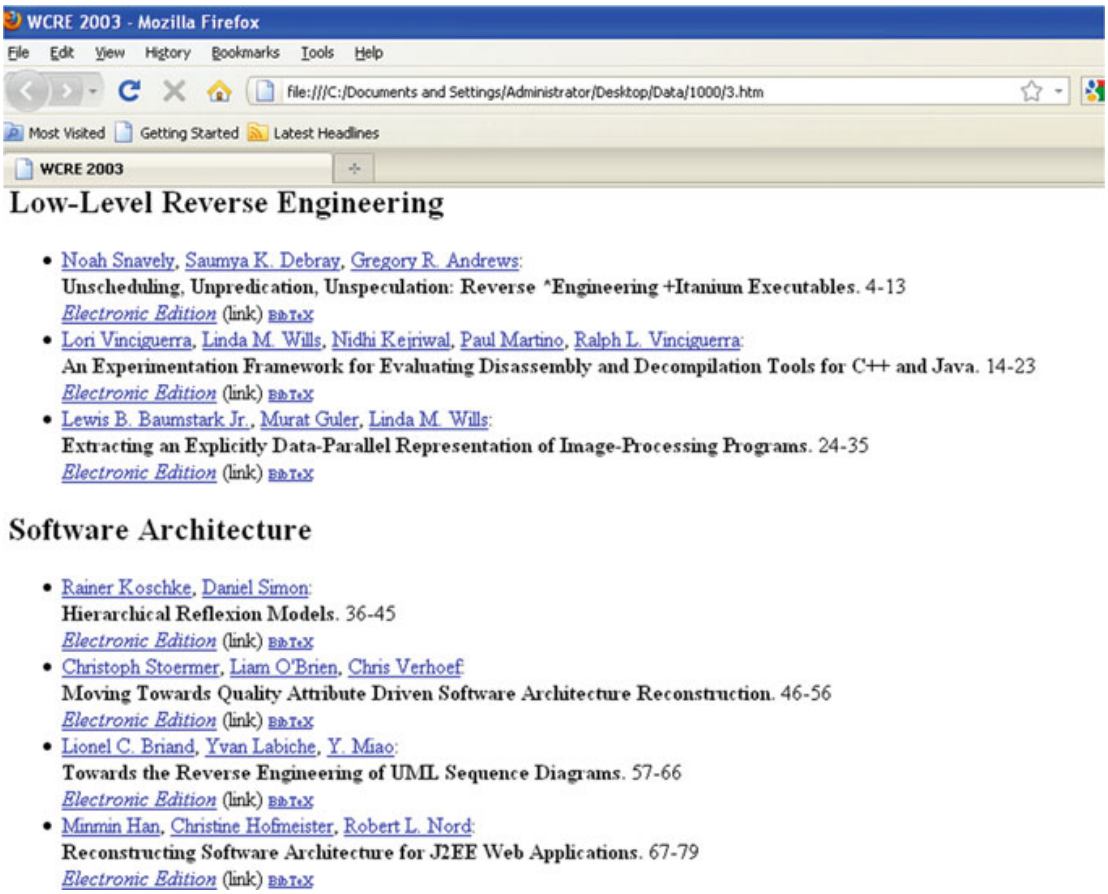


Fig. 2 html pages including raw data

4.3 Results

In this section, we produce the ranking result of our language model and some key terms are used for showing the results of both language models. These terms include Name (title of research paper), Cit (citations), LM (language model) and InLM. The results about one of the query from the query set are given in Table 1.

These are the average results of given base queries using the InLM, including the citations and entropy of the papers for particular input query strings.

In Table 2, the comparison of two models has been taken under consideration, one is previously implemented LM for average of 10 queries and showed the value

Table 1 Top 2 conferences against "data mining" query

S#	Name	Cit	InLM	Name	Cit	LM
1	Region restricted clustering geographic data mining	6	1.875	Data mining track editorial	0	0.675
2	Measuring effectiveness agile methodologies data mining knowledge discovery	4	1.627	Optimization of language data mining	0	0.607

Table 2 Average citations results of top 10 queries

Models	Average citations for 10 queries
LM (VSM)	184.92
InLM	457.36

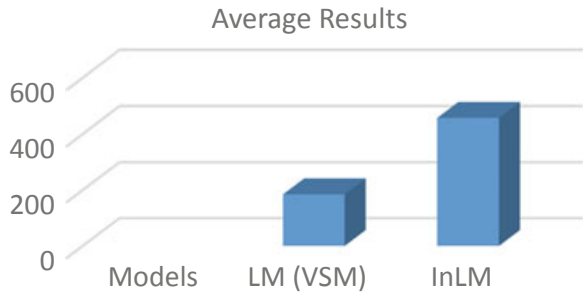


Fig. 3 Graphical representation

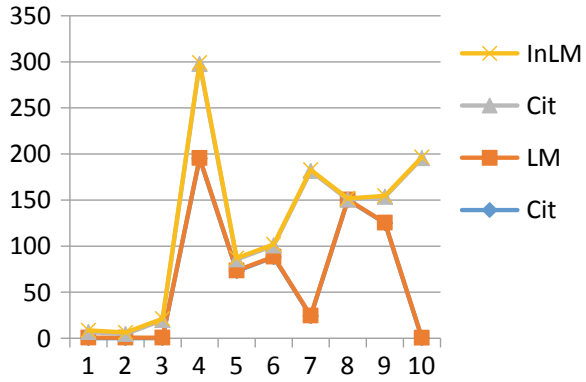


Fig. 4 Comparison of results for data mining query

of 184.92 after setting parameters. On the contrary, the proposed model InLM with the same parameters has given a value of 457.36 with more accuracy and precision for average value of 10 queries.

Figure 3 shows the models along the horizontal axis and number of citations along the vertical axis. The graph clearly mentions that the model InLM has better average results, including entropy results, in more than 400 citations than the LM which is less than 200 citations.

In Fig. 4, the result of InLM is better than LM for "data mining" query.

In Fig. 5, graph depicts score comparison of LM and our proposed model InLM, which shows much better results than LM and has a better ranking for the conference papers.

In Fig. 6, line chart shows ranking result of LM and Fig. 7 shows the result for the InLM.

It has been clearly seen that InLM clearly shows better ranking including entropy of the paper for query. In addition, classification results are shown in Table 3.

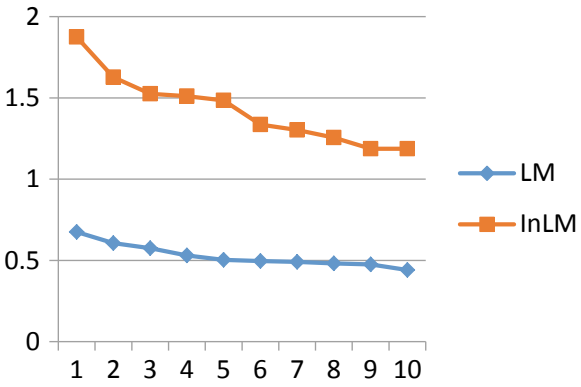


Fig. 5 Comparison of language models results for data mining query

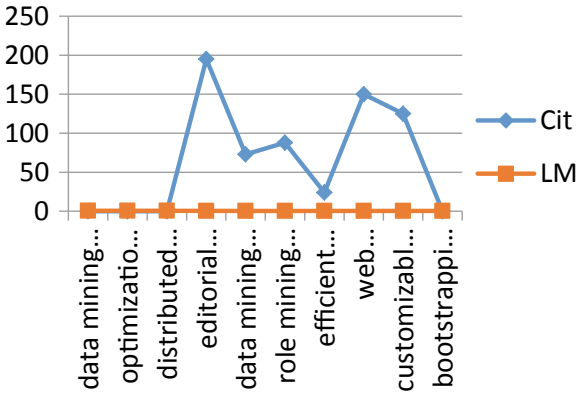


Fig. 6 Language model results of ranked conference papers

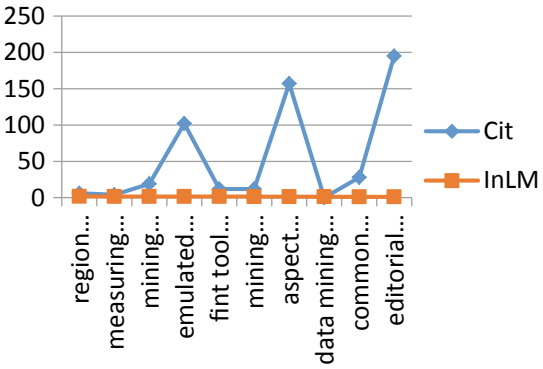


Fig. 7 Language model results of ranked conference papers

Table 3 Classification results

Precision	Recall	Measure-F
82%	79%	81%

5 Conclusions

In this work, the ranking of the conferences is performed using the entropy measure. The InLM system computes entropy of the conferences to generate a ranking list. Further, a comparative analysis is performed with LM. The InLM shows better results in comparison to LM. Furthermore, time factor and names of authors' papers can also be incorporated to improve the ranking.

References

1. Da Silva Almendra, V., Enăchescu, D., Enăchescu, C.: Ranking computer science conferences using self-organizing maps with dynamic node splitting. Scientometrics **102**(1), 267–283 (2015)
2. Effendy, S., Yap, R.H.: Investigations on rating computer sciences conferences: an experiment with the Microsoft Academic Graph dataset. In: Proceedings of the 25th International Conference Companion on World Wide Web, pp. 425–430 (2016)
3. Küngas, P., Karus, S., Vakulenko, S., Dumas, M., Parra, C., Casati, F.: Reverse-engineering conference rankings: What does it take to make a reputable conference? Scientometrics **96**(2), 651–665 (2013)
4. Reinartz, S.J., Urban, D.: Finance conference quality and publication success: A conference ranking. J. Empir. Finance **42**, 155–174 (2017)
5. Sidiropoulos, A., Katsaros, D., Manolopoulos, Y.: Generalized Hirsch h-index for disclosing latent facts in citation networks. Scientometrics **72**(2), 253–280 (2007)
6. Sidiropoulos, A., Manolopoulos, Y.: A new perspective to automatically rank scientific conferences using digital libraries. Inf. Process. Manag. **41**(2), 289–312 (2005)
7. Singh, A.P., Shubhankar K., Pudi, V.: An efficient algorithm for ranking research papers based on citation network. In: Proceedings of the 2011 3rd Conference on Data Mining and Optimization (DMO) (2016)
8. Vasilescu, B., Serebrenik, A., Mens, T., van den Brand, M.G., Pek, E.: How healthy are software engineering conferences? Sci. Comput. Program. **89**, 251–272 (2014)

MapReduce mRMR: Random Forests-Based Email Spam Classification in Distributed Environment

V. Sri Vinitha and D. Karthika Renuka

Abstract The furthermost standard message transfer system used on the internet for communication is email. These days spam is a serious concern that causes major problems in today's internet. Spam emails are uninhibited messages that are sent to a large number of beneficiaries arbitrarily. Owing to an overgrowing rise in reputation, the number of unsolicited data has also increased promptly and has led to many security concerns. Although the sufficient number of spam filtering techniques exists, nowadays spammers start discovering innovative practices to escape data that are filtered using the spam filters. Spammers use this communication source for spreading the malware in the name of an executable file. These spam emails waste user's system memory, computing power, and bandwidth of the network. Spam emails have been initiated to progressively damage the integrity of email and destroy the online experience. The research revealed that if the classification algorithms are used with feature selection then that will return the exact results than the standard classification. In this paper, feature selection is done through minimum redundancy and maximum relevance (mRMR) and the classification is done by means of Random Forests in the MapReduce environment. The performance is compared using various measures, namely sensitivity, correctness, and accuracy with the Random Forests in the distributed environment using Spambase dataset.

Keywords Email spam · Minimum redundancy maximum relevance (mRMR) · Random Forests · Spambase dataset

V. Sri Vinitha (✉)
Department of IT, Bannari Amman Institute of Technology, Sathyamangalam 638401, India
e-mail: srivinithavellingiri@gmail.com

D. Karthika Renuka
Department of IT, PSG College of Technology, Coimbatore 641004, India
e-mail: karthirenu@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_18

1 Introduction

Nowadays, the internet has been mostly used for the purpose of exchanging information. Email has started rising on the internet because of its simplicity, cost-effectiveness, and commercial tool for exchanging information and business communication, and so on [1]. It is exposed to many threats because of its easiness. One among them is email hazard which becomes a big trouble over the internet. Spam email is sent to an enormous amount of users since it charges very less amount for transferring data between the different users. When the user receives the wide range of spam emails then it becomes extremely tough to predict whether the received emails are spam or ham, as it consumes more bandwidth on the internet, wasting user's time and memory. Gradually, spam emails are directed through "zombie systems", or through worm-contaminated PCs around the world. Advanced worms permit spammers to access the user's PC indirectly for malicious purposes. At present, there are some important work-related emails in the spam folder. Globally, the amount of users using email is increased from over 88.6 million in 2014 to over 200.3 million in 2018. Today, various types of spam email exist, and certain things include making money through advertisement and so on. [2]. The sender of a spam email looks like real mail that is not judged by some of the existing filtering techniques. The unwanted spam emails getting into user's inbox should be avoided and to increase the accuracy of email classifiers on the user side, spam filtering techniques are used.

2 Related Works

Nowadays, most people have the habit of using the internet for browsing data, communication purpose, business promotion, and an innovation, providing business services and sharing knowledge with others. Henceforth, the internet has started developing promptly. One of the common threats is spam. Spam message is sent to massive number of recipients for the purpose of advertising, marketing products, job offers, and so on with large incomes for spammers. It also wastes user's time and occupies more memory. These days most of the emails received are spam. Therefore, Bassiouni et al. [3] proposed a technique for classifying spam emails with better performance using ten classification algorithms. The ten different classifiers such as logistic regression, decision tree, Bayes net, support vector machine, naïve Bayes, k-nearest neighbors, Random Forests, artificial neural network, random tree, and radial basis function are used for classification. The ten classification algorithm provides accuracy such as 92.4, 90.3, 89.8, 91.8, 89.8, 90.7, 95.4, 92.4, 91.5, and 82.6. The experimental result shows that the system achieves a correctness of about 95.45% by using Random Forest's approach for Spambase UCI datasets.

Email is the most active and fastest method of correspondence to trade data over the web. It is so popular for its cost-effectiveness, easiness, and fastest mode of communication. Since there is an expansion in the number of users using the diverse social sites, there is a massive increase in unwanted messages. Spam can cause the severe financial disaster, loss of data transmission, and wastage of memory. Spam emails have been created to progressively damage the integrity of email and destroy online experience. The problem of spam mail getting into user's inbox is resolved by using spam filtration approach. In this paper, Kaur et al. [4] suggested the N-gram technique for feature selection and improved multilayer perceptron for spam classification. The result comparison is performed on emails collected from Enron dataset shows that the proposed N-gram-based K-enhanced MLP demonstrated higher performance than existing MLP along with low error rate. Future work uses optimization techniques and some other feature selection techniques for selecting the best features for numerical data, text data, and image corpus.

Currently, the internet has turn out to be an essential one in everyone's life. With this increased use, the number of email users is also enlarged nowadays. It is predicted that 90% of the emails sent are spam. This leads to a problem initiated by spontaneous bulk email stated as spam. It is expected that 90% of emails are sent every day with the intention of advertising any product or spreading any kind of malware to the user PCs without any notification. Spam email is sending unwanted communications to different email users. This spam email causes severe harms for internet users such as the delay in sending legitimate messages, wastage of network bandwidth, storage, and computation power. To avoid this problem, various spam filtering techniques exist. Vijayasekara et al. [5] propose machine learning algorithm naives Bayesian with three-layer structure for classifying spam in bulk emails. The feature selection method is used to select the best attributes and categorize spam emails using naïve Bayesian algorithm which helps to increase the correctness of the system.

To select the best attributes from the high-dimensional data, it uses feature selection techniques. The methodologies used for feature selection should be efficient to manipulate multivariate historical data. To elucidate this problem, Radovic et al. [6] recommend the significant and repeated data with some modification for the assessment process. The features can be discrete as well as continuous. F-statistics can be considered for significant data and Pearson correlation or mutual information for assessing repeated data. The recommended technique is measured using gene expression temporal datasets and provides high performance compared to other systems. Forthcoming work will be applied to various algorithms with feature selection approach that chooses the slightest repetition and more importance.

At present, electronic mail has become the popular mode of communication on the internet which is used to exchange information to a large number of beneficiaries. It is used for getting and transferring data among different users. Because of its fastest delivery, less time, and low cost for sending emails to make people use it for advertising products, sending fraudulent mails and phishing mails resulted in

unwanted spam emails. Spam is the unsolicited, unwanted mail that sends messages to a large number of recipients. It also wastes user's time, transmission capacity, and user's system space. Spam filtering approaches categorize the spam email based on the information it contains and sends the spam mails into spam folder. Easwaramoorthy et al. [7] proposed the system consisting of the training and testing phase. In the preliminary phase, the hybrid Firefly GSO algorithm takes the input mails to choose the best attributes from the huge dataset. Then the selected features are classified using three classification algorithms, namely decision tree, neural networks, and naïve Bayes algorithm. Then the result is evaluated using UCI repository Spambase dataset through various metrics.

3 Proposed Methodology

The process of the proposed email spam classification technique as in Fig. 1 can be described in the succeeding phases:

- Prepare the UCI repository Spambase dataset for the proposed system
- In the MapReduce framework, do the feature selection and classify the spam emails

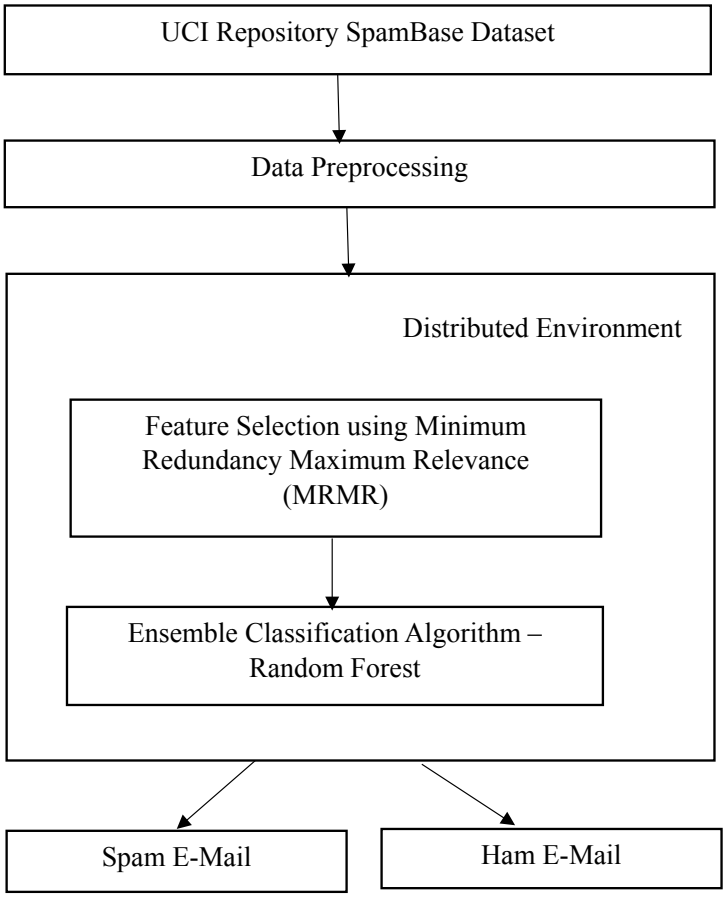


Fig. 1 Framework of the proposed spam email classification

- Feature selection is done by using minimum redundancy and maximum relevance technique to select the best attributes from Spambase dataset
- Then the selected attributes are classified by using Random Forest classification algorithm
- Finally, compare the proposed framework with Random Forest classification in a distributed environment.

3.1 Dataset

In this paper, UCI repository Spambase dataset has been taken for email spam classification. Totally, there are 4601 emails in the Spambase. In this 39.4%, that is, 1813 emails are spam and remaining 60.6%, that is, 2788 are ham emails [8]. There are 58 features in UCI repository Spambase dataset out of which 57 attributes are continuous and one with class label (d). The description of the attributes is given in Table 1 as.

3.2 Data Preprocessing

In real world, Spambase dataset contains some missing values, and this missing data may lead to incorrect or even misleading statistics. Data preprocessing procedure is used to remove noise, inconsistent, and incomplete data in order to obtain a quality result. It is the initial step which makes sure that data is prepared to be analyzed [9]. There are various preprocessing steps available, and here data normalization is done in order the make the dataset suitable for further processing.

Data discretization: Discretization states the method of converting continuous attributes into discretized variables. It transforms data from numeric into nominal data type. The generic representation of Spambase dataset is given by:

$$\text{Spambase Dataset} = d_k; \quad 1 \leq k \leq p \tag{1}$$

Table 1 Description of Spambase dataset

Attribute number	Type of attribute
1–48	word_freq_WORD
49–54	char_freq_CHAR
55	capital_run_length_average
56	capital_run_length_longest
57	capital_run_length_total
58	Class attribute

The Spambase dataset is given to the discretization function to transfer the input data into discretized one. Discretization converts the data into specific range. The highest and smallest value of every attribute is identified and the I interval is followed by taking the proportion between the differed value and the I value.

- For example, at-first, the deviation is calculated for every k value

$$\mathrm{Dev}(k) = \frac{\mathrm{Max}(d_k) - \mathrm{Min}(d_k)}{2} \tag{2}$$

- After calculating deviation values for each row, values are converted to the following condition:

$$\left.\begin{array}{ll} 0, & \mathrm{input} < 1(\mathrm{Dev}(k)) \\ 1, & \mathrm{input} < 2(\mathrm{Dev}(k)) \\ 2, & \mathrm{input} < 3(\mathrm{Dev}(k)) \\ 4, & \mathrm{input} < 4(\mathrm{Dev}(k)) \end{array}\right\} \tag{3}$$

Then, the value which comes under particular range is substituted by interval value in order to convert the input data into discretized data. After discretization function, the training dataset is converted to above conversion Eq. (3) and then each value is converted into 8421 binary conversions as discretized format.

3.3 MapReduce

MapReduce is a framework of Hadoop system designed to examine large dimensions of files in a parallel mode on commodity machines. Its objective is to get used to the available computational resources and process data in a massively scalable way in two processing phases, that is, map and reduce phase.

The entire process of MapReduce can be illustrated using the following notation [10]:

$$\mathrm{Map}{:}(k1, v1) \rightarrow \mathrm{list}\ \{k2, v2\} \tag{4}$$

$$\mathrm{Shuffle}\ \&\ \mathrm{Sort}{:}\ \mathrm{list}\{\mathrm{k2},\ \mathrm{v2}\} \rightarrow (\mathrm{k2}, \mathrm{list}\{\mathrm{v2}\}) \tag{5}$$

$$\mathrm{Reduce}{:}(\mathrm{k2},\ \mathrm{list}\{\mathrm{v2}\}) \rightarrow \mathrm{list}\{\mathrm{k3}, \mathrm{v3}\} \tag{6}$$

In the Map step, Spambase dataset after preprocessing into discretized data is loaded once into a mapper function as in Eq. (4). Then the feature selection is done by using minimum redundancy maximum relevance and the features are delivered by means of a list of key-value objects. The output of mapper will be the sequence of new objects with maximum relevance features without any repetition.

In the Reduce step, the input of a reducer is the intermediate object as in Eq. (5), which shares the same features that will be collected together. The features are classified by means of Random Forest algorithm and the output is obtained by the majority voting of the classification trees that have been formed. The final output will be produced as in Eq. (6).

3.3.1 Minimum Redundancy Maximum Relevance (mRMR)

The process used for obtaining the subset of relevant attributes is a feature selection process, that is, the dimensionality reduction method [11]. Its objective is to attain a better result with the minimum number of features. The filter method used in this paper selects the feature with minimal repetition and maximal importance. mRMR depends on two metrics: the least repetition among the attributes that have chosen and the greatest significance among candidate features and target class. In the Mapper step, the Spambase dataset D is given as input with N samples (4601 instances) and M features (57 attributes) $X = \{x_1, x_2, x_3, \ldots. x_m\}$ and the target classification variable c. The feature selection technique finds the subset of m attributes, R^m from the M-dimensional data, R^M.

Maximal relevance is selecting the features from the Spambase dataset with the maximum significance to the target class c. Either correlation or mutual information may be the measure used to define dependency of variables. Max-relevance is an average of all collective information values among individual attribute x_i in the Spambase dataset and class c as in Eq. (7).

$$max\, D(S, c), D = \frac{1}{|S|} \sum_{x_i \in S} I(X_i, c) \tag{7}$$

The subsets that are identified by the maximum relevance often contain relevant repeated data. mRMR challenges to provide the solution to this problem by eliminating those repeated subsets. The minimal redundancy property is included to choose mutually exclusive features from the Spambase dataset as in Eq. (8).

$$\min\, R(s) = \frac{1}{|S|^2} \sum_{X_i X_j \in S} I\left(X_j, X_j\right) \tag{8}$$

Minimum redundancy maximum relevance is combined to select the features that have the high association with the criterion without any repeated data that can be denoted as in Eq. (9). This gives the subset of features S^* selected using mRMR, which is the output of Map step.

$$S^* = \arg\max_{S \subseteq F}[\max D(S, c) - \min R(s)] \tag{9}$$

mRMR Algorithm

```
INPUT: Data, Features
// Data – Spambase Dataset
// Features - 57 features in Spambase Dataset
OUTPUT: SubsetofFeatures // Set of features selected from Spambase dataset.
for features fi in SpambaseDataset do
        rel = mutualInfomation(fi, classlabel);
        redundant = 0;
        for features fj in SpambaseDataset do
                redundant += mutualInformation(fi, fj);
        end for
        mRMRData[fi] = rel - redundant;
end for
SubsetofFeatures = sort(mRMRData). take(Features);
```

The algorithm for minimum redundancy maximum relevance is given above [11]. In the mRMR system, the core problem is associated with the calculation of the mutual information among two components: either exists as a couple of input features or feature by means of the class. This task is figured among every single pair of features; some of the pairs of features are still unrelated for the final outcome, which is insignificant in high-dimensional problems.

3.3.2 Random Forests

Random Forest is an ensemble learning technique that works best for high-dimensional data and was developed by Leo Breimanand Adel for classification and regression not pruned trees. The bagging technique uses a decision tree for selecting arbitrary features to improve classification performance by combining the results of different classifiers. It uses randomized node optimization. Based on the arbitrary choice of data and variables, the progress of the trees can be determined.

The selected feature S^* from mapper phase is given as input to the reducer phase. With the selected feature, it builds an m number of decision trees and forms a forest. Then the new object which is necessary to be classified is placed under each of the

decision trees for classification. The mail is categorized as ham or junk mail based on each trees decision by voting as in Fig. 2. The prediction of the spam emails is attained by majority voting of the decision trees that have been formed, which will be the results of the reducer phase.

Every tree is developed as follows [12]:

1. Spambase dataset after feature selection contains S^* features; the training set will be sample M cases from the subset of features S^*.
2. Choose the best split between the M features.
3. Develop each one of the trees to the biggest possible level without pruning.
4. The spam emails are estimated by grouping the predictions of m trees by majority voting technique.

Certain features of the Random Forest algorithm are given as follows:

1. It provides better accuracy and also lighter compared to other ensemble algorithms.
2. It does not overfit the data because it does not have any insignificant data.
3. It is sturdy to noise and outliers and can also handle uneven datasets.

Pseudocode for Random Forests Algorithm

```
Begin RandomForests
      Input:   N:number of nodes
               F: Spambase Dataset with the subset of features S*
               T: Total amount of trees to be created
      Output: MV : the class label spam or ham with majority voting
       While ending condition is incorrect do
            The underneath phases are used to build the tree from the training data:
            1)  Select the features f arbitrarily from F; where f ≪ F
            2) For node t, compute the best split point between the f features
            3) Based on the best split, divide the node into two child nodes
            4) Iterate the steps 1, 2 and 2 up to n amount of nodes has been reached
         Construct the forest by iterating steps from 1–4 for T periods
       End While
         Display all the created trees
         The new sample is applied to all the created trees beginning from the initial node
         Allocate the new object to the external node with the resultant class label.
         Aggregate the votes of all the nodes in the trees.
         Display the final output MV, with the highest vote.
End RandomForests
```

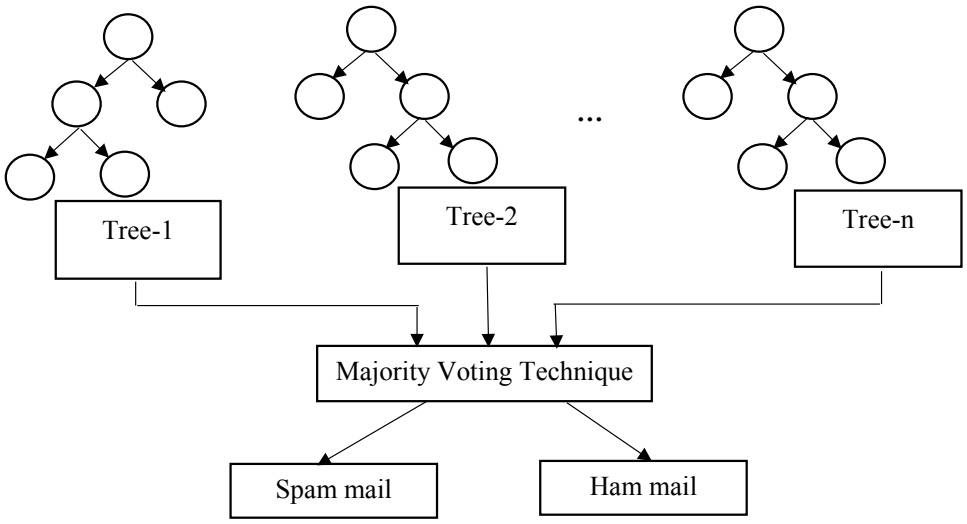


Fig. 2 Random Forests algorithm

4 Performance Analysis

Table 2 shows the performance comparison for Spambase dataset using Random Forests in the distributed environment, and MapReduce minimum relevance maximum redundancy (mRMR) feature selection with Random Forests algorithm. The different assessment metrics have been used for investigating the performance.

Table 2 Comparison of precision, recall and accuracy

Model	Precision	Recall	Accuracy
Random Forests in distributed environment	0.837624	0.846754	0.840721
MapReduce mRMR feature selection with Random Forests	0.852746	0.864657	0.860824

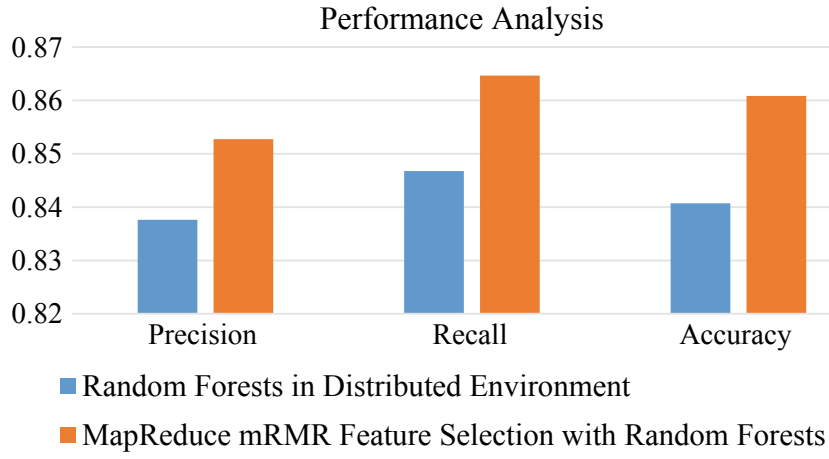


Fig. 3 Performance analysis

By analyzing the plotted graph in Fig. 3, the performance of the proposed email spam classification is significantly improved by selecting the best features using minimum redundancy maximum relevance (mRMR) in MapReduce environment, and the classification is done using the Random Forest.

5 Evaluation Metrics

The efficiency of the recommended methodology for discovering spam emails has been estimated by means of some metrics such as precision, recall, and accuracy. The assessment metrics can be defined as [13]:

Precision

It is the fraction of spam emails that are correctly categorized as spam from the given input emails. Precision illustrates the exact correctness and can be represented using Eq. (10) as:

$$\text{Precision} = \frac{\text{Spam emails correctly classified}}{\text{Spam emails correctly classified} + \text{Ham emails wrongly classified}} \tag{10}$$

Recall

It is the fraction of spam emails that are categorized as spam from the given input emails. Recall illustrates the completeness and is represented using Eq. (11) as:

$$\text{Recall} = \frac{\text{Spam emails correctly classified}}{\text{Spam emails correctly classified} + \text{Spam emails wrongly classified}} \tag{11}$$

Accuracy

It is determined as the percentage of addition of all true positives and true negatives to the entire amount of emails. Accuracy can be denoted by means of Eq. (12) as:

$$\text{Accuracy} = \frac{\text{Spam emails correctly classified} + \text{Ham emails correctly classified}}{\textit{Total emails}} \tag{12}$$

6 Conclusion

Email is an efficient, fastest, and cost-effective communication system in today's world. In the email, spam is the common problem, which sends data to the large number of recipients for distributing malware, advertising products, phishing, and so on. Spam emails take excess time in deleting junk mails, surplus disk space, and network bandwidth. Therefore, the filter is needed with great precision to filter the unsolicited emails. To filter the spam emails, here feature selection is done using MapReduce minimum redundancy maximum relevance (mRMR) to select the best features for classification. The selected features are classified by using the Random Forests algorithm. In the Random Forests, by means of voting technique, it classifies the junk mails and ham emails. The comparison indicates that the proposed email spam classification gives better accuracy than classification using Random Forests in the distributed environment.

Acknowledgements The authors sincerely thank the University Grants Commission (UGC), Hyderabad for granting funds to carry out this work.

References

1. Zhang, Y., He, J., Xu, J.: A new anti-spam model based on e-mail address concealment technique. J. Nat. Sci. **23**(1), 79–83 (2018)
2. Khalaf, O.I., Abdulsahib, G.M., Salman, A.D.: Handling dimensionality reduction in spam e-mail classification. J. Adv. Res. Dyn. Control Syst. **10**(1), 691–697 (2018)
3. Bassiouni, M., Ali, M., El-Dahshan, E.A.: Ham and spam e-mails classification using machine learning techniques. J. Appl. Secur. Res. **13**(3), 315–331 (2018)
4. Kaur, J., Priyanka: Feature selection based efficient machine learning technique for email spam predicition. Int. J. Eng. Appl. Sci. Technol. **2**(12), 13–19 (2018)
5. Vijayasekaran, G., Rosi, S.: Spam and email detection in big data platform using naives bayesian classifier. Int. J. Comput. Sci. Mob. Comput. (IJCSMC) **7**(4), 53–58 (2018)
6. Radovic, M., Ghalwash, M., Filipovic, N., Obradovic, Z.: Minimum redundancy maximum relevance feature selection approach for temporal gene expression data. BMC Bioinform. **18** (8), 1–14 (2017)
7. Easwaramoorthy, S., Thamburasa, S., Aravind, K., Bhushan, S.B., Rajadurai, H.: Heterogeneous classifier model for e-mail spam classification using FSO feature selection method. In: International Conference on Inventive Computation Technologies (ICICT), pp. 1–6 (2017)
8. Awad, M., Foqaha, M.: Email spam classification using hybrid approach of Rbfneural network and particle swarm optimization. Int. J. Netw. Secur. Appl. (IJNSA) **8**(4), 1–12 (2016)
9. Sri Vinitha, V., Karthika Renuka, D., Bharathi, A.: E-mail spam classification using machine learning in distributed environment. J. Comput. Theor. Nanosci. **15**(5), 1688–1694 (2018)
10. Nesi, P., Pantaleo, G., Sanesi, G.: A hadoop based platform for natural language processing of web pages and documents. J. Vis. Lang. Comput. **31**, 130–138 (2015)
11. Ramirez-Gallego, S., Lastra, I., Martinez-Rego, D., Bolon-Canedo, V., Benitez, J.M., Herrera, F., Alonso-Betanzos, A.: Fast-mRMR: fast minimum redundancy maximum relevance algorithm for high-dimensional big data. Int. J. Intell. Syst. **00**, 1–19 (2016)

12. Ozarkar, P., Patwardhan, M.: Efficient spam classification by appropriate feature selection. Glob. J. Comput. Sci. Technol. Softw. Data Eng. **13**(5), 49–57 (2013)
13. Vaishnavi, N., Thiyagarajan, K.: A study on prediction of malicious program using classification based approches. Int. J. Comput. Sci. Mob. Comput. IJCSMC **7**(5), 38–46 (2018)

The Impact of Sustainable Development Report Disclosure on Tax Planning in Thailand

Sathaya Thanjunpong and Thatphong Awirothananon

Abstract This paper aims to examine the impact of sustainable development report disclosure (hereafter called SDRD) on the tax planning (hereafter called TP) of listed companies, which exclude the financial sector, in the Stock Exchange of Thailand. The data of this paper is based only on the year 2016 and the sample size consists of 337 companies from seven industries. The questionnaires from Global Report Initiative are used for evaluation of SDRD. The TP is also measured by the ratio of total tax expenses to total assets. Overall, this paper finds that the SDRD had a statistically negative effect on the TP. This indicates that companies with good SDRD practices could have a higher level of TP. Regarding control variables, financial leverage and capital intensity had a statistically positive effect on the TP, while profitability and family control had a statistically negative effect on the TP. This paper further divides the sample into family and non-family companies to examine whether there is any different effect of SDRD on the TP. The results further indicated that the relationship between SDRD and TP was significantly negative for the family companies. The relationship was, thus, weak and insignificant for the non-family companies. The results are useful to market regulators since they could make some decisions about adjusting some rules and regulations or give some incentives to encourage the Thai listed companies to perform better sustainable development practices.

Keywords Listed companies · Sustainable development report disclosure · Tax planning · Thailand

This paper is our original contributions and has not been submitted or accepted for publication anywhere else. All remaining errors are ours.

S. Thanjunpong · T. Awirothananon (✉)
Faculty of Business Administration, Maejo University, Sansai, Chiang Mai 50290, Thailand
e-mail: thatphong@hotmail.com

S. Thanjunpong
e-mail: sathayat@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_19

1 Introduction

Corporate social responsibility (hereafter called CSR), as a consequence of corporate governance (hereafter called CG), in Thailand particularly seems to have become more important. The Corporate Governance Center is established in the year 2002 by the Stock Exchange of Thailand (hereafter called SET) to support the listed companies in developing their CG system by complying with the international assessment regarding the improvement of CG [8]. As a result, the voluntary CG rating was subsequently introduced in the year 2008. The SET was set up [31] to continue the long-term sustainable growth of the capital market in the year 2007 and launched the "Thailand Sustainability Investment" scheme in the fourth quarter of 2015 by applying three criteria, which are economic, social, and governance criteria, respectively. The Thai listed companies and related organisations consequently become more involved with CSR. This is supported by the additional sections on the disclosure reports and the annual report; however, it is not officially required to disclose the CSR section. There is also a significant increase in the number of activities and projects of listed companies. Sustainable development report disclosure (hereafter called SDRD) is also a report published by companies about the practice of measuring, disclosing, and being accountable for their performance from their everyday activities, while they try to operate their business towards the sustainable development. This report could provide balanced and reasonable representations of sustainability performance, including both positive and negative contributions. It could also be considered as non-financial reporting, or triple bottom line reporting, or CSR reporting. The quality of SDRD is expected to increase the overall value creation, which is the increase in the performance of organisation.

The tax planning (hereafter called TP) is also along the continuum depending upon how aggressive the activity is in reducing total tax expenses. The tax avoidance (hereafter called TA) is some activities that reduce companies' cash-effective tax rate within the legal framework of tax regulations over a long period of time [12, 29]. These activities take some advantages of the loopholes in the tax regulations in order to reduce companies' tax liabilities [19, 34]. Companies, for example, could choose some advantageous methods in the financial reporting and could reduce their tax liabilities; consequently, the TA for this paper is a part of TP, although in some prior studies, including Chen et al. [6], Lestari and Wardhani [24], Zhang et al. [39], and Ying et al. [38], the effective tax rate is popularly used as the TP measurement. This paper, therefore, has developed a new measurement of TP. It is measured by the ratio of total tax expenses to total assets. This measurement is popularly used as the TP measure in Thailand [35, 36]. Under the assumption of corporation with the same profitability we should have the same total tax expenses at any condition, as well as an earning management of the statement of financial position is more difficult than that of the income statement. This leads to

the conclusion that companies with CSR should have the total tax expenses as well. The companies with higher SDRD practices would, therefore, have higher total tax expenses, which indicate the lower TP. The SDRD quality is also expected to improve the overall value creation, which is the reduction in TP.

Many previous papers show that SDRD has a significant impact on the increasing TP [5, 9, 22, 28]. In developed countries, the authors of [22] investigated a legitimacy theory during the 2001–2006 period in the Australian corporations by comparing the CSR disclosures of tax aggressive corporations with those of non-tax aggressive corporations. Their results show that CSR has a statistically and positively significant effect on the TA. Davis et al. [9] further investigated the relationship between CSR and corporate tax payment in the US-listed companies during the year 2002–2011. Their results show that CSR has a significantly positive effect on the TP. Preuss and Preuss [28] also found that CSR and corporate tax payments act as substitutes in the European public companies. In the developing region, Chen [5] found that the CSR disclosure significantly strengthens the possibility of TP in the Chinese-listed companies from 2008 to 2014. However, many prior studies show that SDRD has a significant impact on the decreasing TP. In the developed countries, Hoi et al. [18] found a statistically negative relationship between irresponsible CSR and TA in the US public companies during the 2003–2009 period. Lanis and Richardson [23] further showed that companies with the higher level of CSR disclosure could have the lower level of corporate TA for the 2008–2009 financial year. The CSR disclosure, therefore, has a significant negative effect on the TA in the Australian listed companies. In the Asia region, Sari and Tjen [30] recently revealed that the CSR disclosure has a significantly negative effect toward the TA in the Indonesian-listed companies during the 2009–2012 period.

Some earlier papers, including Laguir et al. [20], Nadiah et al. [27], and [33], further showed that SDRD is not significant or cannot summarise the relationship between CSR and TA. In the developed region, Laguir et al. [20] revealed that the relationship between various CSR dimensions and TA is either positive or negative in the French companies during the 2003–2011 period. In the developing countries, Nadiah et al. [27] further found the relationship between CSR disclosure and corporate TA is either positive or negative in Malaysia in the year 2014. Susanti [33] also investigated the effect of CSR on the TA for manufacturing listed companies in the Indonesia Stock Exchange in the year 2012–2014. These results show that CSR does not have any influence on the TA. On the other hand, SDRD is not associated with the TA. Recent empirical studies find a mixed evidence of the impact of SDRD on the TP. This paper, therefore, aims to search the benefit of SDRD implementation to the TP.

2 Research Methodology

The population of this paper are listed companies, which are excluded financial sectors (58 companies), in the Stock Exchange of Thailand (hereafter called SET) in the year 2016. The purposive sampling method is used for sample selection. The selection is done in accordance with two criteria as follows:

(a) Companies that publish financial statements, annual report, and Form 56-1 in their website and from the SETSMART database with completed period of 2016.
(b) Companies that have the sustainable development report during the study period and their necessary data of TP is available.

According to the above two criteria, the sample consists of 337 companies in this paper. These companies are categorised into seven industries according to SET categorisation. It comprises (1) 35 agro and food industry companies, (2) 28 consumer products companies, (3) 64 industrial companies, (4) 74 property and construction companies, (5) 24 resources companies, (6) 81 services companies, and (7) 31 technology companies, respectively.

The dependent variable is the TP, which is calculated as the total tax expenses for the year divided by the total assets. The most important independent variable is SDRD, which is measured by the Global Report Initiative (hereafter called GRI) index score. This index score is calculated from 18 separate criteria to quantify the overall GRI principles from the Global Sustainability Standards Board [14]. The scorecard criteria span three dimensions of the GRI principles, which are economic dimension that has five criteria, environmental dimension that has nine criteria, and social dimension that has four criteria, respectively, as shown in Table 1.

This paper acquires SDRD index scores from Form 56-1, SETSMART database, and annual reports for the year 2016. These index scores are also adjusted to consider the subtleties of Thai laws and regulations. The inclusion of each attribute is scored on a binary basis as “yes” (included) or “no” (not included). Each “yes” answer is equal to the value of one point, while “no” is equal to the value of zero point. This paper also includes some control variables that have been shown to have a statistically significant impact on the TP [36]. This paper uses firm size (hereafter called SIZE), which is measured by the natural logarithm of total assets. The financial leverage variable (hereafter called LEV) is also included in this paper. The LEV variable, which is a proxy of the differences in the financial structure of companies, is computed as the percentage of total debt to total assets. This paper also includes the return on equity (hereafter called ROE) by dividing the net profit to total equity, as an indicator of companies’ financial performance. In addition, this paper uses the capital intensity variable (hereafter called CAP) by dividing the property, plant, and equipment to total assets, as a proxy of companies’ working capital management.

Table 1 SDRD index score

Dimensions	Details
Economy dimension	1. The results on economic growth as net profit of companies were shown
	2. The results on a growth of community economy as nearby community incomes created partly by companies were demonstrated
	3. The results on a partly growth of national economy as corporate income tax and others related to business operation were exhibited
	4. The results on a circular flow of capital among those diverse stockholders in different business companies were presented
	5. The results on corporate information related to policy and campaign against corruption inside their business organisations were proven
Environment dimension	1. Companies that affected human life and exiting organism were demonstrated
	2. Companies that affected ecosystem and quality of water, soil, and air were exhibited, respectively
	3. Companies that affected our natural resources were shown
	4. Companies that affected input factors, namely, energy, and water were demonstrated, respectively
	5. Companies that followed relevant laws and environmental regulations were provided
	6. A disclosure of costs of environmental disposal
	7. Companies that dealt with waste and garbage disposal were provided
	8. Companies that are concerned with their productions and environmental-friendly logistics were demonstrated
	9. Companies that handled with diversified activities like campaign, public relation events, publications, and donation against environmental damage were exhibited, respectively
Society dimension	1. Companies that are concerned with labour forces and employment were demonstrated
	2. Companies that are concerned with human rights were provided
	3. Companies that are concerned with their public accountability were shown
	4. Companies that are concerned with their productions were described

Source Improved from the Global Sustainability Standards (14)

There is a large number of Thai-listed companies that are still family-owned companies, where the family members being the owners have substantial control over both ownership and management [7, 32]. This paper, therefore, uses family control (hereafter called FAM) to separate between family and non-family companies as some previous papers, including Tantiyavarong [35], Chen et al. [4], Landry et al. [21], and Landry et al. [36], show that family companies have a significant impact on the TP. In the developed countries, Chen et al. [4] and Landry

et al. [21] revealed that the relationship between family companies are less TA than their non-family companies because family companies being more concerned with the reputation damage and potential penalty. Bauweraerts and Vandernoot [1] and Martinez and Ramalho [25], however, found that the family companies are more TA than non-family. Recently, Gaaya et al. [13] revealed that the family ownership has a positive effect on corporate TA in Tunisian-listed companies during 2008 and 2013.

In the Asian region, Tantiyavarong [35] found that family companies have positive effects on the TP during the 2001–2007 period. In addition, Thanjunpong and Bangmek [36] further showed that the relationship between family companies is more tax planning than their non-family companies in Thailand during the 2012–2014. Masripah et al. [26], however, found that the entrenchment effect of controlling shareholder has negative effect on the TA in manufacturing industries on the Indonesia Stock Exchange companies during the 2010–2013 period. Furthermore, Bousaidi and Hamed [3] found that the concentration ownership exhibits a negative association with the effective tax rate, which indicates there is a higher TP in Tunisian-listed companies during the 2006–2012 period. Therefore, the FAM variable, which is a dummy variable, has a value of one for the family companies and zero otherwise for non-family companies. The family companies are defined as when a stake of 25% or above of companies' outstanding shares belongs to the major family shareholders.

3 Research Methods

This paper uses the multiple regression analysis technique to determine the relationship between SDRD and TP. According to the objective of this paper, the multiple regression on the full sample model will be utilised as follows:

$$\text{Model 1}: \ \text{TP}_i = \beta_0 + \beta_1 \text{SDRD}_i + \beta_2 \text{SIZE}_i + \beta_3 \text{LEV}_i + \beta_4 \text{ROE}_i + \beta_5 \text{CAP}_i + \beta_6 \text{FAM}_i + \sum_{j=7}^{12} \beta_j \text{IND}_j$$

In addition, this paper separates the full data to two sub-groups (family and non-family companies) to examine whether there is any different effect of SDRD on the TP among the family companies (Model 2.1) and the non-family companies (Model 2.2) as follows:

$$\text{Model 2}: \ \text{TP}_i = \beta_0 + \beta_1 \text{SDRD}_i + \beta_2 \text{SIZE}_i + \beta_3 \text{LEV}_i + \beta_4 \text{ROE}_i + \beta_5 \text{CAP}_i + \sum_{j=6}^{11} \beta_j \text{IND}_j$$

4 Results and Discussion

The descriptive statistics for the full and separate sub-samples for family and non-family companies in the SET for the year 2016 are reported in Table 2. It also presents the mean values and the *t*-statistics that test for the mean differences of all variables between family and non-family companies. The results show that the average value of TP for the full sample was 0.011%, which was a low level (representing the TP was a high level), while the values for both family and non-family companies were 0.011 and 0.009%, respectively. The results also reveal that the TP value shows statistically significant difference between family and non-family companies at the 0.10 level. The results show that the family companies had a lower TP than the non-family companies. This implies that the family companies had higher level of TP than the non-family companies.

The values of descriptive statistics in Table 2 show that the mean value of SDRD for the full sample was 11.580. It also shows that the mean value of SDRD for the family companies was 11.730, which is higher when compared with the non-family companies, 11.110. However, the *t*-statistics for the mean difference of SDRD between family and non-family companies was not statistically significant. This indicates that both family and non-family companies had the same level of SDRD implementation. Additionally, the average value of SIZE for the full sample was 22.606, while the mean values for both family and non-family companies were 22.700 and 22.330, respectively. The results indicated that the mean value of SIZE for the family companies was greater than that for the non-family companies, since the *t*-statistics of the mean differences were statistically significant at the 0.05 level. Moreover, the average value of LEV for the full sample was 0.438. The mean value for the family companies, which was 0.437, was also higher than that for the non-family companies (0.442). However, the results indicated that the non-family companies had the same LEV value on the average as the family companies, since

Table 2 Descriptive statistics

Variables	Full sample (n = 337)		Family companies (n = 254)		Non-family companies (n = 83)		*t*-statistics of mean difference
	Mean	SD	Mean	SD	Mean	SD	
TP	0.011	0.013	0.011	0.014	0.009	0.010	1.865*
SDRD	11.580	3.048	11.730	3.118	11.110	2.789	1.612
SIZE	22.606	1.513	22.700	1.578	22.330	1.263	2.154**
LEV	0.438	0.252	0.437	0.258	0.442	0.237	− 0.168
ROE	0.070	0.776	0.077	0.256	0.046	1.505	0.189
CAP	0.340	0.237	0.343	0.235	0.330	0.244	0.422
FAM	0.754	0.431					

Note One, two, and three asterisks indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively

the *t*-statistics for the mean differences were not statistically significant. Furthermore, the mean value of ROE for the full sample was 0.070, while the values for both family and non-family companies were 0.077 and 0.046, respectively. Furthermore, the average value of CAP for the full sample was 0.034, while the mean values for both family and non-family companies were 0.343 and 0.330, respectively. Nevertheless, both family and non-family companies had the same level of CAP as the *t*-statistics for the mean difference was not statistically significant. The mean value of FAM for the full sample was 0.754.

The Pearson's product-moment correlation, as shown in Table 3, is computed to examine the correlation among all variables of this paper. The correlation between SDRD and TP, for example, was −0.121, which was statistically and negatively significant. It indicated that higher SDRD practices could lower the TP, representing the TP of companies is at a high level. The correlations among explanatory variables were between −0.025 and 0.369. For instance, the correlations between CAP and other explanatory variables were not statistically significant. The correlations of FAM and ROE variables with others were also not statistically significant. Overall, the correlation among independent variables was also moderately low (below ± 0.7), indicating that there is no any multicollinearity problem [17]. This paper further considers whether data has the normal distribution or not. According to Berenson et al.'s [2] statement, the sampling distribution becomes almost normal, regardless of the shape of population since the sample size is large enough (which is greater than 30 observations). The data of this paper, therefore, has the normal distribution.

As previously mentioned, this paper uses the multiple regression analysis technique for analysing data and tests for the heteroscedasticity problem. According to the White's [37] test, all models, including the full sample and both family and non-family companies, did not have the problem of heteroscedasticity. The problem of autocorrelation also does not exist in all the models, as the Durbin–Watson [10, 11] statistics are between 1.50 and 2.50 [16]. The values of variance inflation factor (hereafter called VIF) in Table 4 for all models are further less than 10, indicating that the multicollinearity does not exist [15]. Overall, the results in Table 4 show

Table 3 Pearson correlation matrix

Variables	TP	SDRD	Size	LEV	ROE	CAP	FAM
TP	1.000	− 0.121**	0.026	0.197***	− 0.128**	0.060	− 0.101
SDRD		1.000	0.369***	0.034	0.079	0.043	0.088
SIZE			1.000	0.335***	0.068	− 0.015	0.105
LEV				1.000	− 0.005	− 0.025	− 0.009
ROE					1.000	− 0.018	0.017
CAP						1.000	0.023
FAM							1.000

Note One, two, and three asterisks indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively

Table 4 Effects of sustainable development report disclosure and family control on the tax planning

Independent variables	Full sample (n = 337)			Family companies (n = 254)			Non-family companies (n = 83)		
	β	t-statistics	VIF	β	t-statistics	VIF	β	t-statistics	VIF
Intercept	− 0.013	− 1.025		− 0.041	− 3.151***		0.048	2.416**	
SDRD	− 0.000	− 1.992**	1.204	− 0.000	− 1.657*	1.339	− 0.000	− 0.652	1.203
SIZE	0.000	0.260	1.501	0.002	2.494**	1.718	− 0.003	− 3.235***	1.266
LEV	0.010	3.418***	1.172	0.003	0.984	1.279	0.016	3.347***	1.239
ROE	− 0.002	− 2.183**	1.013	− 0.022	− 7.055***	1.129	− 0.000	− 0.460	1.067
CAP	0.006	1.910*	1.197	0.007	2.151**	1.176	0.003	0.590	1.352
FAM	− 0.003	− 1.855*	1.063						
IND1	− 0.001	− 0.395	2.012	− 0.003	− 0.974	2.085	0.006	1.103	1.888
IND2	0.003	1.012	1.855	0.002	0.607	2.032	− 0.001	− 0.159	1.256
IND3	− 0.002	− 0.870	2.642	− 0.002	− 0.711	2.702	0.001	0.179	2.509
IND4	0.000	0.160	2.694	− 0.002	− 0.555	2.506	0.006	1.572	3.446
IND5	0.000	0.010	1.732	− 0.003	− 0.744	1.747	0.003	0.635	1.745
IND6	− 0.002	− 0.955	2.824	− 0.006	− 1.930*	2.705	0.001	0.263	3.383
Durbin–Watson	1.813			2.080			2.122		
F-statistics	2.964***			6.904***			2.031**		
R^2	0.099			0.239			0.239		
Adjusted R^2	0.066			0.204			0.121		

Note One, two, and three asterisks indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively

that the coefficient for SDRD and TP was statistically significant at the 0.05 level. It indicated that companies with good SDRD practices could decrease the TP, representing the TP of companies is at a high level. These findings are consistent with prior studies, including Hoi et al. [18], Sari and Tjen [30], and Chen [5]. When dividing the data into family and non-family companies, the result indicated that there was no significant relationship between SDRD and TP for the non-family companies. On the other hand, the relationship between SDRD and TP for the family companies was negatively and statistically significant at the 0.10 level. This finding is consistent with previous papers, including Chen et al. [4], Landry et al. [21], and Thanjunpong and Bangmek [36]. Moreover, sustainable development is related to higher disclosure and transparency, higher accounting quality, and audit quality. Therefore, high SDRD companies might be able to decrease their TP, resulting in the TP of companies at a high level.

For the control variables, the relationship between SIZE and TP was not significant for full sample, as presented in Table 4. When dividing full sample into two sub-groups (family and non-family companies), the results show that the relationship between SIZE and TP was positively and statistically significant at the 0.05 level for the family companies, while the relationship was negatively and statistically significant at the 0.01 level for the non-family companies. Similarly, the effects of LEV on the TP for the full sample and the non-family companies were positively and statistically significant at the 0.01 level. Moreover, the effects of ROE on the TP for the full sample and the family companies were negatively significant at the 0.01 level. In addition, the effects of CAP on the TP for the full sample and the family companies were positively and statistically significant, while the effect of CAP on the TP for the non-family companies was not statistically significant. Finally, the relationship between FAM and TP was negatively statistically significant at the 0.10 level. This relationship is consistent with prior studies, including Chen et al. [4], Landry et al. [21], and Thanjunpong and Bangmek [36]. The relationship between IND and TP, however, was not statistically significant for the full sample.

This paper further does an additional analysis by excluding the industry variable, which is a dummy variable, for a robustness investigation. All models are also tested for the problems of autocorrelation and multicollinearity. The results in Table 5 show that the Durbin–Watson [10, 11] statistics are between 1.50 and 2.50, resulting that the autocorrelation problem does not exist [16]. Each variable of all models also has the VIF value less than 10, indicating that the problem of multicollinearity does not exist [15]. The results in Table 5 show that the coefficient for SDRD and TP was negatively and statistically significant at the 0.05 level. It indicates that the TP could be reduced, when companies have a good SDRD implementation. This implies that companies have a high level of TP. These findings are consistent with previous papers, including Hoi et al. [18], Sari and Tjen [30], and Chen [5]. Once dividing the full sample into family and non-family companies, the result indicated that relationship between SDRD and TP for the non-family companies was not statistically significant. However, the relationship between SDRD and TP for the family companies was negatively and statistically

Table 5 Effects of sustainable development report disclosure and family control on the tax planning (robustness analysis)

Independent variables	Full sample (n = 337)			Family companies (n = 254)			Non-family companies (n = 83)		
	β	t-statistics	VIF	β	t-statistics	VIF	β	t-statistics	VIF
Intercept	− 0.013	− 1.269		− 0.040	− 3.339***		0.035	1.863**	
SDRD	− 0.001	− 2.141**	1.178	− 0.001	− 2.044**	1.295	− 0.000	− 0.471	1.135
SIZE	0.000	0.418	1.329	0.001	2.502**	1.480	− 0.002	− 2.614**	1.115
LEV	0.010	3.146***	1.141	0.003	0.894	1.242	0.017	3.594***	1.183
ROE	− 0.002	− 2.188**	1.009	− 0.021	− 6.821***	1.084	− 0.000	− 0.524	1.011
CAP	0.004	1.328	1.004	0.005	1.663*	1.005	0.000	0.028	1.047
FAM	− 0.003			− 1.713*			1.016		
Durbin–Watson	1.780			2.005			2.016		
F-statistics	4.883***			13.303***			3.412***		
R^2	0.082			0.211			0.181		
Adjusted R^2	0.065			0.196			0.128		

Note One, two, and three asterisks indicate statistical significance at the 0.10, 0.05, and 0.01 levels, respectively

significant at the 0.05 level. These results were similar to the above results, which were reported in Table 4, and prior studies. Consequently, sustainable development is associated with higher disclosure and transparency, higher accounting quality, and audit quality. As a result, high SDRD companies might be able to decrease their TP, representing the TP of companies is at a high level.

5 Conclusion

This paper investigates the effect of SDRD on the TP of listed companies, excluding financial sectors, in the SET. The sample size consists of 337 companies, based only on the data for the year 2016. The questionnaires of GRI are used for evaluation of SDRD. The TP is also measured by the ratio of total tax expenses to total assets. Overall, this paper finds that the SDRD had a significant and negative effect on the TP. This indicates that companies could have a higher level of TP, when they have higher SDRD implementation. Regarding some control variables, the financial leverage and capital intensity variables had a positive effect on the TP. However, the profitability and family control variables had a significantly negative impact on the TP. In addition, the relationship between SDRD and TP was significantly negative for the family companies. The relationship was, however, weakly insignificant for the non-family companies. As a result, the results are useful to tax officers for making a decision of initial audit. They, for instance, should pay attention on companies that have low SDRD implementation. Moreover, the results are useful to the SET and Securities and Exchange Commission (hereafter called SEC). They could further make some decisions about adjusting some rules and regulations or give some incentives to encourage the Thai-listed companies to perform better SDRD practices. For example, market regulators, including SET and SEC, should focus on family companies and encourage them to improve their SDRD implementation by giving them some incentives, and the results indicate that the relationship between SDRD and TP for the family companies is negatively significant.

References

1. Bauweraerts, J., Vandernoot, J.: Are family firms more tax aggressive than non-family firms? Empirical evidence from Belgium. Int. J. Manag. **30**(4), 235–243 (2013)
2. Berenson, M.L., Levine, D.M., Krehbiel, T.C.: Basic Business Statistics: Concept and Application, 12th edn. Pearson Education, New Jersey (2012)
3. Boussaidi, A., Hamed, M.S.: The impact of governance mechanisms on tax aggressiveness: empirical evidence from Tunisian context. J. Asian Bus. Strategy **5**(1), 1–12 (2015)
4. Chen, S., Chen, X., Cheng, Q., Shevlin, T.: Are family firms more tax aggressive than non-family firms? J. Financ. Econ. **95**(1), 41–61 (2010)

5. Chen, X.: Corporate social responsibility disclosure, political connection and tax aggressiveness: evidence from China's capital markets. Open J. Bus. Manag. **6**(1), 151–164 (2018)
6. Chen, X., Hu, N., Wang, X., Tang, X.: Tax avoidance and firm value: evidence from China. Nankai Bus. Rev. Int. **5**(1), 25–42 (2014)
7. Chienwittayakun, J., Mankin, D.: Strategic management planning process (SMPP) as an organization development intervention (ODI) to align values, goals and objectives and improve employee teamwork, engagement and performance: a case study of a family-owned business in Thailand. ABAC J. Vis. Action Outcome **2**(1), 74–92 (2015)
8. Corporate Governance Center: Good governance assessment of listed companies. The Stock Exchange of Thailand, Bangkok (2003)
9. Davis, A.K., Guenther, D.A., Krull, L.K., Williams, B.M.: Do socially responsible firms pay more taxes? Account. Rev. **91**(1), 47–68 (2016)
10. Durbin, J., Watson, G.S.: Testing for serial correlation in least squares regression: I. Biometrika **37**(3/4), 409–428 (1950)
11. Durbin, J., Watson, G.S.: Testing for serial correlation in least squares regression. II. Biometrika **38**(1/2), 159–177 (1951)
12. Dyreng, S.D., Hanlon, M., Maydew, E.L.: Long-run corporate tax avoidance. Account. Rev. **83**(1), 61–82 (2008)
13. Gaaya, S., Lakhal, N., Lakhal, F.: Does family ownership reduce corporate tax avoidance? The moderating effect of audit quality. Manag. Audit. J. **32**(7), 731–744 (2017)
14. Board, Global Sustainability Standards: G4 Sustainability Reporting Guidelines. Global Reporting Initiative, Amsterdam (2013)
15. Greene, W.H.: Econometric Analysis, 8th edn. Pearson, New York (2018)
16. Gujarati, D.N.: Basic Econometrics, 4th edn. McGraw-Hill, New York (2004)
17. Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E.: Multivariate Data Analysis, 7th edn. Pearson Education, New Jersey (2010)
18. Hoi, C.K., Wu, Q., Zhang, H.: Is corporate social responsibility (CSR) associated with tax avoidance? Evidence from irresponsible CSR activities. Account. Rev. **88**(6), 2025–2059 (2013)
19. Hope, O.-K., Ma, M., Thomas, W.B.: Tax avoidance and geographic earnings disclosure. J. Account. Econ. **56**(2–3), 170–189 (2013)
20. Laguir, I., Staglianò, R., Elbaz, J.: Does corporate social responsibility affect corporate tax aggressiveness? J. Clean. Prod. **107**, 662–675 (2015)
21. Landry, S., Deslandes, M., Fortin, A.: Tax aggressiveness, corporate social responsibility, and ownership structure. J. Acc. Ethics Public Policy **14**(3), 611–645 (2013)
22. Lanis, R., Richardson, G.: Corporate social responsibility and tax aggressiveness: a test of legitimacy theory. Acc. Audit. Account. J. **26**(1), 75–100 (2012)
23. Lanis, R., Richardson, G.: Corporate social responsibility and tax aggressiveness: an empirical analysis. J. Account. Public Policy **31**(1), 86–108 (2012)
24. Lestari, N., Wardhani, R.: The effect of the tax planning to firm value with moderating board diversity. Int. J. Econ. Financ. Issues **5**(1S), 10–11 (2015)
25. Martinez, A.L., Ramalho, G.C.: Family firms and tax aggressiveness in Brazil. Int. Bus. Res. **7**(3), 129–136 (2014). https://doi.org/10.5539/ibr.v7n3p129
26. Masripah, M., Diyanty, V., Fitriasar, D.: Controlling shareholder and tax avoidance: family ownership and corporate governance. Int. Res. J. Bus. Stud. **8**(3), 167–180 (2015)
27. Nadiah, A.H., Syakirah, W.N.W.K., Mastora, Y., Rohayu, Y., Rozainun, A.A.: Corporate social responsibility (CSR) disclosure and its impacts towards corporate tax aggressiveness. J. Appl. Environ. Biol. Sci. **7**(5S), 10–15 (2017)
28. Preuss, A., Preuss, B.: Corporate tax payments and corporate social responsibility: complements or substitutes? Empirical evidence from Europe. Bus. Econ. J. **8**(4), 1–8 (2017)
29. Sandmo, A.: The theory of tax evasion: a retrospective view. Natl. Tax J. **58**(4), 643–663 (2005)
30. Sari, D., Tjen, C.: Corporate social responsibility disclosure, environmental performance, and tax aggressiveness. Int. Res. J. Bus. Stud. **9**(2), 93–104 (2016)

31. Social Responsibility Center: Sustainable Development Meaning. Wanida Printing Limited Partnership, Nonthaburi (2014)
32. Suehiro, A., Wailerdsak, N.: Family business in Thailand its management, governance, and future challenges. ASEAN Econ. Bull. **21**(1), 81–93 (2004)
33. Susanti, M.: Corporate social responsibility, size and tax avoidance. Int. J. Econ. Perspect. **11** (1), 1639–1650 (2017)
34. Tang, T., Firth, M.: Can book–tax differences capture earnings management and tax Management? Empirical evidence from China. Int. J. Acc. **46**(2), 175–204 (2011)
35. Tantiyavarong, T.: A Study of the Determinants of Tax Planning and the Association between Tax Planning and Firm Value: An Empirical Evidence of Thailand. Chulalongkorn University, Bangkok (2009)
36. Thanjunpong, S., Bangmek, R.: The influence of board of directors, audit committee and ownership structure impact on tax planning: an empirical evidence of Thailand. J. Acc. Prof. **13**(37), 29–44 (2017)
37. White, H.: A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica **48**(4), 817–838 (1980)
38. Ying, T., Wright, B., Huang, W.: Ownership structure and tax aggressiveness of Chinese listed companies. Int. J. Acc. Inf. Manag. **25**(3), 313–332 (2017)
39. Zhang, C., Cheong, K.C., Rasiah, R.: Corporate tax avoidance and performance: evidence from China's listed companies. Inst. Econ. **8**(3), 61–83 (2016)

Clustering and Labeling Auction Fraud Data

Ahmad Alzahrani and Samira Sadaoui

Abstract Although shill bidding is a common fraud in online auctions, it is however very tough to detect because there is no obvious evidence of it happening. There are limited studies on SB classification because training data are difficult to produce. In this study, we build a high-quality labeled shill bidding dataset based on recently scraped auctions from eBay. Labeling shill biding instances with multidimensional features is a tedious task but critical for developing efficient classification models. For this purpose, we introduce a new approach to effectively label shill bidding data with the help of the robust hierarchical clustering technique CURE. As illustrated in the experiments, our approach returns remarkable results.

Keywords Auction fraud · Shill bidding · Hierarchical clustering · CURE · Silhouette · Data labeling

1 Introduction

In the last three decades, we witnessed a significant increase in exchanging goods and services over the Web. According to the World Trade Organization, the worldwide merchandise during the period 1995–2015 was over 18 billion.[1] Online auctions are a very profitable e-commerce application. For instance, in 2017, eBay claimed that the net revenue attained 9.7 billion US dollars, and the number of active users hits 170 million.[2] Regardless of their popularity, e-auctions remain very vulnerable to cybercrimes. The high anonymity of users, low fees of auction services, and flexibility

[1] https://www.wto.org/english/res_e/statis_e/its2015_e/its2015_e.pdf.
[2] https://www.statista.com.

A. Alzahrani (✉) · S. Sadaoui
Computer Science Department, University of Regina, Regina, SK, Canada
e-mail: alzah234@uregina.ca

S. Sadaoui
e-mail: sadaouis@uregina.ca

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_20

of bidding make auctions a great incubator for fraudulent activities. The Internet Crime Complain Center announced that auction fraud is one of the top cybercrimes [2]. As an example, the complaints about auction fraud in only three states, California, Florida, and New York, reached 7,448 in 2016 [2]. Malicious moneymakers can commit three types of fraud, which are pre-auction fraud, such as auctioning of black market merchandise, in-auction fraud that occurs during the bidding time, such as Shill Bidding (SB), and post-auction fraud, such as fees stacking. Our primary focus is on the SB fraud whose goal is to increase the profits of sellers by placing many bids through fake accounts and colluding with other users. SB does not leave any obvious evidence, unlike the two other auction fraud. Indeed, buyers are not even aware that they have been overcharged.

Identifying relevant SB strategies, determining robust SB metrics, crawling and preprocessing commercial auction data, and evaluating the SB metrics against the extracted data make the study of SB fraud very challenging as demonstrated in our previous technical paper [1]. In addition, labeling SB instances with multidimensional features is a critical phase for the classification models. In the literature, labeling training data is usually done manually by the domain experts, which is quite a laborious task and prone to errors. Due to the lack of labeled SB training datasets, the prime contribution of this paper is to produce high-quality labeled SB data based on commercial auction transactions that we extracted from eBay and preprocessed [1]. As illustrated in Fig. 1, we introduce a new approach to effectively label SB data with the help of data clustering. First, we split the SB dataset into several subsets according to the different bidding durations of the extracted auctions. Second, we efficiently partition each SB subset into clusters of users with similar bidding behavior. Last, we apply a systematic labeling method to each cluster to classify bidders into normal or suspicious.

Hierarchical clustering is significantly preferable over partitioning clustering because it provides clusters with a higher quality [10]. This type of data clustering has been utilized successfully in numerous fraud studies [8, 10]. In fact, we employ the Clustering Using REpresentatives (CURE) technique to produce the best differentiation between normal and suspicious activities. CURE [11] has proved over the years to be a highly efficient clustering method in terms of eliminating outliers and producing high-quality clusters, especially for large-scale training datasets. The

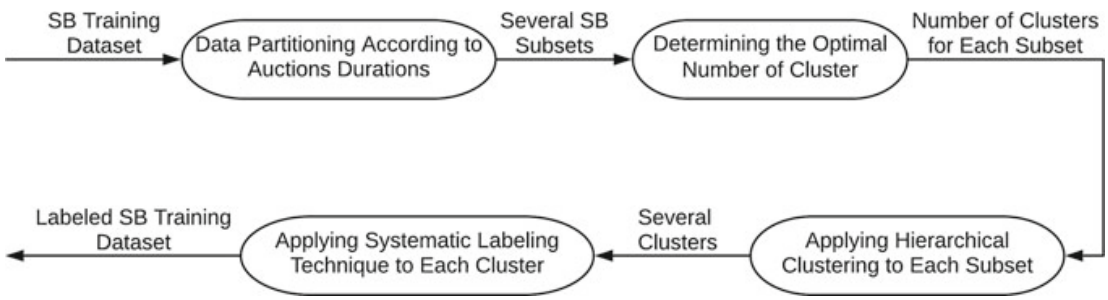


Fig. 1 The labeling process of shill bidding training data

labeled SB dataset that we produced can be utilized by the state-of-the-art supervised classification methods. Furthermore, the accuracy of new predictive models can also be tested using our SB training dataset.

The rest of the paper is organized as follows. Section 2 discusses related work on data clustering in the context of online auctions. Section 3 describes the SB patterns used in this research as well as the SB data produced from the commercial website eBay. Section 4 explains how to apply the hierarchical clustering CURE according to different bidding durations. Section 5 exposes a new approach to label the clusters of bidders. Finally, Sect. 6 summarizes the results of this study and highlights important research directions as well.

2 Related Work

Many researchers utilized data clustering to examine auctions from different angles, such as studying the dynamics of auction prices and bidder behavior. For instance, [13] introduced an approach to model and analyze the price formation as well as its dynamics to characterize the heterogeneity of the price formation process. The proposed functional objects represent the price process by accommodating the structure format of bidding data on eBay. Then, the curve clustering is used to partition auctions by grouping similar price profiles. Finally, differential equations are used to specify the price of each group.

Another work [7] measured the similarity of bidder behavior using specific attributes, such as bidder feedback rating, average increment difference, and number of bids. Then, a centroid-based hierarchical clustering approach is presented to group similar bidders. Each produced cluster is then labeled manually according to the overall bidder behavior in that cluster.

The study [14] suggested an SB detection model utilizing k-mean clustering technique. The latter groups similar buyers in one class to differentiate between general buyers and shill bidders. There are four features that represent a buyer: "how long the buyer has been in the auction," "the times of buyer bids," "the average response time of the buyer," and "the absolute average discrepancy of buyer bids." Based on these features, k-mean classifies bidders into one of the two clusters: general buyers and shill bidders. However, we believe the second and third features do not really reflect SB since a buyer might be very interested in winning the auction. Besides, there are stronger patterns that highly identify SB. Since SB behavior is somehow similar to real bidding [5], there is a possibility that some SB samples fell in the normal class.

In [4], the authors proposed a two-step clustering model to recognize bidding strategies. The hierarchical clustering is the first step to produce a dendrogram and an agglomeration schedule table to find the best number of clusters. The next step employs k-mean clustering to provide more details about the bidders' strategies in each cluster. The experiment was operated on an outdated data (2003) collected

from Taobao.com. According to the agglomeration coefficients, the optimal number of clusters is three: "early bidding strategy", "snipe bidding strategy," and the third one groups bidders that enter the auction early and remain for a long time. The first cluster has shown low values for the given features, which indicates that the bidders' strategy is to enter and exit auctions early and participate infrequently. The second cluster has displayed high values for some features and low values for others. This illustrates that bidders enter and exit auctions late and rarely participate. The final cluster has administered the bidders' strategy where bidders enter early and stay for a long time and highly participate in the auctions.

Zhang et al. [20] proposed a model based on hedonic regression and Fuzzy Logic Expert System (FLES) to analyze bidder behavior. The hedonic regression is used to select key variables that are passed to FLES to produce a knowledge base about the relationships between variables, like auction characteristics. Since the examined data have no relational information, k-mean is employed to obtain the minimum squared-error clusters. So, each training sample is classified to low, medium, or high membership degree. The issue here is that the study is based on an outdated dataset (2004). Also, there is a potential for fabricating feedback ratings conducted between shill bidders and fraudulent bidder rings [7].

More recently, [12] applied k-mean clustering to categorize bidders' habits. The observations are obtained by the k-mean and then passed to the Baum–Welch algorithm and Hidden Markov model. Three main clusters were suggested according to the values of the given features, which are low, medium, and high cluster values. A bidder habit with values beyond these clusters values is considered as a fraudster. The experiment showed that only two simple features were given to identify the clusters: the number of auctions that a bidder participated in and the number of submitted bids by that bidder. Thus, if more features were considered to define the clusters, then the samples' distribution on each cluster might be changed. As a result, this may influence the outcomes of the detection model. Also, the clustering is based on a dataset that is not adequately described, and only ten samples were used for explaining the results.

Lastly, [10] applied a hierarchical clustering to group users with similar bidding behavior. The centroid linkage is used as the similarity measure. The described SB patterns were computed for all the bidders of each of the generated clusters. Then, the authors introduced a semi-automatic approach to label each cluster according to the general behavior of users in that cluster and the weights of the fraud patterns.

3 Shill Bidding Overview

SB is a well-known auction fraud, and yet it is the most difficult to detect since it behaves similarly to normal bidding [5, 8]. The aim of this fraud is to increase the price or desirability of the auctioned product through imitation accounts and collusion

with other users. In other words, shill bidders do not tend to win the auction but to increase the revenue of the seller. SB leads buyers to overpay for the items, and for high priced items, buyers will lose a substantial amount of money. As mentioned in [5], excessive SB could lead to market failure. Thus, e-auctions may lose their credibility [5]. In fact, several sellers and their accomplices have been prosecuted due to SB activities, including

- In 2007, a jewelery seller was accused of conducting SB fraud on eBay and had to pay $400,000 for a settlement. Also, he and his employees were prevented from engaging in any online auctioning activities for 4 years.[3]
- In 2010, a seller faced a £50,000 fine after being found outbidding himself on eBay. He claimed that "*eBay let me open up the second account and I gave all my personal details and home address to do so.*"[4]
- In 2012, the online auction Trade Me had to pay $70,000 for each victim after the investigation discovered SB fraud conducted by a motor vehicle trader in Auckland. The fraud was carried out for 1 year and caused a significant loss for the victims. Trade Me blocked this trader from using their site, and referred the case to the Commerce Commission for a further investigation.[5]
- In 2014, a lawsuit was filled against Auction.com by VRG in California claiming that the website allowed SB. The bid of $5.4 million should have secured the property as the plaintiff declared, and yet the winning price was 2 million more. Auction.com was accused of helping the property loan holder, which is not fair for genuine bidders. The California state passed a law on July 1, 2015, which requests the property auctioneers to reveal bids they submit on a seller's behalf.[6] The spokeswoman for the California Association of Realtors said "*To the best of our knowledge, we are the only state to pass this sort of legislation, even though we believe shill bidding to be prevalent all over the country.*"

By examining throughly the literature on the SB strategies [6, 8, 17], we compiled in Table 1, the most relevant SB patterns. Each pattern, which is a training feature, represents a unique aspect of the bidding behavior in auctions. The feature uniqueness will lead to an improved predictive performance.

[3] https://www.nytimes.com/2007/06/09/business/09auction.html.

[4] http://www.dailymail.co.uk/news/article-1267410/Ebay-seller-faces-fine-bidding-items-raise-prices.html.

[5] https://www.trademe.co.nz/trust-safety/2012/9/29/shill-bidding.

[6] https://nypost.com/2014/12/25/lawsuit-targets-googles-auction-com.

Table 1 SB patterns and their characteristics

Name	Definition	Category	Source	Weight
Bidder tendency (BT)	Engages exclusively with few sellers instead of a diversified lot	Bidder	User history	0.5
Bidding ratio (BR)	Participates more frequently to raise the auction price	Bid	Bidding period	0.7
Successive outbidding (SO)	Successively outbids himself even though he is the current winner	Bid	Bidding period	0.7
Last bidding (LB)	Becomes inactive at the last stage to avoid winning	Bid	Last bidding stage	0.5
Early bidding (EB)	Tends to bid pretty early in the auction to get users attention	Bid	Early bidding stage	0.3
Winning ratio (WR)	Participates a lot in many auctions but rarely wins any auctions	Bidder	Bidder history	0.7
Auction bids (AB)	Tends to have a much higher number of bids than the average of bids in auctions selling the same product	Auction	Auction history	0.3
Auction starting price (ASP)	Offers a small starting price to attract genuine bidders	Auction	Auction history	0.3

4 Production of Shill Bidding Data from Online Auctions

4.1 Auction Data Extraction and Preprocessing

To obtain a reliable SB training dataset, it must be built from actual auction data. Nevertheless, producing high-quality auction data is itself a burdensome operation due to the difficulty of collecting data from auction sites on one hand, and the challenging task of preprocessing the raw data on the other hand. The latter consumes a significant time and effort, around 60–80% of the entire workload [15]. In our technical study [1], we employed the professional scraper Octoparse[7] to collect a large number of auctions for one of the most popular products on eBay. The extracted dataset contains all the information related to auctions, bids and bidders. We crawled completed auctions of the iPhone 7 for 3 months (March to June 2017). We chose iPhone 7 because it may have attracted malicious moneymakers due to the following facts:

- Its auctions attracted a large number of bids and bidders.
- It has a good price range with the average of $578.64 (US currency). Indeed, there is a direct relationship between SB fraud and the auction price [5].
- The bidding duration varies between 1 (20.57%), 3 (23.2%), 5 (16.23%), 7 (38.3%), and 10 (1.7%) days. In long duration, a dishonest bidder may easily mimic usual bidding behavior [5]. However, as claimed in [3], fraudulent sellers may receive a positive rating in short duration. Thus, we considered both durations.

[7] https://www.octoparse.com.

Table 2 Preprocessed auctions of iPhone 7

No. of auctions	807
No. of records	15145
No. of bidder IDs	1054
No. of seller IDs	647
Avg. winning price	$578.64
Avg. bidding duration	7
No. of attributes	12

Table 2 presents the statistics after preprocessing the scraped auction data. This operation was very time consuming as it required several manual operations [1]: (1) removing redundant and inconsistent records, and also records with missing bidder IDs; (2) merging several attributes into a single one; (3) converting the format of several attributes into a proper one; (4) assigning an ID to the auctions. For example, the two attributes date and time in each auction are converted into seconds; as an example, 1-day and 10-day durations are converted into 86,400 and 864,000 seconds, respectively.

4.2 SB Data Production

The algorithms to measure the SB patterns are presented in [1, 17]. Each metric is scaled to the range of [0, 1]; a high value indicates a suspicious bidding behavior. We evaluated each metric against each bidder in each of the 807 auctions [1]. Therefore, we obtained a SB training dataset with a total of 6321 samples. A sample is described as a vector of 10 elements: the eight SB features along with Auction, and Bidder identification numbers. Once labeled, an instance will denote the conduct (normal or suspicious) of a bidder in a certain auction.

5 Hierarchical Clustering of Shill Bidding Data

Since the produced SB data are not labeled, data clustering, an unsupervised learning method, can be utilized to facilitate the labeling operation. Clustering is the process of isolating instances into K groups w.r.t. their similarities. The clustering techniques fall into one of the following categories: (1) Partitioning-based, such as K-medoids and K-means; (2) Hierarchical-based, such as BIRCH, GRIDCLUST and CURE; (3) Density-based, such as DBSCAN and DBCLASD; (4) Grid-based, such as STING and CLIQUE. In our work, we select agglomerative (bottom-top) hierarchical clustering where instances are arranged in the form of a tree structure using a proximity matrix.

5.1 CURE Overview

Among the hierarchical clustering methods, we choose CURE because it is highly performant in handling large-scale multidimensional datasets, determines non-spherical shapes of the clusters, and efficiently eliminates outliers [11]. Random sampling and partitioning techniques are utilized to handle the large-scale problem and to speed up the clustering operation. Each instance is first considered as an individual cluster, and then the cluster with the closest distance/similarity is merged into it in order to form a new cluster [11]. Two novel strategies have been introduced in CURE:

- **Representative Points (RPs)**, which are selected data points that define the cluster boundary. Instead of using a centroid, clusters are identified by a fixed number of RPs that are well dispersed. Clusters with the closest RPs are merged into one cluster. The multiplicity of RPs allows CURE to obtain arbitrary clustering shapes.
- **Constant shrinking factor** (α), which is utilized to shrink the distance of RPs toward the centroid of the cluster. This factor reduces noise and outliers.

The worst-case computational complexity of CURE is estimated to $O(N^2 \log N)$, which is high when N is large (N is the number of instances) [18]. Since the SB data clustering is an offline operation, so the running time is not an issue. The only disadvantage of CURE is that the two parameters RP and α have to be set up by users.

To run the experiments, we utilize the Anaconda Navigator environment for running Python 3, and incorporate the CURE program developed by Freddy Stein and Zach Levonian. CURE code (in Python) is available at GitHub.com.[8]

5.2 SB Data Preparation

Since the bidding duration is used as a denominator for the two patterns Early Bidding and Last Bidding, the large gap between different durations greatly affects the computation results. The computed value of the fraud pattern for 10 days is far smaller than for 1 day. So, before applying CURE, we first partition the SB dataset into five subsets according to the five durations (1, 3, 5, 7, and 10 days) as presented in Table 3.

5.3 Optimal Number of Clusters

It is always difficult to decide about the optimal number of clusters of a training dataset. Besides, normal and shill bidder behavior is somehow similar. Thus, deter-

[8] https://github.com/levoniaz/python-cure-implementation/blob/master/cure.py.

Table 3 SB dataset partitioning according to bidding durations

Subset	1 Day	3 Days	5 Days	7 Days	10 Days
No. of auctions	166	187	131	309	14
No. of Instances	1289	1408	1060	2427	137

mining the best number of clusters is an essential step to achieve a better interpretation for classifying similar SB instances [19]. There are several methods to address this problem, such as Elbow, Dendrogram, the Rule of Thumb and Silhouette. In our study, we employ the Silhouette method where each group is represented as a silhouette based on the separation between instances and the cluster's tightness [16]. The construction of a silhouette requires the clustering technique to generate the partitions and also to collect all approximates between instances [16]. K-mean clustering algorithm has been successfully utilized for this task due to its simplicity and effectiveness [19]. Consequently, we apply K-mean to estimate the number of clusters for each of the five SB subsets.

Next, for each subset, we examine the silhouette scores of 19 clusters (2–20 clusters), and choose the best number based on the best silhouette score. We have noticed that once the peek (denoting the optimal number) of the silhouette score is reached on a certain number of clusters, the silhouette score gradually decreases with the increasing of the number of clusters. In Fig. 2, we give an example for the 7-day bidding duration for which the optimal number of clusters is eight since the highest silhouette score (0.4669) is obtained on that number. In Table 4, we expose the best number of clusters for each of the five SB subsets. The total number of produced clusters is 29.

5.4 Cluster Generation

CURE has three parameters that need to be setup: representative points (RPs), shrinking factor (α) and an optimal number of clusters. Based on the results of silhouette, we have obtained the optimal number of clusters for each of the five SB subsets. The two parameters RPs and α are defined by selecting the configuration that provides the best instance distribution among the specified clusters. Thus, CURE is applied with different values of RPs and α starting from the default values (5 for RPs and 0.1 for α). We throughly conducted trial-and-error experiments for all the clusters (29 clusters in total) of the five SB subsets to determine the best values of the parameters (Table 4). The best parameters' values are selected based on the best distribution of a subset population between the defined clusters. As an example, in Table 5, we present the results for the eight clusters that we have generated previously for the 7-day bidding duration subset. The best value configuration is shown in bold. Each

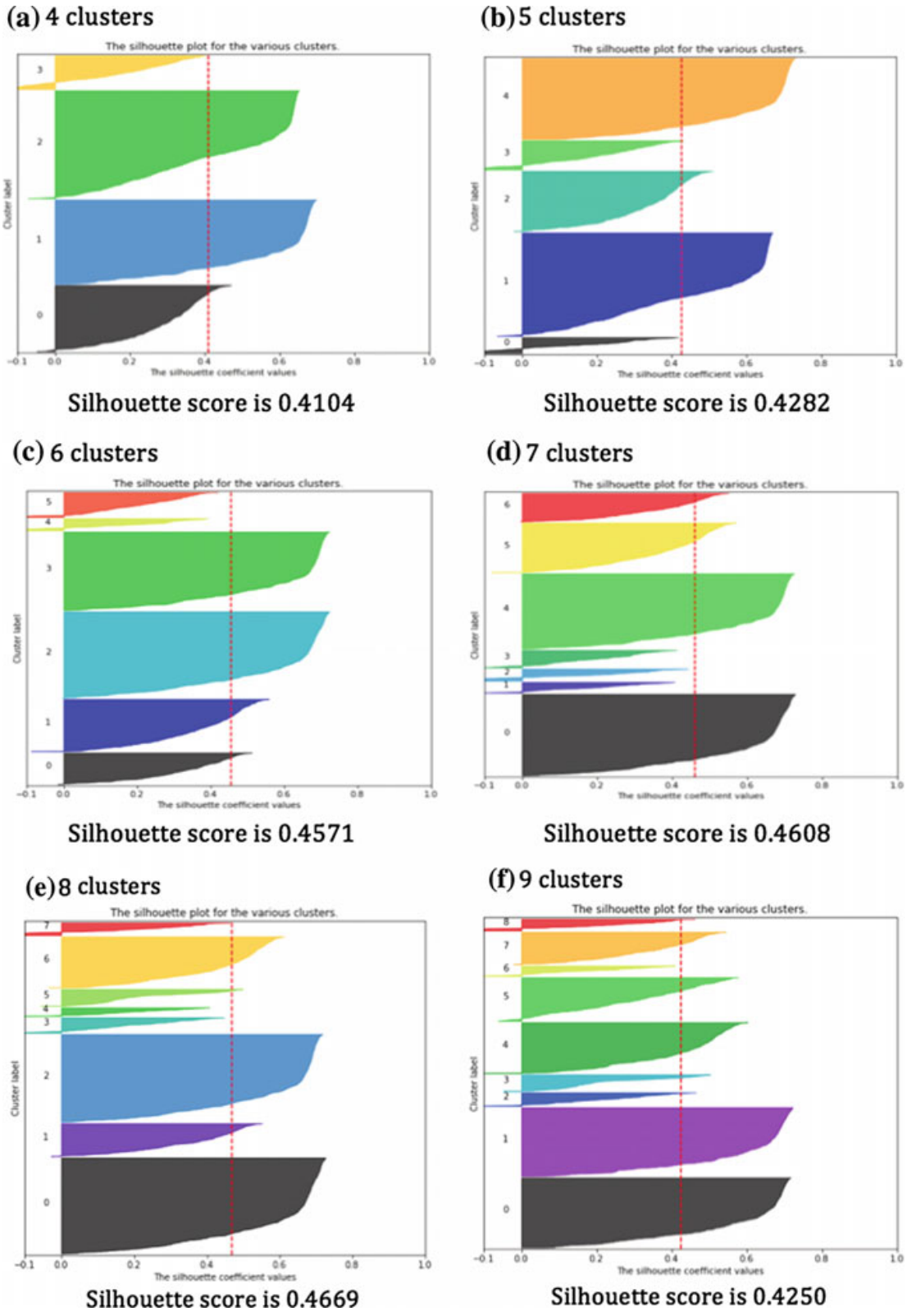


Fig. 2 Optimal number of clusters for 7-day bidding duration. Silhouette score is examined 19 times. We show the top Silhouette scores

Table 4 Optimal number of clusters and optimal CURE parameters

SB subset	No. of samples	No. of clusters	Silhouette score	RPs	α
1 Day	1289	7	0.4597	5	0.05
3 Days	1408	7	0.4672	5	0.01
5 Days	1060	5	0.4758	5	0.05
7 Days	2427	8	0.4669	10	0.001
10 Days	137	2	0.5549	5	0.1

Table 5 CURE clustering for 7-day bidding duration (8 clusters)

RP	α	Cl. # 1	Cl. # 2	Cl. # 3	Cl. # 4	Cl. # 5	Cl. # 6	Cl. # 7	Cl. # 8
5	0.1	136	1438	1	2	657	190	2	1
5	0.05	657	328	2	1408	1	28	2	1
5	0.01	1438	640	1	1	17	1	1	328
5	0.001	21	166	2	1	2	25	2209	1
10	0.1	2	1	657	1410	22	8	137	190
10	0.05	2	133	2	1	2	31	2066	190
10	0.01	1	1	1	1	135	31	2067	190
10	**0.001**	**189**	**654**	**1410**	**3**	**137**	**31**	**1**	**2**

cluster consists of users with similar bidding behavior. As we can see, the clusters 4, 7, and 8 have very few bidders, which are most probably outliers i.e., suspicious.

6 Labeling Shill Bidding Data

In Algorithm 1, we show the steps to label the bidders of a given cluster. A cluster belongs to a certain SB subset. As shown in Table 6, for each subset, we first compute the mean and STandard Deviation (STD) of each fraud pattern for all the instances in that subset. Then, we compute the average of the means (Avg. Means) and average of the STDs (Avg. STDs) of all the patterns for that subset. We consider the value of $(Avg.Means + \frac{1}{2}Avg.STDs)$ since it produces the best decision line that separates between normal and suspicious instances as depicted in Fig. 3. Then, we calculate the average of the means of all the patterns for the cluster. So, if the average mean of the cluster is greater than the decision line of the subset, then instances are labeled as suspicious (1) in that cluster, otherwise, they are labeled normal (0).

Table 6 Characteristics of SB subsets

Subset	1 Day	3 Days	5 Days	7 Days	10 Days
Mean of each pattern per subset					
BT	0.1434	0.1394	0.1419	0.1455	0.1162
BR	0.1287	0.1328	0.1235	0.1273	0.1021
SO	0.0996	0.1047	0.0872	0.1149	0.0620
LB	0.4624	0.4511	0.4676	0.4678	0.4746
EB	0.4314	0.4192	0.4318	0.4348	0.4575
WR	0.3812	0.3718	0.3810	0.3533	0.3496
AB	0.2120	0.1936	0.2403	0.2567	0.2926
ASP	0.5007	0.4301	0.4478	0.4801	0.7123
Avg. means	0.2949	0.2802	0.2901	0.2975	0.3208
STD of each pattern per subset					
BT	0.1973	0.1884	0.1984	0.2019	0.1811
BR	0.1246	0.1330	0.1243	0.1377	0.1165
SO	0.2764	0.2811	0.2583	0.2917	0.2215
LB	0.3773	0.3753	0.3917	0.3783	0.3931
EB	0.3775	0.3742	0.3921	0.3802	0.3968
WR	0.4356	0.4373	0.4402	0.4345	0.4398
AB	0.2323	0.2426	0.2646	0.2658	0.2575
ASP	0.4931	0.4831	0.4863	0.4908	0.4510
Avg. STDs	0.3142	0.3143	0.3194	0.3226	0.3071

Algorithm 1 : Labeling Bidders in a Cluster

Require: AvgMeans and AvgSTDs of the corresponding SB subset
1: Compute $MeanCluster$
2: **if** $(MeanCluster \geq (AvgMeans + \frac{AvgSTDs}{2}))$ **then**
3: **for** x=1 to $NumberBiddersCluster$ **do**
4: $LabelBidder_x$ = 1 (Suspicious)
5: **end for**
6: **else**
7: **for** x=1 to $NumberBiddersCluster$ **do**
8: $LabelBidder_x$ = 0 (Normal)
9: **end for**
10: **end if**

To validate our approach, we choose randomly 5 auctions among the 7-day duration auctions (in total 309) and select randomly one bidder in each auction (Table 7). As we can observe from this table, the shill bidding instances were successfully labeled by our approach. For example, bidder "g***r" has four fraud patterns with very high values; among them, two have a high weight and one a medium weight. Therefore, the activity of this bidder in auction ID # 2370 is suspicious. On the other

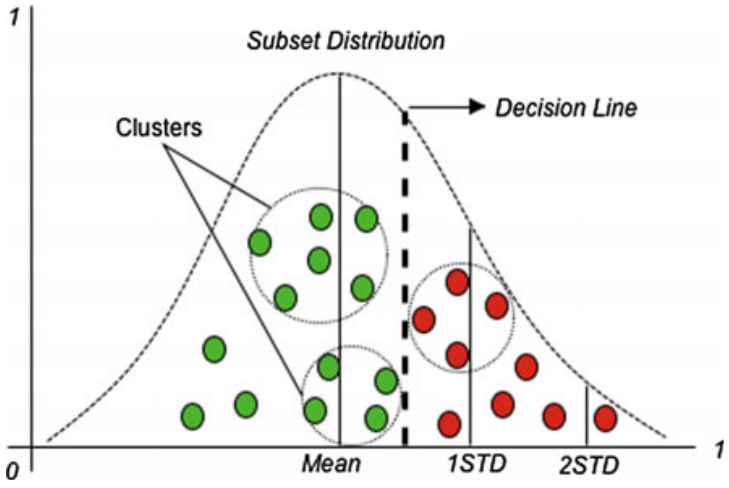


Fig. 3 Decision line of a subset and its labeled clusters

Table 7 SB instances and their labels

AuctionID	1009	900	2370	432	1370
BidderID	z***z	k***a	g***r	0***0	o***–
BT	0.75	0.4705	0.8333	0.5	0.04615
BR	0.3461	0.3076	0.2	0.3333	0.0857
SO	1	0	1	0	0.5
LB	0.5667	0.1909	0.0350	0.2199	0.2966
EB	0.5409	0.1909	0.0239	0.0043	0.2060
WR	0.75	0.4	1	0.5	0
AB	0	0	0.3333	0	0.0526
ASP	0	0	0.9935	0	0
Generated label	**1**	**0**	**1**	**0**	**0**

Table 8 Final results of instance labeling

SB subset	1 Day	3 Days	5 Days	7 Days	10 Days	Total
No. of normal instances	1135	1303	975	2098	135	5646
No. of suspicious instances	154	105	85	329	2	675

hand, the bidder "k***a" has all his fraud patterns with very low values; this indicates that this bidder behaved normally in the auction ID # 900. All these results are consistent with the labels produced by our approach.

Table 8 provides all the final results of the labeling task of our SB training dataset. There are 5646 instances categorized as normal and 675 instances as suspicious.

7 Conclusion and Future Work

There are limited classification studies on the SB fraud due to the difficulty of producing training data on one hand and labeling multidimensional instances on the other hand. Our aim in this paper is to effectively label SB instances based on the hierarchical clustering CURE that showed a remarkable capability for partitioning the online behavior of bidders. First, we divide the SB dataset into several subsets according to the different bidding durations of the auctions that we scraped from eBay. Then, we efficiently partition each SB subset into clusters of users with similar bidding behavior. At last, we apply a systematic labeling approach to each cluster to classify bidders into normal or suspicious.

In the following, we highlight two important research directions:

- The generated SB dataset is highly imbalanced, which will negatively impact the performance of classifiers as demonstrated in numerous studies such as [9]. The decision boundary of the fraud classifiers will be biased toward the normal class, which means suspicious bidders will be poorly detected. Handling the class imbalance problem is a continuous area of study [21]. In our research, we will investigate this problem by testing different types of techniques, such as data sampling and cost-sensitive learning, to determine the most suitable technique for our SB dataset.
- Ensemble learning has produced a reliable performance for many practical applications. The goals defined by ensemble learning are lowering the model's error ratio, avoiding the overfitting problem, and reducing the bias and variance errors. The most common ensemble methods are Boosting and Bootstrap Aggregation (Bagging). Thus, we will employ ensemble learning to develop a robust SB detection model, and examine the most fitting ensemble strategy for our SB dataset.

Acknowledgements The first author would like to thank the Saudi Arabian Cultural Bureau in Canada and the Umm Al-Qura University for their financial support.

References

1. Alzahrani, A., Sadaoui, S.: Scraping and preprocessing commercial auction data for fraud classification (2018). arXiv:1806.00656
2. Center, I.C.C.: 2015 internet crime report. In: 2015 IC3 Report. IC3 (2016)
3. Chang, J.S., Chang, W.H.: Analysis of fraudulent behavior strategies in online auctions for detecting latent fraudsters. Electr. Commer. Res. Appl. **13**(2), 79–97 (2014)
4. Cui, X., Lai, V.S.: Bidding strategies in online single-unit auctions: their impact and satisfaction. Inf. Manag. **50**(6), 314–321 (2013)
5. Dong, F., Shatz, S.M., Xu, H.: Combating online in-auction fraud: clues, techniques and challenges. Comput. Sci. Rev. **3**(4), 245–258 (2009)
6. Dong, F., Shatz, S.M., Xu, H.: Reasoning under uncertainty for shill detection in online auctions using dempster-shafer theory. Int. J. Softw. Eng. Knowl. Eng. **20**(07), 943–973 (2010)
7. Ford, B.J., Xu, H., Valova, I.: Identifying suspicious bidders utilizing hierarchical clustering and decision trees. In: IC-AI, pp. 195–201 (2010)

8. Ford, B.J., Xu, H., Valova, I.: A real-time self-adaptive classifier for identifying suspicious bidders in online auctions. Comput. J. **56**(5), 646–663 (2012)
9. Ganguly, S., Sadaoui, S.: Classification of imbalanced auction fraud data. In: Canadian Conference on Artificial Intelligence, pp. 84–89. Springer (2017)
10. Ganguly, S., Sadaoui, S.: Online detection of shill bidding fraud based on machine learning techniques. In: International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems, pp. 303–314. Springer (2018)
11. Guha, S., Rastogi, R., Shim, K.: CURE: an efficient clustering algorithm for large databases. In: ACM Sigmod Record, vol. 27, pp. 73–84. ACM (1998)
12. Gupta, P., Mundra, A.: Online in-auction fraud detection using online hybrid model. In: 2015 International Conference on Computing, Communication and Automation (ICCCA), pp. 901–907. IEEE (2015)
13. Jank, W., Shmueli, G.: Studying heterogeneity of price evolution in ebay auctions via functional clustering. In: Business Computing. Handbook of Information Systems Series, pp. 237–261 (2009)
14. Lei, B., Zhang, H., Chen, H., Liu, L., Wang, D.: A k-means clustering based algorithm for shill bidding recognition in online auction. In: 2012 24th Chinese Control and Decision Conference (CCDC), pp. 939–943. IEEE (2012)
15. Ravisankar, P., Ravi, V., Rao, G.R., Bose, I.: Detection of financial statement fraud and feature selection using data mining techniques. Decis. Support Syst. **50**(2), 491–500 (2011)
16. Rousseeuw, P.J.: Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. **20**, 53–65 (1987)
17. Sadaoui, S., Wang, X.: A dynamic stage-based fraud monitoring framework of multiple live auctions. Appl. Intell. **46**(1), 197–213 (2017)
18. Xu, D., Tian, Y.: A comprehensive survey of clustering algorithms. Ann. Data Sci. **2**(2), 165–193 (2015)
19. Yu, H., Liu, Z., Wang, G.: An automatic method to determine the number of clusters using decision-theoretic rough set. Int. J. Approx. Reason. **55**(1), 101–115 (2014)
20. Zhang, J., Prater, E.L., Lipkin, I.: Feedback reviews and bidding in online auctions: An integrated hedonic regression and fuzzy logic expert system approach. Decis. Support Syst. **55**(4), 894–902 (2013)
21. Zhang, S., Sadaoui, S., Mouhoub, M.: An empirical analysis of imbalanced data classification. Comput. Inf. Sci. **8**(1), 151–162 (2015)

Big Data Security Challenges and Preventive Solutions

Nirmal Kumar Gupta and Mukesh Kumar Rohil

Abstract Big data has opened the possibility of making great advancements in many scientific disciplines and has become a very interesting topic in academic world and in industry. It has also given contributions to innovation, improvements in productivity and competitiveness. However, at present, there are various security risks involved in the process of collection, storage and use. The leakage of privacy caused by big data poses serious problems for the users; also the incorrect or false big data may lead to wrong or invalid analysis of results. The presented work analyzes the technical challenges of implementing big data security and privacy protection, and describes some key solutions to address the issues related with big data security and privacy.

Keywords Big data · Big data analysis · Big data security · Privacy protection

1 Introduction

In today's world, a large number of people share their social information and behavior using the internet and it has led to the explosion of data generated. The constant advance of technologies has allowed an "explosive" growth in the amount of data generated from different sources, for example social networks, mobile devices, sensors, X-ray machines, telescopes, space probes, applications logs, climate predictions, geo-positioning systems and, in general terms, everything that can be classified within the definitions of the internet of things [1]. According to

N. K. Gupta (✉)
Department of Computer Science and Engineering, Jaypee University Anoopshahr, Anoopshahr, India
e-mail: nirmal.gupta@mail.jaypeeu.ac.in

M. K. Rohil
Department of Computer Science and Information Systems, Birla Institute of Technology and Science, Pilani, India
e-mail: rohil@pilani.bits-pilani.ac.in

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_21

statistics, an average of 40,000 search queries occur every second; in other words, it can be said that over 3.5 billion searches are processed per day by Google search [2]. At the same time, various monitoring and sensing equipment are also generating data continuously. There is also a large amount of data in various fields such as scientific computing, healthcare, finance and retail. This phenomenon has aroused widespread concern.

The generation of this big data makes data analysis and application more complicated and difficult to manage. According to statistics, the amount of data generated globally over the past 3 years is more than the previous 400 years of data, which include documents, pictures, videos, web pages, e-mail, microblogging, and other types, which include mostly unstructured data as compared to structured data [3]. Gartner had forecasted that 4.9 billion connected objects to be in use by 2015, up 30% from 2014 and will reach 25 billion by 2020 [4]. At present, big data has become another information industry growth point in the field of information technology after cloud computing.

Like other current information systems, big data also involves security risks in the process of storage, processing, transmission and similarly it needs the data and privacy to be protected. The services like storage and management of data are provided in big data and cloud computing by the service providers themselves. But, the problem with big data as compared to cloud computing is that in case of cloud computing the user still has some control over their data to some extent; for example, through the use of cryptographic methods and through the other trusted computational methods. However, in the context of big data the businesses like Facebook not only produce the data but also provide the services like storage and management of data themselves. Therefore, it is extremely difficult to restrict the use of user information by means of technology and to protect the privacy of users' data [5].

At present, many organizations are aware of the big data security issues, and actively take action to focus on big data security issues. This paper focuses on the security challenges brought by the current big data technology and elaborates on the key technologies used for big data security and privacy protection. It should be pointed out that while introducing new security issues and challenges, big data also brings new opportunities in the field of information security. That is, big data-based technologies for information security can be used in turn for big data security and privacy protection.

2 Big Data Research Overview

2.1 Big Data Sources and Characteristics

Big data may have various sources from where it may be generated. Based upon the source of generation of big data, it can be divided into following categories [6]:

1. From the people: all kinds of data generated by people in the process of performing activities over the internet. The generated data can be in the form of text, images, videos or of any other type.
2. From the machines: This includes the data which is generated by different computers and information processing units which may be in the form of files, databases, multimedia, and so on, and also includes automatically generated information such as logs.
3. From the devices: data collected by various types of digital devices, such as the digital signals continuously generated by the camera; the different data related to human beings generated through various medical devices; the large amount of data which is generated by the astronomical telescopes.

2.2 Big Data Analysis Goal

At present, big data analysis is applied to various diversified areas such as science, medicine and commerce. Overall, the goals of big data analytics fall into the following categories:

(1) *Gain knowledge through extensive analysis*

People have a long history of data analysis. There may be various reasons. The first and most important reason for analyzing the data is to get knowledge from it. Since there is a large amount of unprocessed real sample information, it can effectively abandon individual differences and help people through the mining, and more accurately grasp the common purpose behind the things. Depending on the knowledge they have discovered, one can predict more accurately the nature or social phenomena that will occur. Typical examples include the ability to retrieve information about the flu through Google's search using data mining [7]; predicting stock quotes based on Twitter information [8]; and so on.

(2) *Grasp individual laws through long-term analysis*

Individual activities have distinct personal characteristics while satisfying certain common characteristics. Through long-term multi-dimensional data accumulation and gaining information through that data, various companies get the insights of users' behavior and this may help them to accurately describe the individual user's profiles. In this way, it helps them to provide more accurate products and services according to users' individual needs. It also helps companies to accurately provide recommendations related to advertisements.

For example, Google analyzes users' habits and hobbies through its big data products, helping advertisers to evaluate the efficiency of their advertising campaigns, and it is estimated that there may be hundreds of billions of dollars in the market in the future [9].

(3) *Control epidemic through analysis*

Many times the timely information obtained through the analysis of big data can provide more valuable information regarding spread of epidemics, than by disease-prevention centers. For example, during the 2009 flu pandemic, big data analysis was performed by Google to get the timely information. Generally, patients do not go to the doctor immediately just after getting infected, but their search and discussion trends can be analyzed to get information about most influenced geographical areas.

2.3 Big Data Technology Framework

Big data processing involves data collection, management, analysis and display. Figure 1 is a schematic view of the relevant technology, including four stages.

1. Data acquisition and preparation

The data sources of big data are diversified, including all kinds of structured, unstructured and semi-structured data such as databases, texts, pictures, videos and web pages. Therefore, the first step in big data processing is to gather data from the data source and pre-process it to provide a consistent, high-quality data set for subsequent processes.

Since there exist various sources of big data, therefore, there may be different models for its description and these may even contradict. Therefore, it becomes important to clean the data during the data integration process so that similar, repetitive or inconsistent data can be removed. In the literature, data cleaning and

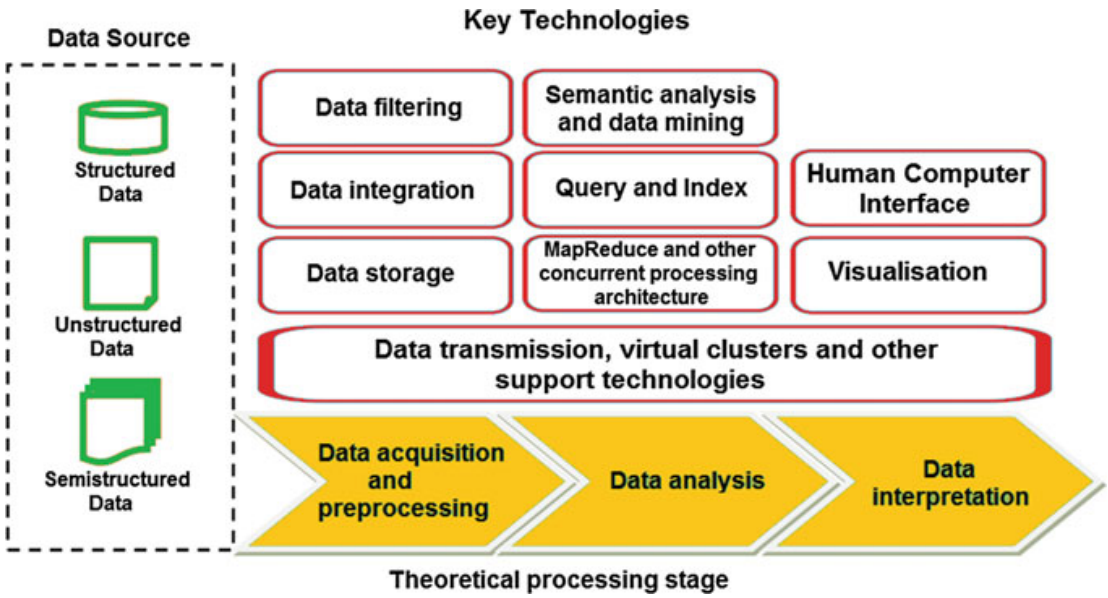


Fig. 1 Big data technology architecture

integration technology is aimed at the characteristics of big data, and proposes the cleaning of unstructured or semi-structured data and the integration of very large-scale data [9].

Data storage and big data applications are closely related. Real-time applications using big data require strong computational ability so that large amount of data could be processed in less amount of time. Therefore, big data processing system for real-time systems should quickly respond to the requests generated. This requires robust computing power for big data. Stream processing mode is more suitable for such applications. Most of the other applications need storage for subsequent deeper data analysis process. This may increase the storage cost. Usually big data uses distributed architecture to increase data throughput and reduce storage cost. Representative studies include file system GFS, HDFS and Haystack; NoSQL database Mongodb, CouchDB, HBase, Redis, Neo4j and so on [10].

2. Data analytics

Traditional data analysis may not work for big data as it was devised for structural data sources, but big data mostly consists of semi-structured or unstructured data. This presents a big challenge to big data analysis, but it is also the core process of big data applications. Based on the different levels, it can be broadly divided into three categories: computing architecture, queries and indexes, and data analysis and processing.

In terms of computing architecture, MapReduce [11] is a widely used big data set computing model and framework. In order to adapt to some analysis requirements that require high task completion time, work in [12] optimizes its performance. In [13], a data flow analysis solution based on MapReduce architecture, MARISSA, was proposed to support real-time analysis task. Dede et al. [14] proposed Mastiff, a big data analysis program based on time. Chandramouli et al. [15] proposed a TiMR framework based on MapReduce to deal with the real-time streaming for applications with high real-time demand such as advertisement push.

In query and indexing, traditional relational database query and indexing techniques are limited due to the large amount of unstructured or semi-structured data contained in big data, and NoSQL database technology received more attention. For example, Chandramouli et al. [16] proposed a hybrid data access architecture, HyDB, as well as a method of concurrent data query and optimization. Wang et al. [17] optimizes the query of key-value type database.

In data analysis and processing, the main technologies involved include semantic analysis and data mining. Owing to the diversification of data in big data environment, it is difficult to unify the terms to mine information when semantic analysis of data is concerned. In [18], for the big data environment, a high-efficiency terminology standardization method for solving the term variation problem is proposed. Keteta et al. [19] studied the heterogeneity of semantic ontology in semantic analysis. Traditional data mining technology is mainly aimed at structured data, so it is very important to study unstructured or semi-structured

data mining technology. Kang et al. [20] proposed a mining technique for image files, and Kang et al. [21] proposed a large-scale TEXT file retrieval and mining technology.

3. Data interpretation

The purpose of data interpretation is to represent data analysis results in a way which can serve the user's purpose. The major technologies which make it possible are visualization and human–computer interaction. There have been some visualization studies for large-scale data [22], which solve the display problem of large-scale data through data projection, dimension degradation or display wall. As human visual sensitivity limits the effectiveness of larger screens, a human-centric human–computer interaction design will also be an important technique to address the display of big data analytics results.

4. Other support technologies (data transmission and virtual cluster)

Although big data applications emphasize data-centric computing and push calculations to data execution, data transmission is still essential throughout the process, such as the transmission of some scientific observations from observation points to data centers. In [23] the authors study efficient transmission architectures and protocols for big data features.

In addition, because virtual clusters have the advantages of low cost, flexible construction and easy management, people can choose more convenient virtual clusters to complete the various processing tasks during big data analysis. Therefore, virtual machine cluster optimization research for big data applications is needed [24].

3 Big Data Security Challenges

Big data provides a great technology which has its significance in various fields and the security requirements in these fields are also changing. During the various activities performed over data during its collection, refinement and mining, many security threats are also associated. During this process the data may be destroyed, leaked, tampered, which can put the user privacy or corporate secrets to be compromised. In general, the security related to big data has the following characteristics and challenges.

3.1 Technical Challenge

The technical challenges of big data applications are mainly in the following three aspects:

First, as we all know, the application of big data is based on the possession of massive data. This involves the challenges of data storage technology and the technical challenges for data processing and analysis, including the data processing capabilities of computer hardware and supercomputer algorithm technology. The report learned from interviews with technical experts indicates that they are not optimistic about the recent overcoming of the technical challenges described above. Second, in the educational application of big data, data collection and problem-solving analysis are the core links, and application developers have to face the challenges of data acquisition technology and problem-solving analysis technology. The third aspect is the data compatibility challenge, the inconsistency of data encoding and format in different data storage systems, resulting in data sharing difficulties between different systems. The main reason for this problem is the lack of unified planning for the construction and purchase of various systems which results in inability to form unified data platform.

3.2 Mobile Data Security Challenge

In the world of today, various new applications related to social media have been emerged; besides this it is not uncommon for e-commerce portals and mobile applications to generate enormous amount of data. Analyzing such large amount of data in addition to data generated through internet of things network is a big challenge for companies. It also put responsibilities over the companies as the demand for data security capabilities on enterprises also increases greatly. In addition, the increase in data secrecy results in more sensitive analysis of data which is being transferred between mobile devices. Personal information may also be compromised because of the malware which may track user location or steals the confidential information. It requires increasing the safety related to user's privacy and data. Since the mobile devices have grown rapidly in recent years, it has created a challenge for big data security that how the samples of malware could be tracked which may enable analyzing these malware samples and finding the relationship between them.

3.3 Easy Target for Attackers

The various resources and personalized services are being provided through the network access in a flexible manner. In such a large networked society, the voluminous data attract the hackers because of the potential value associated with it. Such big data is large and has interrelated information stored, and due to this an attacker finds it easy to get more information by successfully attacking once, which further reduces the attacking cost and increases their profit value. A good example of such situation is when hackers use information in big data and take control of

important systems to take advantage from it for the time being. Hackers can also deploy AI technologies to take full control of the system by slowly penetrating it over time. Such strategy may work, as it would perform minor changes and such changes would appear natural. Thus, its detection could be avoided.

3.4 User Privacy Protection Challenge

Worldwide people use different kinds of web services and while using such services they need to access the network which stores their credentials like name, username, password, contact information, address and so on. It also includes their other personal information, behavior, habit, political or religious inclination and location; this has increased the risk of user privacy and data disclosure. Through the use of various data mining approaches such data can be easily recorded. If these data are found to be useful, then companies in the relevant areas of customer's need and habit can target to the specific users to achieve greater benefits. The traditional approach of privacy protection is based on anonymity of user data which is publicly available. But in reality, the user privacy protection cannot be achieved only through data protection through anonymity. There exist many other requirements and characteristics for user privacy. But the problem is that most of the existing privacy protection models and algorithms are only for traditional relational data, and cannot be directly applied to big data applications.

3.5 Safe Storage of Massive Data Challenge

The quantity of big data has increased tremendously, which includes structured and unstructured data. It has increased to such a level that the previous storage systems are not capable to meet the needs of big data applications. The current disk technology has some limitations and this limit is around 4 terabytes per disk. So for such a massive data of exabyte levels, this limitation creates a problem, because just to store 1 exabyte requires around 25,000 disks. It will really create the hardware issue of how to attach such large number of disks to a single computer, even if the computer is capable of processing that much of data.

To satisfy the demands of data storage of such large amount, various new storage technologies have been developed. Storage technologies such as direct attached storage (DAS), network attached storage (NAS) and SAN (storage area network) are being used to solve the data storage-related issues. Another technology called NoSQL storage technology is also used to capture, manage and process big data. Using NoSQL data storage can be extended and performance also can be improved, but still some issues exist. These issues include access control and

privacy control issues, issues related to technical vulnerabilities, security issues for authorization and verification, data management and confidentiality issues, and so on.

3.6 Challenge of Analysis

With the increase of storage capabilities, prior selection of data became insignificant. This can be seen as a real chance, allowing to keep focus on the potential future uses and which are not always fully defined at the time of their acquisition. In particular, many issues that were considered nonexistent later become accessible before the use of their potential of significant advantage (competitive advantage for example). It should be kept in mind, however, that more data is not always better data. It depends on whether or not they are heterogeneous, and whether they are representative of what is being sought. In addition, as the number of parameters increases, the number of erroneous correlations also increases. The analysis part will have to take into account these essential aspects.

Heterogeneous (structured, unstructured) data or incomplete or uncertain data for which specific treatments are needed will also be stored. Moreover, in this regard specific treatments are already required for the more standard data, even if some old methods remain effective for the volume of the existing data, which may cause theoretical and practical difficulties, unknown in advance. Thus, the simple statistical tests [25] become inoperative for large sample sizes. We can also mention the difficulty of multi-dimensional analysis on large sets of data, which arises during their interpretation. The visualization is considered until now as an extremely powerful tool, but it risks becoming inoperative by simple graphic saturation effect. In addition, real-time analysis of continuous flow of data from different sources also poses specific challenges. All these problems involve the development of new statistics for big data, for example, requiring a review of basic calculations such as statistical tests and correlations [26]. Of course, these technical analysis tools cannot be isolated from the computer tools and techniques dedicated to big data, for example, NoSQL, Hadoop, MapReduce or Spark.

3.7 Big Data Security Trust Challenge

Although big data has provided various opportunities to its users, but still it lacks the complete trust of the people using it. The visibility of social profiles generated by users varies across different types of networks and these are crawled by the search engines and therefore they become visible by the other users whether they have account or not. Therefore, here a trust issue arises from the user that how safe their data privacy is. This requires trust measures to be integrated with big data. These trust measures should not be treated as a static measure. That means, as the

data evolves the trust measures should also be updated accordingly. Yin and Tan [27] in their research have put the fact that semi-supervised learning methods that start with ground truth data are able to provide higher accuracy and trust on the source data. Another fact is that the different people have different personal opinions regarding various factors affecting their life and when there exist differences with statistics it leads to the market doubts about statistics.

4 Key Technologies for Big Data Security and Privacy Protection

At present, it is urgent to carry out research on key technologies of big data security in view of security challenges such as user privacy protection, data content credibility and access control faced by the big data. This section introduces some related key research areas for this.

4.1 Data Anonymity Protection Technology

For structured or unstructured data in big data, the core key technologies and basic means for data protection to achieve its privacy protection are still in the stage of development and improvement. Take the typical k anonymity scheme as an example. The early schemes [28] and their optimization schemes [29] group quasi-identifiers by data processing such as tuple generalization and suppression. The quasi-identifiers in each packet are the same and contain at least k tuples, so each tuple is at least indistinguishable from $k - 1$ other tuples. Since the k-anonymous model is for all attribute collections, it is not defined for a specific attribute, and it is prone to insufficient anonymity of a certain attribute. If the value of a sensitive attribute in an equivalence class is the same, the attacker can effectively determine the value of the attribute. This research is for static, one-time release. In reality, data publishing often faces scenarios in which data is continuously and repeatedly released. It is necessary to prevent an attacker from analyzing the data associations that are published multiple times, and destroying the original anonymity of the data [30].

In big data scenarios, data anonymity protection is more complicated: an attacker can get data from multiple sources, not just the same source. For example, in the Netflix application, people [31] found that an attacker could identify the target's Netflix account by comparing the data to the publicly available imdb. According to this, the user's political inclinations and religious beliefs are obtained (obtained through the user's viewing history and comments and scoring analysis of certain movies). Such issues are subject to further research.

4.2 Social Network Anonymity Protection Technology

The data generated by social networks is one of the important sources of big data, and it contains a large amount of user privacy data. Because social networks have the characteristics of graph structure, their anonymous protection technology is very different from structured data.

The social networks also require anonymity protection. Here, some typical requirements are user anonymity and attribute anonymity. While using social networks the user identity and attribute information is also required to be hidden while publishing these. The related data of different users should not disclose their relationship and anonymity between users is required. It is also known as edge anonymity. Hide the relationship between users when publishing. The attacker tries to use the various attributes of the node (degrees, tags, some specific connection information, etc.) to re-identify the identity information.

The current side-anonymity schemes are mostly based on additions and deletions of edges. The method of randomly adding and deleting exchange edges can effectively implement edge anonymity. Among them, Ying et al. [32] keep the eigenvalues of the adjacency matrix and the corresponding second eigenvalues of the Laplacian matrix in the anonymity process. Zhang et al. [33] group according to the degree of the nodes, and select the nodes with the same degree. The problem with this type of method is that the randomly added randomness is too scattered and sparse, and there is a problem of insufficient protection of the anonymous side.

4.3 Data Watermarking

Digital watermarking refers to a method in which identification information is embedded in a data carrier in an imperceptible manner without affecting its use, and is more commonly found in multimedia data copyright protection. There are also some watermarking schemes for databases and text files.

The method of adding watermarks in databases and documents is very different from the multimedia carrier, which is determined by the characteristics of data disorder and dynamics. The basic method is that there can be redundant information in the data or it can bear certain precision errors. For example, Agrawal et al. [34] have an error tolerance range based on numerical data in the database, embedding a small amount of watermark information into the least significant bits randomly selected from these data. Sion et al. [35] proposed a scheme based on statistical features of data sets, embedding one-bit watermark information in a set of attribute data to prevent attackers from destroying the watermark. In addition, by embedding database fingerprint information in the watermark [36], the owner of the information and the object being distributed can be identified, which is beneficial for tracking the leak in a distributed environment. Watermarking based on text content [37] depends on modifying the content of the document, such as adding spaces,

modifying punctuation, natural language-based watermarking [38] and so on, through the understanding of semantics to achieve changes, such as word substitution or sentence changes.

4.4 Data Traceability Technology

Owing to the diversified sources of data, it is necessary to record the source of the data and its distribution to provide support for later mining and decision-making. Data provenance technique has been extensively studied in the database field long before the big data concept emerged. The basic purpose is to help people determine the source of each data in the data warehouse, for example, which data items in tables are computed, so that it is convenient to check the correctness of the results at a very small cost. The basic method of data tracing is notation, such as marking the data in the data warehouse in [39] to record the query and propagation history of the data in the data warehouse. Data traceability techniques can also be used for traceability and recovery of files. For example, the work in [40] created a prototype system of data origin storage systems by extending the Linux kernel and file system, which can automatically collect origin data.

Further data traceability technologies can play an important role in the field of information security. However, data traceability technology also faces the following challenges in the protection of big data security and privacy:

1. *The balance between data traceability and privacy protection.* Using traceability to provide big data security protection requires first to obtain big data source using analysis of big data. Then the next step becomes to define the security policy and provide the required security mechanism. Often, the source of big data is privacy-sensitive and users are not interested that data to be accessed by the analysts also. Therefore, the problem is how to balance these two requirements simultaneously so that data traceability and privacy protection of the data both can be achieved.
2. *Security protection of data security technology itself.* The data tracing techniques currently employed are unable to handle the security issues correctly. The problem is that how to determine whether the tag associated with the data is itself correct or not. The other problem is that the tag information itself may not be securely bound with the data content and there may be other similar issues. Also in case of big data, since it is implemented on such a large scale, high speed and diverse characteristics such problems become more important.

5 Conclusion

The arrival of the era of big data has opened great opportunities. Big data not only has impact on everyone's social and economic behavior but also has influenced their way of living and thinking. Although big data is an important solution to various problems, it has also brought new security issues into existence. From the perspective of privacy protection, trust and access control of big data, this paper analyzes various security features and problems in the big data environment, namely, mobile data security, attack targets, user privacy protection challenges and security, storage issues, data security evolution, trust security issues and so on, and also discusses the preventive solutions for them. However, generally speaking, the current research on the protection of big data security and privacy is not sufficient. Only through the combination of technical means and relevant policies and regulations combined can better solve the big data security and privacy protection issues.

References

1. Xia, F., Yang, L.T., Wang, L., Vinel, A.: Internet of things. Int. J. Commun. Syst. **25**(9), 1101–1102 (2012)
2. Google search statistics. http://www.internetlivestats.com/google-search-statistics/
3. Lee, I.: Big data: dimensions, evolution, impacts, and challenges. Bus. Horiz. **60**(3), 293–303 (2017)
4. Nguyen, B., Simkin, L.: The Internet of Things (IoT) and marketing: the state of play, future trends and the implications for marketing. J. Mark. Manage. **33**(1–2), 1–6 (2017)
5. Boyd, D., Crawford, K.: Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. Inf. Commun. Soc. **15**(5), 662–679 (2012)
6. Guo-Jie, L., Xue-Qi, C.: Research status and scientific thinking of big data. Bull. Chin. Acad. Sci. **27**(6), 647–657 (2012)
7. Wu, X., Zhu, X., Wu, G.Q., Ding, W.: Data mining with big data. IEEE Trans. Knowl. Data Eng. **26**(1), 97–107 (2014)
8. Arias, M., Arratia, A., Xuriguera, R.: Forecasting with twitter data. ACM Trans. Intell. Syst. Technol. **5**(1), 8 (2013)
9. Arasu, A., Chaudhuri, S., Chen, Z., Ganjam, K.: Experiences with using data cleaning technology for bing services. IEEE Data Eng. Bull. **35**(2), 14–23 (2012)
10. Sarma, A.D., Dong, X.L., Halevy, A.: Data integration with dependent sources. In: Proceedings of the 14th International Conference on Extending Database Technology, ACM, pp. 401–412 (2011)
11. Elomari, A., Maizate, A., Hassouni, L.: Data storage in big data context: a survey. In: International Conference on Systems of Collaboration (SysCo), pp. 1–4. IEEE (2016)
12. Dean, J., Ghemawat, S.: MapReduce: simplified data processing on large clusters. Commun. ACM **51**(1), 107–113 (2008)
13. Verma, A., Cherkasova, L., Kumar, V.S., Campbell, R.H.: Deadline-based workload management for mapreduce environments: pieces of the performance puzzle. In: Network Operations and Management Symposium (NOMS), pp. 900–905, IEEE (2012)

14. Dede, E., Fadika, Z., Hartog, J., Govindaraju, M., Ramakrishnan, L., Gunter, D., Canon, R.: Marissa: Mapreduce implementation for streaming science applications. In: IEEE 8th International Conference on E-Science (e-Science), 2012, pp. 1–8, IEEE (2012)
15. Guo, S., Xiong, J., Wang, W., Lee, R.: Mastiff: a mapreduce-based system for time-based big data analytics. In: IEEE International Conference on Cluster Computing (CLUSTER), 2012, pp. 72–80, IEEE (2012)
16. Chandramouli, B., Goldstein, J., Duan, S.: Temporal analytics on big data for web advertising. In: IEEE 28th International Conference on Data Engineering, pp. 90–101. IEEE (2012)
17. Chen, C.P., Zhang, C.Y.: Data-intensive applications, challenges, techniques and technologies: a survey on Big Data. Inf. Sci. **275**, 314–347 (2014)
18. Wang, Y., Lu, W., Wei, B.: Transactional multi-row access guarantee in the key-value store. In: IEEE International Conference on Cluster Computing (CLUSTER), 2012, pp. 572–575. IEEE (2012)
19. Hwang, M., Jeong, D.H., Jung, H., Sung, W.K., Shin, J., Kim, P.: A term normalization method for better performance of terminology construction. In: International Conference on Artificial Intelligence and Soft Computing, pp. 682–690. Springer, Berlin (2012)
20. Ketata, I., Mokadem, R., Morvan, F.: Biomedical resource discovery considering semantic heterogeneity in data grid environments. In Integrated Computing Technology, pp. 12–24. Springer, Berlin (2011)
21. Kang, U., Chau, D.H., Faloutsos, C.: Pegasus: mining billion-scale graphs in the cloud. In: 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5341–5344. IEEE (2012)
22. Kola, A., More, H., Soderman, S., Gubanov, M.: Generating Unified Famous Objects (UFOs) from the classified object tables. In: 2017 IEEE International Conference on Big Data (Big Data), pp. 4771–4773. IEEE (2017)
23. Tang, J., Liu, J., Zhang, M., Mei, Q.: Visualizing large-scale and high-dimensional data. In: Proceedings of the 25th International Conference on World Wide Web, pp. 287–297. International World Wide Web Conferences Steering Committee (2016)
24. Hashem, I.A.T., Yaqoob, I., Anuar, N.B., Mokhtar, S., Gani, A., Khan, S.U.: The rise of "big data" on cloud computing: review and open research issues. Inf. Syst. **47**, 98–115 (2015)
25. Meyer-Schönberger, V., Cukier, K.: Big Data: A Revolution That Will Transform How We Live, Work, and Think. Eamon Dolan, England (2013)
26. Gandomi, A., Haider, M.: Beyond the hype: big data concepts, methods, and analytics. Int. J. Inf. Manage. **35**(2), 137–144 (2015)
27. Yin, X., Tan, W.: Semi-supervised truth discovery. In: Proceedings of the 20th International Conference on World Wide Web, pp. 217–226. ACM (2011)
28. Sweeney, L.: k-anonymity: a model for protecting privacy. Int. J. Uncertain. Fuzziness Knowl. Syst. **10**(5), 557–570 (2002)
29. LeFevre, K., DeWitt, D.J., Ramakrishnan, R.: Incognito: efficient full-domain k-anonymity. In: Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data, pp. 49–60. ACM (2005)
30. Xiao, X., Taom, Y.: M-invariance: towards privacy preserving re-publication of dynamic datasets. In: Proceedings of the 2007 ACM SIGMOD International Conference on Management of Data, pp. 689–700. ACM (2007)
31. Narayanan, A., Shmatikov, V.: Robust de-anonymization of large sparse datasets. In: IEEE Symposium on Security and Privacy, 2008. SP 2008, pp. 111–125. IEEE (2008)
32. Ying, X., Wu, X.: Randomizing social networks: a spectrum preserving approach. In: Proceedings of the 2008 SIAM International Conference on Data Mining, pp. 739–750. Society for Industrial and Applied Mathematics (2008)
33. Zhang, L., Zhang, W.: Edge anonymity in social network graphs. In: International Conference on Computational Science and Engineering, 2009. CSE'09, vol. 4, pp. 1–8. IEEE (2009)
34. Agrawal, R., Haas, P.J., Kiernan, J.: Watermarking relational data: framework, algorithms and analysis. VLDB J. Int. J. Very Large Data Bases **12**(2), 157–169 (2013)

35. Sion, R., Atallah, M., Prabhakar, S.: On watermarking numeric sets. In: International Workshop on Digital Watermarking, pp. 130–146. Springer, Berlin (2002)
36. Guo, F., Wang, J., Li, D.: Fingerprinting relational databases. In: Proceedings of the 2006 ACM Symposium on Applied Computing, pp. 487–492. ACM (2006)
37. Pease, A., Niles, I., Li, J.: The suggested upper merged ontology: a large ontology for the semantic web and its applications. In: Working Notes of the AAAI-2002 Workshop on Ontologies and the Semantic Web, vol. 28, pp. 7–10 (2002)
38. Atallah, M.J., Raskin, V., Hempelmann, C.F., Karahan, M., Sion, R., Topkara, U., Triezenberg, K.E.: Natural language watermarking and tamperproofing. In: International Workshop on Information Hiding, pp. 196–212. Springer, Berlin (2002)
39. Cui, Y., Widom, J., Wiener, J.L.: Tracing the lineage of view data in a warehousing environment. ACM Trans. Database Syst. **25**(2), 179–227 (2000)
40. Muniswamy-Reddy, K.K., Holland, D.A., Braun, U., Seltzer, M.I.: Provenance-aware storage systems. In: USENIX Annual Technical Conference, General Track, pp. 43–56 (2006)

Role and Challenges of Unstructured Big Data in Healthcare

Kiran Adnan, Rehan Akbar, Siak Wang Khor and Adnan Bin Amanat Ali

Abstract Unprecedented growth in the volume of unstructured healthcare data has immense potential in valuable insight extraction, improved healthcare services, quality patient care, and secure data management. However, technological advancements are required to achieve the potential benefits from unstructured data in healthcare according to the growth rate. The heterogeneity, diversity of sources, quality of data and various representations of unstructured data in healthcare increases the number of challenges as compared to structured data. This systematic review of the literature identifies the challenges and problems of data-driven healthcare due to the unstructured nature of data. The systematic review was carried out using five major scientific databases: ACM, Springer, ScienceDirect, PubMed, and IEEE Xplore. The inclusion of articles in review at the initial stage was based on English language and publication date from 2010 to 2018. A total of 103 articles were selected according to the inclusion criteria. Based on the review, various types of healthcare unstructured data have been discussed from different domains of healthcare. Also, potential challenges associated with unstructured big data have been identified in healthcare for future research directions in the technological advancement of healthcare services and quality patient care.

Keywords Unstructured data · Healthcare · Big data · Systematic literature review · Challenges of healthcare data

K. Adnan (✉) · R. Akbar · S. W. Khor · A. B. A. Ali
Faculty of Information and Communication Technology,
Universiti Tunku Abdul Rahman, 31900 Kampar, Malaysia
e-mail: kiranadnan@1utar.my

R. Akbar
e-mail: rehan@utar.edu.my

S. W. Khor
e-mail: khorsw@utar.edu.my

A. B. A. Ali
e-mail: adnanbinamanat@1utar.my

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_22

1 Introduction

Rapid advancement in healthcare systems from disease-centric systems to patient-centric systems involves patient health records, integration of data from various health and social sources, advancement in mhealth, smart health and tele-health devices. Advanced Healthcare systems are required to support clinical decision support system, personalized and precision medicine, diagnostics systems, health and wellness monitoring, patient care and predictive analytics. This data-driven healthcare support leads to orchestrate data preprocessing, analytics, interpretation, management and curation techniques for improved healthcare services. Most of the healthcare data are unstructured such as medical prescriptions (hand-written, free-text), clinical notes, Electronic Health Records (EHR), Electronic Medical Records (EMR), Patient Health Record (PHR), Medical Imaging (Magnetic Resource Imaging, MRI, Photoacoustic imaging, Fluoroscopy, Positron emission tomography), discussion forums, social media, sensor data, medical sounds and video data, genomics, and in many other transactional, biometric and application data. By definitional perspective of unstructured data, it has no schema or predefined model, attitudinal and behavioral, and without specified format. This heterogeneity, variability, and diversity of unstructured data make it difficult to get valuable insights from unstructured data that if analyzed in a controlled manner, can improve healthcare analytics and extraction of more real values. In past years, the unstructured data growth rate is increasing exponentially. 90% of digital universe data are unstructured. 57% of all healthcare data will be useful if it is properly tagged and analyzed whereas only 3.1% of healthcare data are providing the highest value. It is a challenging and daunting task to find a needle in a haystack [1, 2].

The role of multifaceted unstructured big data in healthcare is very challenging and gaining popularity because it contains more rich information that is helpful to provide novel insights as well as improve technological capabilities in healthcare systems. Unstructured data in healthcare have great potential but a lot of challenges remain to be addressed. Several techniques have been proposed in the literature regarding big data management, analysis, and representation of data but very less number of articles dealt with the unstructured data in healthcare. The literature about unstructured big healthcare data emphasizes technological advancement need to be more focused. The main objective of this literature review is to highlight the most affected areas of healthcare big data solutions from unstructured data. This review is an attempt to answer the following question about unstructured data in healthcare, i.e., which domains of healthcare are suffering from the complexity of unstructured data, what challenges are being created by unstructured data, and what is the role of unstructured data in inefficient and poor results in analytics and data-driven decision-making in healthcare. The review discusses the types of unstructured big data in healthcare and the challenges associated with unstructured big healthcare data since 2010. A systematic literature review has been conducted in

this regard to fulfill the objectives of the study. The review of existing literature on the challenges associated with unstructured big healthcare data will help to improve usability and quality of unstructured data in various fields of healthcare. It will highlight the major challenges of unstructured data in healthcare that need advancement in tools and technologies.

The rest of the paper is organized in the following manner: Sect. 2 discusses the methodology used and the results of the systematic literature review. Challenges associated with unstructured big healthcare data in technological advancement have been elaborated in Sect. 3. Next Sect. 4 includes the findings of the review. Finally, Sect. 5 presents the conclusions.

2 Methodology

A systematic literature review has been used to conduct an effective review to highlight the potential challenges in healthcare domain due to unstructured data. The objectives of this research were to discuss the types of unstructured big data in healthcare and to explore the challenges associated with unstructured big healthcare data. This review has comprised of three phases: planning the review, conducting the review and reporting. In the planning phase, the fundamental search has been performed using specific keywords. The second phase involves the searching keywords on specific sources to conduct the review and results are obtained. In the third and last phase, results are presented.

2.1 Searching the Literature

According to the objectives of the study, the best suitable keywords were identified to organize more effective review. “Unstructured data” and “healthcare” were the queries formed from the keywords. This search query was used on some of the popular academic databases: IEEE Xplore, Springer, ScienceDirect, ACM digital library, and PubMed.

2.2 Inclusion Criteria

Articles deal with unstructured data in the healthcare domain, published between 2010 and 2018, using the English language was the inclusion criteria to capture the literature. According to the results of databases and inclusion criteria, a total number of articles (n) selected for review were 1522. Articles were filtered by title reading and the total number of articles reduced to (n) = 566. After that, abstracts

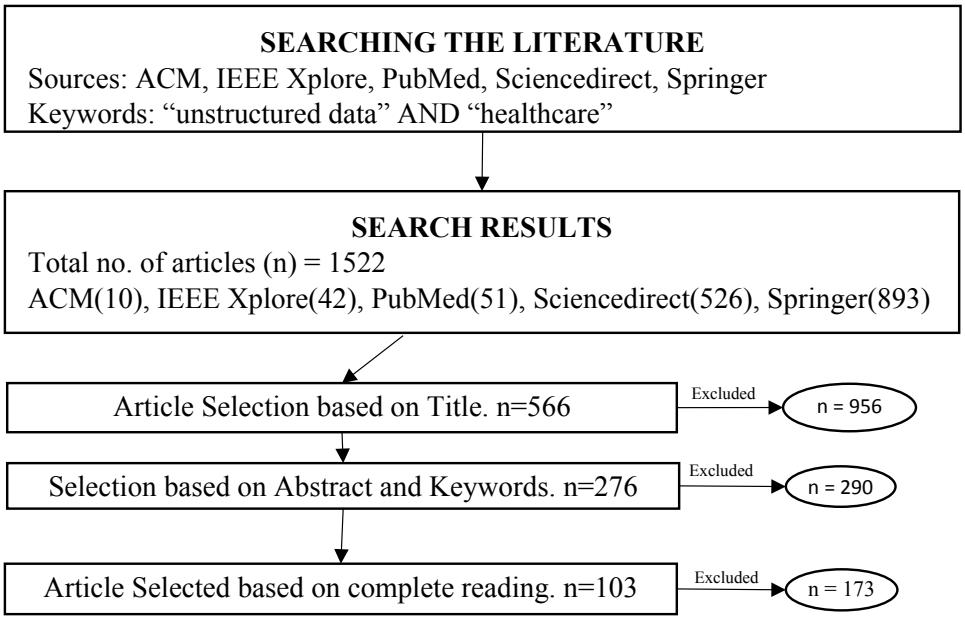


Fig. 1 Literature review process

and keywords were carefully studied and n decreased to 276. Based on the complete article study, 103 articles were selected for the study. Type of healthcare technology, type of unstructured data used, the domain of work and identified challenges are the records kept from each selected paper. Following Fig. 1 describes the whole research process in detail.

2.3 Results

This review was conducted to analyze the challenges in technological advancement due to the huge volume of unstructured healthcare data. The following Fig. 2 shows the research trend of selected articles for this review according to the year of publication. The figure shows the number of publications in each year that deals with unstructured data. An exponential rise can be seen in the following diagram that reflects the importance of unstructured data in the healthcare industry.

The results witness several types of healthcare advancement in four major categories: healthcare analytics (51% of selected articles), decision support systems (12%), health and wellness monitoring (18%), precision and personalized medicine (6%). In selected articles, several types of unstructured data have been used in various domains of healthcare such as disease identification from EHR, pharmacovigilance, drug effects identification, phenotyping, genotyping, and medical recommendation systems. Unstructured data are found in EHR, EMR, PHR, medical imaging, discussion forums, social media, sensor data, medical sounds and video data, genomics, and in many other transactional, biometric and application data. A major portion of selected research articles in this literature review was

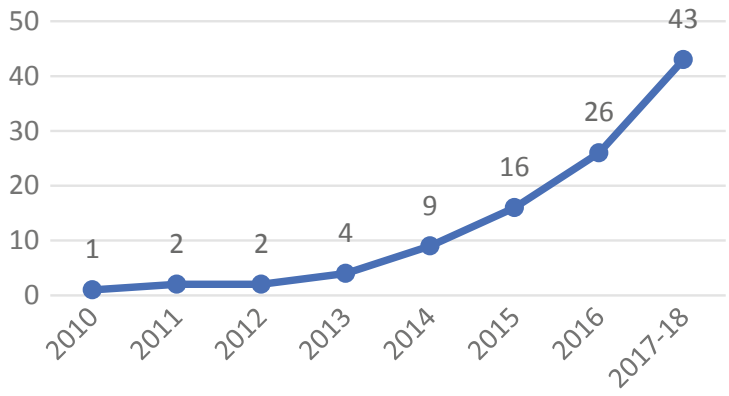


Fig. 2 Research trend of publications

related to the clinical hand-written notes, free-text in EHR, and text from social media. The main purpose of this literature review was to identify the most affected areas due to unstructured data in healthcare where technological advancement is required in order to improve the patient care. In the following section, challenges identified from the literature have been described in each category along with the requirements.

3 Challenges

This section describes the challenges of unstructured big data in healthcare identified in the literature. Selected articles are divided into five major categories with respect to their relevancy. The major categories of challenges are preprocessing, interoperability, Information Extraction (IE), analytics and result interpretation, and some other issues related to data exfiltration, standardization, and data curation. The following subsections describe the challenges associated with unstructured big healthcare data in detail.

3.1 Data Preprocessing

The variety of data is available in healthcare but these data are not fully usable for analysis and other processing due to noise, errors, missing values, and other quality issues. Preprocessing is a step that can reduce the inefficiency of unstructured data. As, 95% of healthcare data are unstructured data, available in various formats and facing more complexity, quality, usability and dimensionality issues. It arises the importance of preprocessing to improve the efficiency and accuracy of healthcare systems. Clinical decision support systems, health and wellness monitoring, disease identification, precision medicine, and analytics are important areas where unstructured data preprocessing techniques are required. Data preprocessing and

data cleaning are required for improved results in predictive analytics from social media text [3], complaint management in health information systems [4], decision support system from medical text and images [5], disease identification from health and wellness monitoring EHR data [6], in analytics and management of unstructured data [7, 8].

Data aggregation prepares the data in the required format as data is coming from various sources with different formats, frequencies, and representations such as structured EHR, semi-structured logs and unstructured images, audio and video data. The volume of unstructured data is far higher than the structured data. Hence, unstructured data aggregation and acquisition are more challenging. Free-text doctors' notes, handwritten clinical notes, patient care data, transactional, biometric and application data, EMR, and many other types of unstructured data are coming from various sources. Collection of these data found in a variety of formats is more challenging to handle [9]. Transformation engines clean, split, translate, merge, validate and sort unstructured data and make data prepared for analysis. Transformation techniques are required to transform the representation of data from unstructured to structured data in healthcare analytics, drug safety surveillance and clinical decision support systems [10–13]. Healthcare domains like identification of drug safety issues, disease identification systems need an automatic conversion of unstructured documents to structured data to increase the availability and usability of data analytics and databases [14–16]. But context-aware transformation based on user requirement and finding the relationships between the data items is still an open issue in healthcare. Personalized medicine, disease identification, adverse drug events identification, and medical knowledge base are the fields of healthcare where aggregation of structured and unstructured data is found critical. Integration of structured and unstructured data [17], improvement in predictive performance by combining structured and unstructured data [18], combining structured and unstructured data for comprehensive representation in predictive analytics [19], advanced analytical techniques [20], mapping of unstructured data to normalized content [21], complexities arise when outcomes of diverse sources need to be combined [22], and classification of combined data [23] are the challenges found in the literature.

Unstructured data require more advanced techniques than structured data because it has no schema, no predefined model, and no formal representation. Advancement is required to prepare unstructured healthcare data with more efficient data preprocessing methods. New big data technologies are emerging with adapted preprocessing. Apache Spark [24], Hadoop [25], and Flink [26] are performing better in processing big data. Apache Spark is making effort particularly in preprocessing and implemented MLlib to bridge the gap. It is highly desirable to develop new preprocessing techniques for data cleaning, discretization, feature selection, and data profiling using these advanced big data processing tools in healthcare.

3.2 Interoperability

Interoperability is another challenge to advance healthcare systems where two or more systems share and use data. Integration of data from different systems without any middleware and exchange of data poses more challenges to healthcare analytics (Table 1).

It requires common interfaces of healthcare systems, standardization of processes and data, and share common characteristics of a variety of data. Lack of uniform coding system across vendors is a challenge to interoperability that needs to be solved in order to manage patient information across hospitals and different applications with considering the ethical and privacy issues of patients' personal data. Following are the challenges identified in this literature review regarding interoperability in healthcare with a variety of healthcare data:

- Integration (Late or Early) of heterogeneous data and common type system for a variety of data types are the common challenges in predictive analytics in precision medicine and clinical diagnostics.
- Conventional databases are not adequate to handle this diversity of data types. Consolidated view of patient-related data using distributed storage is required for advanced information systems.
- Lack of harmonization between structured and unstructured data, evolving interoperability standards and integration of EHRs are creating challenges in analytics, data-driven medicine, precision and personalized medicine, and decision support system.
- All stakeholders need to work together in order to achieve interoperability [46]. Semantic and syntactic interoperability is important and essential for context awareness. Shared use of heterogeneous data, common representation of combined data, and deep understanding of unstructured data is an open issue. However, technical solutions are required for interoperability and integration.

3.3 Information Extraction

Disease diagnosis and its complications, communicable and non-communicable diseases' identification, drug surveillance, surgery details, and patient history are some examples of healthcare data where data are stored in the form of natural language text. Information extraction is a technique to extract structured information from these clinical texts with the help of natural language processing and machine learning. These information extraction techniques are helpful in predictive, prescriptive analytics, clinical care, clinical decision support system, disease identification from clinical notes, and precision medicine. Information extraction challenges due to unstructured healthcare data (i.e., free-text and handwritten clinical notes, EHR and EMR) found in the literature are described in Table 2.

Table 1 Related literature regarding interoperability and integration

Healthcare category	Data type used	Domain	Description of challenges
Analytics	Physician's and pharmacy notes, medical images	Systemic Lupus Erythematosus	Integration of various data formats, interoperability, evolving standards are the main challenges for which Common Type System is required for Clinical Natural Language Processing. The accuracy of predictive analytics can be improved using the advancement of these techniques [21, 27–31]
Analytics	EMR Clinical Free Text	Phenotyping	
Predictive Analytics	EHR Text	Clinical Diagnostics	
Analytics	Genomic Data	Evidence-based precision medicine	
Analytics	Omics data	Health data Spectrum	
Distributed Storage	Medical Images and associative text data	Radiology Information System	Conventional databases are not able to handle the complexity of a variety of data in healthcare. Consolidated View of Patient-relevant data from heterogeneous and disperse systems [32, 33]
Distributed Storage	Free-text, images and Proprietary formats	Medical Information Systems	
Data-driven Medicine	Text and medical images	Epilepsy	Lack of harmonization between structured and unstructured data, evolving interoperability standards, and integration of EHRs are the potential challenges and need solutions for efficient and effective curation [24, 25, 34–40]
Healthcare Systems	Genome and Medical images	European Single Market for health	
Data Analysis, Representation	EHR	Health Information Systems	
Precision and Personalized Medicine	Omics data	Molecular characterization of diseases	
Decision Support system	Behavior data and EHR	Medical Recommendations for Diabetes	
Information Management	Clinical Text	Common type of system development	Deep semantic understanding required structured data [41]. Semantic interoperability and data sharing are required to make computers context-aware [42, 43]. Shared use of heterogeneous data sources [44]. Interoperability, standardization, and integration of data sources are open issues [45]
Smart Healthcare	PHR	IohT	
Decision Support system	Structured and unstructured text	Pharmaceutical IE and Integration	
Medical Image Processing	Medical and diagnostic images	Medical diagnosis	
Decision Support system	Text and images of clinical data	Dental Clinics	

Table 2 Related literature regarding information extraction

Healthcare category	Data type used	Domain	Description of challenges
Analytics	Clinical documents	Open sources IE software	Extracting quality insights from healthcare data sources [47]
Analytics	Social Media streams	Drug and its effects detection	IE regarding drug and its effects from social media text streams is challenging due to ambiguous information [48]
Predictive Analytics	Clinical discharge summaries	Prediction of readmission from discharge summaries	Gaining information regarding readmission from unstructured clinical discharge summaries [49]
Analytics	Clinical Text data	Physiotherapy treatment	IE from unstructured EHR clinical notes [50]
Clinical Care	EHR	Phenotyping	Phenotype extraction from unstructured EHRs [51]
Personalized medicine	EHR and biomedical text literature	Genotype and phenotype extraction for cancer	Finding actionable granularity in disease identification and risk score [52]
Adaptive Analytics	Clinical Text trails of EMR and EHR	Cancer Clinical trials	Subject extraction [53]
Analytics	Unstructured medical notes	Disease identification	Semantically informed IE [54]
Sentiment Analysis	Discussion forums	Medical forum Discourse	Semantic extraction [55]
Exploratory Analytics	Physician and Biomedical data	Dashboard for querying heterogeneous data	Standardization, data formats, and their aggregation will improve semantic data processing [56]
Information retrieval and extraction	Clinical text documents	Software implementation in hospital	Finding relevant data [57]
Analytics and Representation	Imaging, Genetic and healthcare data		Multifaceted proprietary, open-source, and community developments [58]
Decision Support system	Medical and clinical references and publication text	Medical Recommendation System for Mental health	Predefined schema, large manually crafted dictionaries and ontologies are required to extract information from medical domain publications [59]
Clinical Decision Support System	EMR, Unstructured Clinical Notes	Automatic diagnosis prediction	Unstructured data combined with structured and semi-structured data will show good results [60]

(continued)

Table 2 (continued)

Healthcare category	Data type used	Domain	Description of challenges
Disease identification	Emergency dept. text notes	Automated identification of Pediatric Appendicitis Score	Unstructured data processing is challenging to extract useful information [61]
Bio surveillance	Clinical Free text	Detailed clinical and epidemiological variable extraction	Temporal IE and its integration for more elaborative and comprehensive representation of information [62]
Precision Medicine	EHR text	Prescreening and automated identification of HFpEF patients	The immense need for automatic data extraction system for the transformation of unstructured data into structured representation [63]
Data Management	Social Network text	Health data mining	Extracting information and associative features from unstructured text (associative analysis) [64]
Database Analytics		Data linkage and administrative databases	Metadata Management and IE [65]

Entities, relations, events, terms, and other important information can be extracted from text data in healthcare. These subtasks of information extraction are valuable to identify important terms and its associative relations in the medical domain. In healthcare analytics, decision support systems, medicine, and data management area, information extraction deal with several challenges due to unstructured big data. This review identifies those challenges to highlight the potential areas that need solutions for improved healthcare. Following is the summary of challenges:

- Extracting quality insights from EHR clinical notes, documents, clinical discharge summaries, and social media streams for healthcare analytics in treatment and medicine.
- Phenotype and genotype information extraction for disease identification from EHR and biomedical text up to actionable granularity and identify risk score.
- Semantic information extraction from various types of analytics to identify the disease from medical data and advanced intelligent IE systems are required to extract semantic information from multifaceted, open, and diverse unstructured health data.
- Automated extraction of various types of information from EHR clinical notes, EMR, medical and clinical publication text for medical recommendation system and automatic disease prediction is an immense need to represent data into the more elaborative structured form.

- Information extraction from healthcare unstructured data and metadata management can facilitate data linkage and databases. Traditional knowledge-based decision support systems deal with the data that have clear structure [66]. Lack of knowledge-based systems for unstructured data requires advance automatic IE system to build knowledge bases.

3.4 Data Analysis and Representation

Healthcare analytics is comprised of automated computer-based processes and workflows to transform raw health data into meaningful information for effective healthcare decision-making and wellness monitoring. Healthcare analytics is not only facilitating patients but also physicians, clinicians, nursing and other administrative staff. This review explores the challenges of unstructured big data to healthcare where analysis of data needs more accuracy because it is directly related to human life. According to the review, healthcare analytics has been identified as the most critical and most prominent research field and most affected by the heterogeneity and diversity of unstructured data. Table 3 briefly describes the related research.

Application of big data analytics techniques on disease diagnosis, pharmacovigilance, intelligent patient care, phenotyping, disease prediction, and patient health and wellness monitoring is helpful for preventive interventions and treatments. Technological development in healthcare is producing huge volume of different types of data. Patient demographic data, laboratory results, patient history, genomic data, EHR, EMR, PHR, and many other types of data are being produced every day. These data can be structured, semi-structured or unstructured. Due to the heterogeneity, different data formats, no predefined schema or model, and diversity of data sources adding more challenges to unstructured big data analytics in healthcare. This section identifies the potential challenges of unstructured data analytics to predictive, prescriptive, descriptive and diagnostic analytics. According to the literature review, the following is a brief description of the area where unstructured data are critical for analytics and data representation. The challenges of each area are also described categorically.

3.4.1 Smart Healthcare

Most of the data generating by sensors in smart healthcare are unstructured. Sensor-based health and wellness monitoring are generating a huge amount of unstructured data which could not be handled and interpreted manually. This huge amount of unstructured data is generating a gap between its potential and usability. The huge amount of streaming data from sensors is not usable due to heterogeneity and dimensionality of unstructured data. The data analytics pipeline for smart

Table 3 Related Literature regarding Healthcare Analytics and Representation of data

Healthcare category	Data type used	Domain	Description of challenges
Smart Healthcare	Sensor data	Patient Monitoring	Data discovery in heaps of unstructured data where streams of data are coming at very high speed, advanced smart health, and management to automate the curation process and improve the infrastructure for context-aware analytics [67–69]
	Heterogenous dataset text, images, sound	Healthcare Analytics pipeline	
		Analysis of healthcare systems	
Data Representation and Predictive Modeling	Clinical text events	Pharmacovigilance	Summarization and representation of data are facing difficulties due to lack of standardization such as abbreviations. Aggregated data from multiple sources contain diversity, sparsity, and multidimensionality due to data formats and representations. These different representations of EHR data making context-aware analytics challenging [70–73]
Data representation	Clinical Text documents	Phenotyping: Cancer Care	
Data representation	Clinical Notes and EMRs	Congestive heart failure	
Decision Support system	EHR text and Genetic data	Phenotyping using HER	
Healthcare System		Israel Health Policy Research	Due to data quality issues and heterogeneity, unstructured data are considered useless for clean analytics. Missing data and acronyms decreases the efficiency of analytics. Social media data have poor quality and noise, so the analysis of social media data is critical. Hashtag words can be used to improve analytical process but still, it needs health, media, and computer literacy for efficient results [74–79]
Predictive Analytics	EHR clinical narratives	Framingham risk factors identification of Coronary artery disease	
Twitter Analytics	Twitter text	Common disease identification	
eHealth	All	Implementation of eHealth platform	
Healthcare big data		Review of healthcare big data	
Twitter Analytics	Twitter hashtag words	Identification of top diseases	
Text Analytics	EHR Text	Russian language medical data of acute coronary syndrome	The gap between semantic and syntactic structure creating challenges for context understanding of the unstructured text. Although context understanding can improve the results of Analytics. But privacy and security of patient data is also an issue. While extracting maximum information from unstructured data, security issues should also be considered [80–85]
Text Analytics	Clinical Notes	Medical Research	
Privacy and Security	EHR text	Conceptual design of healthcare system	
Visual Analytics	Structured and unstructured text	Visual content correlation analysis	
Predictive Analytics	EHR Text	Eczema Disease identification	
Predictive Analytics	EMR-free clinical text	Coronary Artery Disease identification	

(continued)

Table 3 (continued)

Healthcare category	Data type used	Domain	Description of challenges
Advanced analytics	EMR	Intelligent patient care	Efficient Unstructured Data Analytical techniques are required for technological development which can be achieved using integrative and combined approaches that can handle a variety of data [86–90]
Content Analysis	Medical images, sound, video, text	Big data Analytics Capabilities	
Text Analytics	Biomedical and clinical datasets	Pharmacovigilance: drug safety surveillance	
Text Analytics	Free clinical text	Patient safety event (PSE) reports	
Unstructured Data Analytics	Medical Imaging data	Deep learning	

healthcare applications deals with data discovery and curation, analysis, and result of interpretation phases similar to the traditional analytical process. But in healthcare, good data discovery, data processing, and data management need sound considerations for further development [67, 68]. Good data discovery means right data at the right time for the right context. In order to improve the context awareness in healthcare applications, there is a need to remove the gap between multiple disciplines such as computer science and medical science [69]. Hence to deal with good data discovery, data curation is more helpful to improve understanding of patient physiological and psychological care.

3.4.2 Data Interpretation

Combining structured and unstructured EHR data can improve the performance of predictive analytics. Clinical events in EHR data can be extracted and semantic space classification of similar words can be performed. This combination of structured and unstructured data by concatenating their representation using semantic space for each category of clinical events is more efficient than shared semantic space. Diverse distributed representations of clinical text from EHR using semantic spaces showed pretty notable results for predictive performance [70]. But semantic classification based on similar words can reduce the efficiency for different datasets because there is no standardization for words, and acronym, abbreviations which add more challenges to the semantic classification of data. Pharmacovigilance, phenotyping, and disease identification from health records are examples where different types of information can be extracted for different purposes from the data. EHR, EMR, PHR, and omics data contain rich information for various medical fields but there is a need to explore the ways to leverage these data

for healthcare advancement. In this regard, clinical data and genomic data have been integrated for deep phenotyping of cancer, and the proposed model was evaluated using interviews with domain experts [71]. Since the research has some limitation, that is, nor the representation standard was defined neither the evaluation of the proposed model was on real-time dataset. A comprehensive knowledge-base and precise data modeling can be helpful to leverage unstructured clinical notes of diverse formats across different facilities [72]. Extracting useful data from various types of health records is not sufficient for efficient unstructured data analytics but the interpretation or representation of data is also important for improved unstructured data analytics [73].

3.4.3 Data Quality

Several quality factors have been identified in the literature to improve and asses the quality of big data such as accuracy, completeness, consistency, timeliness, objectivity, interpretability, and accessibility [91, 92]. Unstructured data are adding more challenges to this task due to heterogeneity, lack of structure, noise, no schema, and no predefined model [74, 75, 77, 78]. Social media analytics are helpful for analyzing the psychological issues of the patient or most common diseases of society. Social media analytics is suffering from quality issues more than any other field because social media posts, reviews or comments cannot be standardized. Abbreviations, acronym, short words, and other language ambiguities generate problems for clean analytics [76]. In this regard, hashtag words can improve the efficiency of analytics but still, it needs health, media, and computer literacy to improve healthcare social media analytics [79]. Lack of structure and data heterogeneity and other quality issues can be reduced by continuous data acquisition and data cleaning for efficient healthcare analytics.

3.4.4 Context-Aware Analytics

Understanding the clinical notes in the right context is one of the challenges for unstructured data analytics. Information in the medical domain is coming from various health examiners, each examiner examines in his own way. Hence, understanding the right context is difficult due to non-standardized approaches for unstructured data while aggregation [81]. The structure is one of the important dimensions for unstructured data in terms of quality and context understanding. But the gap between semantic and syntactic structure leads to inefficient results in unstructured data analytics [80]. In this regard, the time-based topic-oriented visualization technique has been designed to find a correlation between unstructured and structured data fields for context-aware analysis [83]. But the study deals with limited text fields, whereas unstructured data in the medical field come in different formats. Also, the study steps to the initial phase of context understanding

by finding the context of written notes using words correlation. More advanced techniques with improved efficiency are required to extract useful information from unstructured clinical notes and other types of data without sacrificing the privacy of the patient [82]. Feature extraction, feature selection, and dimensionality reduction are important areas where advancement for a variety of medical data can significantly improve information extraction and analytics [85]. Anonymization and de-identification can hide the identity of the patient but still revealing or extracting maximum information from an unstructured text can be harmful to the privacy and security of data. Hence, the advancement of approaches is required to shift from data analytics to context-aware analytics by considering the ethical and privacy issues of patient data.

Efficient unstructured data analytical techniques are immense need for technological development which can be achieved using integrative and combined approaches that can handle a variety of data in healthcare [87, 88]. It means data aggregation and unstructured data analytics are important for improved analytics. Cognitive computing where computers and humans can interact in a natural way to predict, analyze, and diagnose using human cognition. Problem-oriented summary of patient EMR with IBM Watson has been proposed and generated a problem list with the help of NLP and ML. Data summarization and semantic search capability are two important factors in cognitive computing [86]. Patient safety event system has been designed where patient safety reports (containing structured and free-text unstructured data) are summarized and search approach was applied to find the most common events [89]. Unstructured healthcare data do not exist only in text form but also MRI, X-ray, and other types of visual data. Using deep learning, these data can be analyzed by several techniques of machine learning [90] but the selection of classification technique(s) for improved accuracy is critical.

3.5 Other Issues

There were some other issues discussed in the literature regarding unstructured data in healthcare.

- The healthcare systems are inadequate to handle data exfiltration due to unstructured data and its constantly evolving nature [93]. Patient data are personal and sensitive data which cannot be shared. But sharing is important to integrate and advanced processing of unstructured data [94]. De-identification and anonymization are used to cure privacy issues, but the deep analysis is required to identify the issues of unstructured big healthcare data and their solutions in technological advancement. Privacy protection of patient data, secure accessibility, extensive query frameworks, and unstructured data management are the challenges that cause inefficiency and inaccuracy of results with a variety of unstructured data in healthcare [95–98]

- Standardization of processes are required to deal with unstructured data quality and care, especially in integration, accessibility, and data extraction [99–103]
- Data curation is more critical for unstructured data than structured data due to several factors such as data quality, usability, heterogeneity, and high dimensionality issues of unstructured data. Automatic data labeling for disease severity in health and wellness monitoring for heterogeneous data [104, 105], semantic data enrichment [106], immense data cleaning of unstructured data for analytics [107], evaluated advanced data curation framework and data fusion techniques for unstructured data [108] are some of the main challenges that need to be considered.

4 Findings

This systematic literature review assessed the potential challenges associated with unstructured data in healthcare. Several literature reviews are available on big data analytics in healthcare, but this literature review is specific to unstructured data in healthcare and explore the challenges being created by unstructured data in healthcare. Best available literature has been identified about the healthcare systems, sources of various types of healthcare unstructured data, domains of healthcare affected by unstructured data, and problems generated by these data. Figure 3 depicts the categorization of unstructured data challenges found in the literature.

The results revealed that unstructured data are critical for healthcare analytics, preprocessing, information extraction, and interoperability, but healthcare analytics is most affected by the unstructured data. These areas need to be addressed to improve healthcare systems. This literature reveals that unstructured data are the most critical challenge for the advancement of healthcare systems because it is affecting almost all areas of healthcare. Following are the main findings of the review:

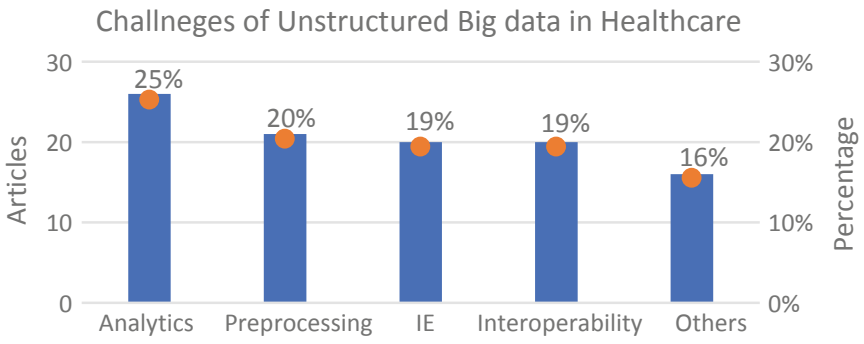


Fig. 3 Challenges of unstructured healthcare data

- Advanced techniques for data preprocessing emerged with big data processing tools are the ultimate requirement for improved healthcare systems. Efficient transformations of unstructured data, semantic data enrichment, profiling, feature selection, cleaning and other activities of preprocessing are to be assessed for multifaceted unstructured data in healthcare. Sequence of preprocessing activities, combining structured and unstructured data, mapping of content, selection of tools are some open issues to develop preprocessing tools in healthcare.
- Semantic and syntactic interoperability is important and essential for context awareness, shared use of heterogeneous data, common representation of combined data, and deep understanding of unstructured data. Standardization of processes and data can reduce the lack of uniform system but privacy and security of patient data should also be considered while achieving interoperability in healthcare systems
- Semantic information extraction in healthcare is a critical challenge to extract disease information from text. It requires a higher accuracy level. Automatic information extraction systems with improved accuracy for unstructured healthcare data with actionable granularity are required to improve the quality of healthcare systems
- Unstructured data analytic techniques are an ultimate requirement of the healthcare industry to improve the quality of care, effective and uninterrupted healthcare services, and diagnostic and decision support tools. Efficient data discovery, context-aware analytics by integrating structured and unstructured data, improving the quality of unstructured data for analysis, and effective interpretation are the ultimate challenges in unstructured data analytics that need to be addressed in future.

Unstructured data exploitation is an important task that can help to reduce the number of technical challenges in healthcare. But the quality, usability, heterogeneity, multidimensionality, scalability, and complexity of unstructured data are making healthcare analytics more challenging. Healthcare systems are unable to achieve higher accuracy and quality in service without considering the importance of unstructured data exploitation.

5 Conclusion

The exponential growth of unstructured data in healthcare forces to adopt more advanced computational systems with more innovative solutions to process the huge volume of unstructured data. Unstructured big healthcare data are rich in content and context as well which will play a pivot role in the development of more advanced and efficient technologies in future healthcare. This review explores the influence of healthcare unstructured data on analytics, clinical decision support systems, health and wellness monitoring, medicine, and smart healthcare.

Various types of unstructured healthcare data such as clinical notes, free-text EHR, handwritten prescriptions, EMR, medical imaging, X-ray, MRI, genomic sequencing, audio, and video medical data are generating challenges in all fields of technology in healthcare. Here, these challenges have been identified and discussed. The results of the literature review have shown that most of the technical challenges of healthcare unstructured data include healthcare analytics, interoperability and integration, preprocessing, and information extraction. Unstructured data have quality and standardization issues which require advanced data cleaning and semantic data enrichment in terms of preprocessing of data. Integration of health data from diverse sources, combining data from various systems to centralize the patient record is future of healthcare which can be achieved using the precise interpretation of data, secure data management, and intelligent healthcare systems. Unstructured data processing and management is an important task in order to improve context-aware analytics and semantic information extraction. The exploitation of unstructured data will maximize the efficiency of healthcare analytics and services.

Acknowledgements This research is funded by Universiti Tunku Abdul Rahman (UTAR) under the UTAR Research Fund (UTARRF): IPSR/RMC/UTARRF/2017-C1/R02.

References

1. Gantz, J., Reinsel, D.: The digital universe in 2020: Big data, bigger digital shadows, and biggest growth in the far east. www.emc.com/collateral/analyst-reports/idc-the-digital-universein-2020.pdf (2012)
2. Turner, V., Gantz, J.F., Reinsel, D., Minton, S.: The digital universe of opportunities: rich data and increasing value of the internet of things. IDC White Paper, No. April, pp. 1–5 (2014)
3. Kiran, A., Vasumathi, D.: Predictive methodology for women health analysis through social media. In: Proceedings of the Second International Conference on Computational Intelligence and Informatics, vol. 712, Springer Singapore, pp. 511–520 (2018)
4. Correia, C., Portela, F., Santos, M.F., Silva, Á.: Data science analysis of healthcare complaints. In: Trends and Advances in Information Systems and Technologies, vol. 747, Springer International Publishing, pp. 176–185 (2018)
5. Kim, K.H., et al.: A text-based data mining and toxicity prediction modeling system for a clinical decision support in radiation oncology: a preliminary study. Korean Phys. Soc. J. **71** (4), 231–237 (2017)
6. Anzaldi, L.J., Davison, A., Boyd, C.M., Leff, B., Kharrazi, H.: Comparing clinician descriptions of frailty and geriatric syndromes using electronic health records: a retrospective cohort study. BMC Geriatr. **17**(1), 1–7 (2017)
7. Saiod, A.K., Van Greunen, D., Veldsman, A.: Electronic health records: benefits and challenges for data quality. In: Handbook of Large-Scale Distributed Computing in Smart Healthcare, Springer, Cham, pp. 123–156 (2017)
8. Gökalp, M.O., Kayabay, K., Akyol, M.A., Koçyiğit, A., Eren, P.E.: Big Data in mHealth. In: Current and emerging mHealth technologies, Springer International Publishing, pp. 241–256 (2018)
9. Austin, C., Kusumoto, F.: The application of Big Data in medicine: current implications and future directions. Interv. Card. Electrophysiol. **47**(1), 51–59 (2016)

10. Angelov, P., Sadeghi-Tehran, P.: A nested hierarchy of dynamically evolving clouds for big data structuring and searching. Procedia Comput. Sci. **53**(1), 1–8 (2015)
11. Kundeti, S.R., Vijayananda, J.: Clinical named entity recognition: challenges and opportunities. In: IEEE International Conference on Big Data (Big Data), pp. 1937–1945 (2016)
12. Raghupathi, W., Raghupathi, V.: Big data analytics in healthcare: promise and potential. Heal. Inf. Sci. Syst. **2**(1), 3 (2014)
13. Liu, M., Hu, Y., Tang, B.: Role of Text Mining in Early Identification of Potential Drug Safety Issues, pp. 227–251. Humana Press, New York, NY (2014)
14. Luo, L., et al.: A hybrid solution for extracting structured medical information from unstructured data in medical records via a double-reading/entry system. BMC Med. Inform. Decis. Mak. **16**(1), 1–14 (2016)
15. van Ooijen, P.M., Jorritsma, W.: Medical imaging informatics in nuclear medicine. In: Quality in Nuclear Medicine. Springer, Cham, pp. 241–267 (2017)
16. Saravana Kumar, N.M., Eswari, T., Sampath, P., Lavanya, S.: Predictive methodology for diabetic data analysis in big data. Procedia Comput. Sci. **50**, 203–208 (2015)
17. Marashi, P.S., Hamidi, H.: Business challenges of big data application in health organization. In: Competitiveness in Emerging Markets. Springer, Cham, pp. 569–584 (2018)
18. Bandyopadhyay, S., et al.: Modeling heterogeneous clinical sequence data in semantic space for adverse drug event detection. In: Data Mining and Knowledge Discovery (2015), p. 31 (2015)
19. Ling, Z.J., et al.: GEMINI: an integrative healthcare analytics system. Proc. VLDB Endow. **7** (13), 1766–1771 (2014)
20. Wang, Y., Kung, L.A., Byrd, T.A.: Big data analytics: understanding its capabilities and potential benefits for healthcare organizations. Technol. Forecast. Soc. Change **126**, 3–13 (2018)
21. Schmidt, D., Budde, K., Sonntag, D., Profitlich, H.J., Ihle, M., Staeck, O.: A novel tool for the identification of correlations in medical data by faceted search. Comput. Biol. Med. **85**, 98–105 (2017)
22. Ong, K.L., De Silva, D., Boo, Y.L., Lim, E.H., Bodi, F., Alahakoon, D., Leao, S.: Big data applications in engineering and science. In: Big Data Concepts, Theories, and Applications. Springer, Cham, pp. 315–351 (2016)
23. Sedghi, E., Weber, J.H., Thomo, A., Bibok, M., Penn, A.M.: A new approach to distinguish migraine from stroke by mining structured and unstructured clinical data-sources. Netw. Model. Anal. Heal. Inf. Bioinf. **5**(1), 30 (2016)
24. Apache Spark™—Unified Analytics Engine for Big Data (online). https://spark.apache.org/. Accessed 09 Oct 2018
25. Apache Hadoop (online). http://hadoop.apache.org/. Accessed 09 Oct 2018
26. Apache Flink: Stateful Computations over Data Streams (online). https://flink.apache.org/. Accessed 09 Oct 2018
27. Gomathi, S., Narayani, V.: Implementing big data analytics to predict systemic lupus erythematosus. In: IEEE Sponsored 2nd International Conference on Innovations in Information, Embedded and Communication systems (ICIIECS), pp. 1–5 (2015)
28. Wu, S.T., et al.: Generality and reuse in a common type system for clinical natural language processing. In: Proceedings of the First International Workshop on Managing Interoperability and Complexity in Health Systems—MIXHS'11, p. 27 (2011)
29. Scheurwegs, E., Luyckx, K., Luyten, L., Daelemans, W., Van den Bulcke, T.: Data integration of structured and unstructured sources for assigning clinical codes to patient stays. J. Am. Med. Informatics Assoc. **23**(e1), 11–19 (2016)
30. Talukder, A.K.: Big data analytics advances in health intelligence, public health, and evidence-based precision medicine. Int. Conf. Big Data Anal. **10721**, 243–253 (2017)
31. Feldman, K., Johnson, R.A., Chawla, N.V.: The state of data in healthcare: path towards standardization. J. Healthc. Inf. Res. **2**(3), 248–271 (2018)

32. Yu, W.D., Kollipara, M., Penmetsa, R., Elliadka, S.: A distributed storage solution for cloud based e-Healthcare Information System. In: IEEE 15th International Conference on e-Health Networking, Applications and Services (Healthcom 2013), pp. 476–480 (2013)
33. Bhaskaran, S., Suryanarayana, G., Basu, A., Joseph, R.: Cloud-enabled search for disparate healthcare data: A case study. In: 2013 IEEE International Conference on Cloud Computing in Emerging Markets, CCEM 2013, pp. 1–8 (2013)
34. Kraus, J.M., et al.: Big data and precision medicine: challenges and strategies with healthcare data. J. Int. Data Sci. Anal. J. **6**(3), 1–9 (2018)
35. Genannt Halfmann, S.S., Mählmann, L., Leyens, L., Reumann, M., Brand, A.: Personalized medicine: What's in it for rare diseases? In: Rare Diseases Epidemiology: Update and Overview, Springer, Cham, pp. 387–404 (2017)
36. Istephan, S., Siadat, M.R.: Unstructured medical image query using big data—an epilepsy case study. J. Biomed. Inform. **59**, 218–226 (2016)
37. Auffray, C., et al.: Making sense of big data in health research: towards an EU action plan. Genome Med. **8**(1), 1–13 (2016)
38. Cuggia, M., Avillach, P., Daniel, C.: Representation of patient data in health information systems and electronic health records. In: Medical Informatics, e-Health, pp. 65–89 (2014)
39. Cruz-Ramos, N.A., Alor-Hernández, G., Sánchez-Cervantes, J.L., Paredes-Valverde, M.A., del Pilar Salas-Zárate, M.: DiabSoft: a system for diabetes prevention, monitoring, and treatment. In: Exploring Intelligent Decision Support Systems, Springer, Cham, pp. 135–154 (2018)
40. Chen, E.S., Sarkar, I.N.: Mining the electronic health record for disease knowledge. In: Biomedical Literature Mining, pp. 269–286 (2014)
41. Wu, S.T., et al.: A common type system for clinical natural language processing. J. Biomed. Semant. **4**(1), 1–12 (2013)
42. da Costa, C.A., Pasluosta, C.F., Eskofier, B., da Silva, D.B., da Rosa Righi, R.: Internet of Health Things: toward intelligent vital signs monitoring in hospital wards. Artif. Intell. Med. **89**, 61–69 (2018)
43. Kozák, J., Nečaský, M., Dědek, J.: Linked open data for healthcare professionals. In: Proceedings of International Conference on Information Integration and Web-based Applications and Services, p. 400 (2013)
44. Ilyasova, N., Kupriyanov, A., Paringer, R., Kirsh, D.: Particular use of BIG DATA in medical diagnostic tasks. Pattern Recognit. Image Anal. **28**(1), 114–121 (2018)
45. Goh, W.P., Tao, X., Zhang, J., Yong, J.: Decision support systems for adoption in dental clinics: a survey. Knowl. Based Syst. **104**, 195–206 (2016)
46. Leyens, L., Reumann, M., Malats, N., Brand, A.: Use of big data for drug development and for public and personal health and care. Genet. Epidemiol. **41**(1), 51–60 (2017)
47. Malmasi, S., Hosomura, N., Chang, L.-S., Brown, C.J., Skentzos, S., Turchin, A.: Extracting healthcare quality information from unstructured data. In: AMIA… Annual Symposium Proceedings/AMIA Symposium, pp. 1243–1252 (2017)
48. Martínez, P., Martínez, J.L., Segura-Bedmar, I., Moreno-Schneider, J., Luna, A., Revert, R.: Turning user generated health-related content into actionable knowledge through text analytics services. Comput. Ind. **78**, 43–56 (2016)
49. Sundararaman, A., Valady Ramanathan, S., Thati, R.: Novel approach to predict hospital readmissions using feature selection from unstructured data with class imbalance. Big Data Res. **1**, 1–11 (2018)
50. Delespierre, T., Denormandie, P., Bar-Hen, A., Josseran, L.: Empirical advances with text mining of electronic health records. BMC Med. Inform. Decis. Mak. **17**(1), 1–15 (2017)
51. Wilcox, A.B.: Leveraging electronic health records for phenotyping. In: Translational Informatics. Springer, London, pp. 61–74 (2015)
52. Simmons, M., Singhal, A., Lu, Z.: Text mining for precision medicine: bringing structure to EHRs and biomedical literature to understand genes and health. In: Translational Biomedical Informatics, vol. 939, pp. 139–166 (2016)

53. Goodman, K., Krueger, J., Crowley, J.: The automatic clinical trial: leveraging the electronic medical record in multisite cancer clinical trials. Curr. Oncol. Rep. **14**(6), 502–508 (2012)
54. Kotfila, C., Uzuner, Ö.: A systematic comparison of feature space effects on disease classifier performance for phenotype identification of five diseases. J. Biomed. Inform. **58**, S92–S102 (2015)
55. Alnashwan, R., Sorensen, H., O'Riordan, A., Hoare, C.: Accurate classification of socially generated medical discourse. J. Int. Data Sci. Anal., pp. 1–13 (2018)
56. Husain, S.S., Kalinin, A., Truong, A., Dinov, I.D.: SOCR data dashboard: an integrated big data archive mashing medicare, labor, census and econometric information. J. Big Data **2**(1), 13 (2015)
57. Jackson, R., et al.: CogStack-experiences of deploying integrated information retrieval and extraction services in a large National Health Service Foundation Trust Hospital. BMC Med. Inf. Decis. **18**(1), 47 (2018)
58. Dinov, I.D.: Methodological challenges and analytic opportunities for modeling and interpreting Big Healthcare Data. Gigascience **5**(1), 1–15 (2016)
59. Hu, B.V., Terrazas, B.: Building a mental health knowledge model to facilitate decision support. In: Knowledge Management and Acquisition for Intelligent Systems, vol. 9806, pp. 198–212. Springer, Cham (2016)
60. Pulmano, C.E., Estuar, M.R.J.E.: Towards developing an intelligent agent to assist in patient diagnosis using neural networks on unstructured patient clinical notes: inaccurate classification of socially generated medical discourseitial analysis and models. Procedia Comput. Sci. **100**, 263–270 (2016)
61. Norman, B., Davis, T., Quinn, S., Massey, R., Hirsh, D.: Automated identification of pediatric appendicitis score in emergency department notes using natural language processing. In: 2017 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI), pp. 481–484 (2017)
62. Chapman, W.W., Gundlapalli, A.V., South, B.R., Dowling, J.N.: Natural language processing for biosurveillance. In: Infectious Disease Informatics and Biosurveillance, vol. 27, pp. 279–310 (2011)
63. Jonnalagadda, S.R., Adupa, A.K., Garg, R.P., Corona-Cox, J., Shah, S.J.: Text mining of the electronic health record: An information extraction approach for automated identification and subphenotyping of HFpEF patients for clinical trials. J. Cardiovasc. Transl. Res. **10**(3), 313–321 (2017)
64. Kim, J.C., Chung, K.: Associative feature information extraction using text mining from health big data. Wirel. Pers. Commun. **105**(2), 691–707 (2018)
65. Clark, A., Ng, J.Q., Morlet, N., Semmens, J.B.: Big data and ophthalmic research. Surv. Ophthalmol. **61**(4), 443–465 (2016)
66. Syomov, I.I., Bologva, E.V., Kovalchuk, S.V., Krikunov, A.V., Moiseeva, O.M., Simakova, M.A.: Towards infrastructure for knowledge-based decision support in clinical practice. Procedia Comput. Sci. **100**, 907–914 (2016)
67. Sakr, S., Elgammal, A.: Towards a comprehensive data analytics framework for smart healthcare services. Big Data Res. **4**, 44–58 (2016)
68. Lee, C., Murata, S., Ishigaki, K., Date, S.: A data analytics pipeline for smart healthcare applications. In: Sustained Simulation Performance 2017. Springer International Publishing, pp. 181–192 (2017)
69. Pramanik, M.I., Lau, R.Y.K., Demirkan, H., Azad, M.A.K.: Smart health: big data enabled health paradigm within smart cities. Expert Syst. Appl. **87**, 370–383 (2017)
70. Henriksson, A., Zhao, J., Dalianis, H., Boström, H.: Ensembles of randomized trees using diverse distributed representations of clinical events. BMC Med. Inform. Decis. Mak. **16** (Suppl 2), 69–79 (2016)
71. Hochheiser, H., Castine, M., Harris, D., Savova, G., Jacobson, R.S.: An information model for computable cancer phenotypes. BMC Med. Inform. Decis. Mak. **16**(1), 1–15 (2016)
72. Wang, Y., et al.: NLP based congestive heart failure case finding: a prospective analysis on statewide electronic medical records. Int. J. Med. Inf. **84**(12), 1039–1047 (2015)

73. Jackson, K.L., et al.: Performance of an electronic health record-based phenotype algorithm to identify community associated methicillin-resistant Staphylococcus aureus cases and controls for genetic association studies. BMC Infect. Dis. **16**(1), 1–7 (2016)
74. Lovis, C., Gamzu, R.: Big Data in Israeli healthcare: hopes and challenges report of an international workshop. Isr. J. Health Policy Res. **4**(1), 4–9 (2015)
75. Jonnagaddala, J., Liaw, S.T., Ray, P., Kumar, M., Chang, N.W., Dai, H.J.: Coronary artery disease risk assessment from unstructured electronic health records using text mining. J. Biomed. Inform. **58**, S203–S210 (2015)
76. Bamwal, A.K., Choudhary, G.K., Swamim, R., Kedia, A., Goswami, S., Das, A.K.: Application of twitter in health care sector for India. 2016 3rd International Conference on Recent Advanced Information Technology, pp. 172–176 (2016)
77. Rinaldi, G.: An introduction to the technological basis of eHealth. In: eHealth, Care and Quality of Life. Springer Milan, pp. 31–67 (2014)
78. Persico, V.: Big data for health. In: Encyclopedia of Big Data Technologies. Springer International Publishing, pp. 1–10 (2018)
79. Grover, P., Kar, A.K., Davies, G.: 'Technology enabled Health'—Insights from twitter analytics with a socio-technical perspective. Int. J. Inf. Manage. **43**(May), 85–97 (2018)
80. Metsker, O., Bolgova, E., Yakovlev, A., Funkner, A., Kovalchuk, S.: Pattern-based mining in electronic health records for complex clinical process analysis. Procedia Comput. Sci. **2017**(119), 197–206 (2017)
81. Khatri, I., Shrivastava, V.K.: A survey of big data in healthcare industry. Adv. Comput. Commun. Technol. **452**, 245–257 (2016)
82. Sarkar, B.K.: Big data for secure healthcare system: a conceptual design. Complex Intell. Syst. **3**(2), 133–151 (2017)
83. Wei, F., et al.: Visual content correlation analysis. In: Proceedings of the first International Workshop on Intelligence Visual Interfaces for Text Analysis—IVITA'10, no. 1, p. 25 (2010)
84. Jayalatchumy, D., Thambidurai, P.: Prediction of diseases using Hadoop in big data—a modified approach. In: Artificial Intelligence Trends in Intelligent Systems. Springer, Cham, pp. 229–238 (2017)
85. Buchan, K., Filannino, M., Uzuner, Ö.: Automatic prediction of coronary artery disease from clinical narratives. J. Biomed. Inform. **72**, 23–32 (2017)
86. Devarakonda, M.V., Mehta, N.: Cognitive computing for electronic medical records. In: Healthcare Information Management Systems, pp. 555–577 (2016)
87. Wang, Y., Kung, L.A., Wang, W.Y.C., Cegielski, C.G.: An integrated big data analytics-enabled transformation model: application to health care. Inf. Manag. **55**(1), 64–79 (2018)
88. Maitra, A., Annervaz, K.M., Jain, T.G., Shivaram, M., Sengupta, S.: A novel text analysis platform for pharmacovigilance of clinical drugs. Procedia Comput. Sci. **36**, 322–327 (2014)
89. Fong, A., Hettinger, A.Z., Ratwani, R.M.: Exploring methods for identifying related patient safety events using structured and unstructured data. J. Biomed. Inform. **58**, 89–95 (2015)
90. Singh, N., Singh, S.: Object classification to analyze medical imaging data using deep learning. In: 2017 International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), pp. 1–4 (2017)
91. Batini, C., Cappiello, C., Francalanci, C., Maurino, A.: Methodologies for data quality assessment and improvement. ACM Comput. Surv. **41**(3), 1–52 (2009)
92. Wahyudi, A., Kuk, G., Janssen, M.: A process pattern model for tackling and improving big data quality. Inf. Syst. Front. **20**(3), 457–469 (2018)
93. Ullah, F., Edwards, M., Ramdhany, R., Chitchyan, R., Babar, M.A., Rashid, A.: Data exfiltration: a review of external attack vectors and countermeasures. J. Netw. Comput. Appl. **101**, 18–54 (2018)
94. Wuyts, K., Verhenneman, G., Scandariato, R., Joosen, W., Dumortier, J.: What electronic health records don't know just yet. A privacy analysis for patient communities and health records interaction. Health Technol. (Berl) **2**(3), 159–183 (2012)

95. Istephan, M.R., Siadat, S.: Extensible query framework for unstructured medical data—a big data approach. In: IEEE International Conference on Data Mining Workshop (ICDMW), pp. 455–462 (2016)
96. Tchagna Kouanou, A., Tchiotsop, D., Kengne, R., Zephirin, D.T., Adele Armele, N.M., Tchinda, R.: An optimal big data workflow for biomedical image analysis. Inf. Med. Unlocked **11**, 68–74 (2018)
97. Meystre, S.M.: De-identification of unstructured clinical data for patient privacy protection. In: Medical Data Privacy Handbook. Springer, Cham, pp. 697–716 (2015)
98. Aqeel-ur-Rehman, Khan, I.U., ur Sadiq ur Rehman, S.: A review on big data security and privacy in healthcare applications. In: Big Data Management. Springer International Publishing, Cham, pp. 71–89 (2017)
99. Gaylis, F., Cohen, E., Calabrese, R., Prime, H., Dato, P., Kane, C.J.: Active surveillance of prostate cancer in a community practice: how to measure, manage, and improve? Urology **93**, 60–66 (2016)
100. Hardy, L.R., Bourne, P.E.: Data science: transformation of research and scholarship. In: Big Data-Enabled Nursing. Springer, Cham, pp. 183–209 (2017)
101. Khennou, F., Khamlichi, Y.I., El Houda Chaoui, N.: Designing a health data management system based hadoop-agent. In: 4th IEEE International Colloquium on Information Science and Technology (CiSt), pp. 71–76 (2016)
102. Vest, J.R., Grannis, S.J., Haut, D.P., Halverson, P.K., Menachemi, N.: Using structured and unstructured data to identify patients' need for services that address the social determinants of health. Int. J. Med. Inf. **107**(August), 101–106 (2017)
103. Hong, N., et al.: Integrating structured and unstructured EHR data using an FHIR-based type system: a case study with medication data. AMIA Joint Summits on Translational Science Proceedings, vol. 2017, pp. 74–83 (2018)
104. Rastegar-Mojarad, M., et al.: Using unstructured data to identify readmitted patients. In: IEEE International Conference on Healthcare Informatics (ICHI), pp. 1–4 (2017)
105. Boursalie, O., Samavi, R., Doyle, T.E.: Machine learning and mobile health monitoring platforms: a case study on research and implementation challenges. J. Healthc. Inf. Res. **2**(1–2), 179–203 (2018)
106. Zillner, S., Neururer, S.: Technology roadmap development for big data healthcare applications. KI Künstliche Intelligenz **29**(2), 131–141 (2015)
107. Giambrone, G.P., Hemmings, H.C., Sturm, M., Fleischut, P.M.: Information technology innovation: the power and perils of big data. Br. J. Anaesth. **115**(3), 339–342 (2015)
108. Banos, O., et al.: An innovative platform for person-centric health and wellness support. Int. Conf. Bioinf. Biomed. Eng. **9044**, 31–140 (2015)

Zip Zap Data—A Framework for 'Personal Data Preservation'

K. Arunkumar and A. Devendran

Abstract In the era of Mobile devices, a huge volume of mobile applications produces vast and variety of data in their own formats. There is every chance that most of these data will be lost or gone forever. Let us explain how, technically any digital data becomes irretrievable if there is a change of data format or a change in application interpreting it or a change in platform feature used by the application and your app did not change to accommodate it. In simple terms, if you try to upgrade your mobile—there are a lot of possibilities that data from your previous mobile could be lost forever. Reason for loss could be your newer device comes with some obsolete platform features and hence some of your applications no longer run in the new device as developer decided to discontinue that app. Even more pressing issue is there is no way to transfer application data from old to a new device and this is very common with mobile apps. The root cause for this is the lack of backup and retrieval mechanisms in devices and applications. We studied various approaches and research works related to this problem and proposed 'Zip Zap Data', a framework for effective backup and recovery of personal digital data.

Keywords Personal data longevity · Personal data preservation

1 Introduction

We are seeing a huge increase in the smartphone adoption across the globe. This is the main reason we see millions of mobile applications in different mobile platform (Android, iPhone, and Windows). These applications were developed by enormous developers/companies across the globe. Most of the applications produce a lot of

K. Arunkumar
Ppltech, Chennai, India
e-mail: arunkumar@gmail.com

A. Devendran (✉)
Dr. M.G.R Education and Research Institute, Chennai, India
e-mail: devendran.alagarsamy@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_23

data related to users, which can be of value to the user or not. Some of them could be really confidential and private. Since there is no standard or protocol in restricting the data representation inside the application, every developer is storing data in their own tool and formats. What it means is even if you get the raw application data file from the device somehow, one cannot make use of it easily. It is only their application code which can access the data and deserialize it properly. The schema of data representation and security of accessing it are the reasons. This part closely ties data with apps on smartphones.

So what is the problem? One fine day when the app developer decides to stop supporting his app in play store, all data of that app is almost gone or at least locked with that current device. There is no way you can get it out meaningfully.

Practical problem: Moving/Upgrading to a new smartphone.

Many of my friends including me have faced the below-stated problem when upgrading to a new model of smartphone. Google, Mac, and device manufacture helps to migrate your contacts, gallery, and platform apps data. But there is no easy way you can get all your apps, especially with data. For example, for the popular messaging app 'WhatsApp' you have to configure your backup option and do the backup functionality regularly. Alternatively, there is an option for exporting data to your pc and later you can import back in the new device. If this is a way for migrating data to a famous app with $\sim$ 00 million active users, consider the case of others (especially free ones) out there in the market. Hundreds and thousands of apps out there in play store are left orphans for different reasons.

Security of apps data: Most of the mobile platforms runs every app in a sandbox to ensure the safety of system resources and data. Following diagrams (Fig. 1) shows the sandboxing in iPhone. Every application needs users permission to access additional system resources like using storage, camera, internet(wifi, mobile data), bluetooth, etc. Also, data from platform apps like SMS, Gallery, Calendar, etc.

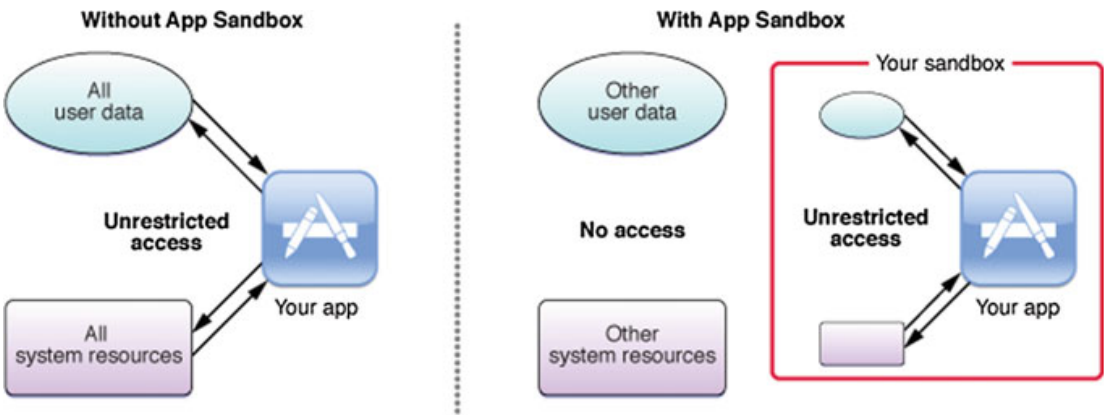


Fig. 1 iPhone App Sandbox model for data security

1.1 Digital Data Preservation

Having said the problem in accessing and interpreting mobile data, can solution to the well know problem of digital realm 'Digital data preservation' help us here?

What is digital preservation? In the definition medium from American Library Association(ALA) => 'The goal of digital preservation is the accurate rendering of authenticated content over time.'

In other words, Digital preservation is the mean of ensuring access of digital data across time (10s, 100s or even 1000s of years). Clearly, this impacts the future of 'history of events', the future of knowledge, culture, evidence, etc.

Couple of statements to understand the seriousness of 'Data Preservation'

1. In his interview on BBC @ 13 February 2015—Vint Cerf, one of 'the fathers of the Internet', expressed his concern that today's digital content might be lost forever. His talk mainly points on the fact that—if technology continues to outpace preservation tactics, future citizens could be locked out of accessing today's digital content enter a 'digital dark age'.
2. "Digital data lasts forever or 5 years, whichever comes first". This is from a Scientific American journal released back in 1995 written by Jeff Rothenberg [10].

Discussing in detail about digital longevity is out of scope.

Fundamental problem with digital data: A sequence of bits is meaningless if it cannot be decoded and transformed to some meaningful representation and viewability. Example: consider documents created with 'WordStar' (a 1980s software for data handling), are unreadable today as there is no software to interpret this data.

Another important question is what to preserve? Clearly, we need to preserve digital data for our future generation. Should all details available in the digital realm be preserved? If not, what should be preserved? Who will define the archival selection—what and whose story gets told, and whose do not. Again discussion on this direction is out of the scope of our work.

The main concern here is the data is being generated by a wide range of electronic devices and the quantities of data being generated are really huge. Now, digital data comes from personal computers, smartphones, game devices, digital cameras, digital recorders, digital TVs, and so on. The main cause for this exploded data quantity is the increase in dependency of our modern society on digital information and devices. Without proper strategies and technology support for preservation of all these data are under huge risk of being lost.

2 Core Problem

In digital preservation paradigm, one of the most difficult challenges is to maintain the interpretability of data across changing technologies. As the sequence of bits is meaningless if it cannot be decoded and transformed into some meaningful representation.

2.1 Internet Era Data Problem

Most of the modern world data be it books, images, videos, music exist as digital data. All these are represented and stored in different formats. Some of these data formats are standard and others are proprietary (purely dependent on software generating them). In the internet-connected world, anyone with proper privileges can access them instantly. The data generated can be static (created and stored once) or dynamic (generated on every time a request for access comes).

The irony is even though most of the content which could be static is often dynamically generated(i.e., through execution of code to create content at runtime).

The challenge is how do we preserve these data over time? It is no longer the three static object or collection of static objects to be preserved. It is the entire stack of technology that is serving content along with its data needs to be frozen and preserved., i.e., We need to freeze and precisely reproduce the execution that dynamically produces the content.

What is the problem there? Preservation and reproduction of software execution are one of the most complex problems. Because this involves perfect alignment of many moving parts to make it work, like OS support for older system libraries, dlls, implementation specific libraries, and hardware and firmware. Preserving this alignment over time is difficult. Many things can change hardware, operating system, linked libraries, configuration and user preferences, geographic location, execution timing, and so on. Even a single change may completely break execution.

2.2 Mobile Devices Data Problem

Across the globe, 52.3% of internet traffic comes from mobile devices and it has been increasing year by year. The study and inference on the mobile device adoption and usage show it [12]. And for data generation, even a greater percentage of the share goes to mobile. Mobile data can be generated by native apps or mobile web apps running in a browser. Recently, we see a shift towards PWA (progressive web app) from native apps, coz of its lightness and features or maybe google is

batting for it. Technically speaking, PWA is more of an optimized mobile web application with a shortcut in app screen.

In general, we can group apps under 2 categories, from data generation point. Both brings its own complexity.

- Native mobile apps (running in sandbox)
- PWA or Web apps (running in browser sandbox)

2.3 Personal Data Preservation

Digital data preservation is a long and still open research area. There are various solutions proposed and implemented till date. All the solutions can be broadly classified into 2 categories (migration and emulation), based on the strategy to ensure longevity [11]. We restricted our focus on 'Personal data preservation' [14] using the emulation approach and proposed a framework for preserving personal data from modern devices especially smartphones.

2.4 Data Privacy—Huge Concern

Data privacy is a huge concern these days. Every big corporate with personally identifiable information or other sensitive information about users is investing lots in ensuring the data privacy of its data. When you should worry—You are collecting, storing, using above mentioned user data – in digital form or otherwise. Consider your healthcare records, financial transactions, biological traits, such as genetic material, residence and geographic records in wrong hands. The challenge of data privacy is to utilize data while protecting individual's privacy preferences and their personally identifiable information.

2.5 Secured Documents Problem

The simplest way to achieve secrecy and longevity is to store the secret in a distributed way on many servers and put a constraint that a minimum number of shares be required to reconstruct the secret. From a long term archival point of view, any retrieval of such secured document needs some piece of software along with the data for the construction of secret key from k or more pieces. There a is lot of research work happening in this area.

3 Related Works

3.1 Parity Cloud Service: A Privacy Protected Personal Data Recovery Service

Here the authors mainly talk about the privacy protection of data in the cloud, an important problem in providing personal data recovery as a service. They proposed a novel mechanism 'data recovery service framework' built on cloud infrastructure. PCS (Parity Cloud Service) can be easily set up on any cloud provider. The main differentiating factor in the proposed framework is that zero user-specific data to be uploaded/saved to the server for data recovery.

3.2 Algorithm for Backup and Recovery of Data Stored on Cloud Along with Authentication of the User

Authors of this work propose a scheme of storing data on the cloud server along with a copy in the remote server. They propose the process of taking a backup copy in remote server regularly. If the main file is lost, it can be recovered from the remote server. To make backup and recovery more secure, user authentication based on attributes, secret keys are used. Every transaction in the system is authenticated per above scheme for security.

3.3 Securing BIG Storage: Present and Future

In this work, the authors reviewed existing works in cloud computing technologies and security threats of BIG DATA storage in cloud environments. Also, existing solutions that address security threats. They also discussed Internet of Things (IOT) and its influence in BIG DATA and propose solution to various threats identified.

3.4 A Framework for Aggregating Private and Public Web Archives

Here the authors talk about the problem of existing public web archive systems and HTTP memento RFC-7089 [13]. The problem with users personalized web pages (Facebook like), feeds and even instance of a private web application (banking like) in existing archiving systems and ways to mitigate them. They worked on amending existing Memento and Timestamp syntaxes of HTTP archiving to accommodate the private data.

3.5 Syncing Mobile Applications with Cloud Storage Services

The authors analyzed various cloud providers APIs and in the context of mobile application data sync. Their experiments show various options for mobile app developers to choose among the providers to sync their app data across multiple devices.

3.6 Commercial Products

Most of the commercial applications doing entire app backup and restore with data need a rooted device for accessing the data specific to the application.

Following list of applications fall in this category:

Helium, MyBackup Pro, **Titanium Backup, Ibackup**, etc.

Google Backup & Restore—is a recent feature from google, which will allow the user to store some of users app data like calendar, Browser, Contacts, Gmail, Photos, and Music to his google drive account and retrieve it when needed. However, this is not a complete solution for the backup/restore.

3.7 Limitation of Existing Backup/Restore Apps

1. They need a rooted device (so that root permission can be granted to the backup app)
2. Google vendor lock and not a complete backup and restore for everything
3. Some of them only backup apps (not the data generated by them) and common user data like contacts, photos, videos, downloads, etc.

4 Our Contribution

The fundamental problem with the existing backup and restore system is closely tied with the security model of android applications. All the files created in an app are stored by default in app's own sandbox, which is not visible to any other app running in the device. Unless you have an app running with root privilege, nothing can see the data created by it.

4.1 Core Idea

The novelty comes in bringing the backup/restore action to the app level and breaking the security barrier of the android system. Also, another important part of the idea is storing data in end user cloud storage and not in any app developer or company cloud storage, which opens doors for privacy vulnerabilities.

To put it simple Zip Zap Data

1. Store user data in the user's cloud space and retrieve it when needed.
2. Store app data along with the app and its metadata like device model, OS, etc.
3. All the data is compressed and encrypted before storing and applied in reverse order @ retrieval time.

4.2 Zip Zap Data Framework

Application specific data:

Most of the useful mobile application has some data starting from user preferences, game scores, to do notes to complex databases. As an app developer, depending on the use case one can store data in a mobile device or some external cloud storage. Some apps were built to operate 100% online and store all its data in the cloud. Even those apps started storing part or some data in a mobile device to provide offline support.

This offline feature in most applications gives seamless behavior to users and crucial for their user engagement in developed/developing countries. The point we try to make here is—every modern app being developed have a lot of application-specific data. It is this data residing in the device needs to be backed up and restored properly to give seamless transition when upgraded.

Synchronize data across multiple device:

These days, it is very normal for a user to have more than one device and very likely that he/she uses the same application in all of them. As an end user—I expect the app should be intelligent to be in the same state across all my devices. This use case brings its own complexity to app developers to write special mechanism to synchronize data across multiple devices of the same user. Even though our work focused on the data backup/retrieval part, a properly implemented system could achieve synchronization across devices with our framework.

4.3 Data in Mobile Application

Almost all the apps even the ones storing data in cloud stores data in device memory (for storing user preferences or to give offline capability). Let us get into

implementation details. As an app developer, storing data in a mobile device can be done in three ways.

- Shared preferences: This is a key, value store provided by android SDK. It is straight forward to use. There are ways in which you can make this data visible outside application sandbox. One has to set 'Context. MODE_WORLD_READABLE' or 'Context.MODE_WORLD_WRITABLE' on the preference for this to work.
- Local files: Any mobile application can create/update/delete raw files (text or any binary files) with standard file system calls. The files created are completely inside the sandbox of application and no other application can see this. These files cannot survive multiple installations even on the same device.
- Local Databases: From a technical point of view, this is very similar to local file case. All the raw files of DB are stored within the sandbox of app.

5 SDK for Backup & Recovery

As developing backup/retrieve capabilities require a considerable time and additional skills on the part of developers. If an easy to use and off the shelf utility is available for this purpose, developers would be too happy to adapt it. Hence, we developed an open-source SDK for this purpose. Zip Zap Data Framework allows developer separate interfaces to handle these three cases of app data stored in devices mentioned above in Sect. 4.3. This SDK solves the fundamental backup/recovery problem along with other benefits listed below.

Solving—Data preservation problem:

As we are storing 'app metadata' along with the data, we are ensuring the retrieval of the data stored any time in future, i.e., the data is readable 'simply…forever'. The fundamental philosophy is to recreate the original environment [11] for the app and its data @ the time of retrieval. Worst cases might involve making a new application or providing dependencies to a lost/obsolete library component.

Solving—App data sync problem:

Integrate app data sync feature into your android, iOS and web apps with just a few lines of code without any server setup or cost. Enable your users to have a seamless experience across device usage. Allow him to move to new devices without any issue of data loss.

Unlike existing cloud provider solutions like parse or azure or others, which require payment from developers for the sync feature, the proposed solution won't require any cost from the developer. This is possible as we are going to use end user's cloud storage (like dropbox, Google Drive, iCloud, SkyDrive, etc.) for storage part and also the data privacy part is taken care.

Avoiding cloud vendor lock in problem:

The key thing in our proposal is to store user's data in user's own cloud space (like dropbox, Google Drive, iCloud, SkyDrive, etc.), which avoids any vendor locking for server-side implementation for giving sync mechanism. As we are storing data in compressed and encrypted form, concerns on the privacy of user data and security are addressed.

5.1 How It Works?

1. The cloud storage of the user is configured in that app once for every device.
2. The SDK part running in every application identifies and send application data be it local files or shared preferences or sqllite DB to the backup and restore app regularly.
3. For optimized data transfer from device to cloud storage, we have version based files maintained across all apps using SDK.
4. The delta of the data to or from the cloud is identified as mentioned above and transfer is done when the user goes online.

5.2 Zip Zap Data Sync Backup and Restore

Zip Zap Data Sync is the open-source app that backups & restores apps and app data. This app needs no rooted devices to backup and restores user app-specific data. It gets the data from the app through SDK and uses it for backup/restore.

Along with the app data, user can get the entire list of installed apps in the device. There is a provision to maintain and retrieve multiple versions of apps something like marking a checkpoint and retrieve from that stable point.

6 Integration and Results

6.1 Integration Steps

SDK is available as an archive file (source code for the same is available in github). Integration can be done in a few steps after configuring the file, shared preference or DB details. It can be done in below four steps.

1. Import the 'Zip Zap SDK.aar' file as a new module
2. Add a dependency in the application gradle file 'compile project(':zipzapsdk')'
3. For restoring data or fetching data from cloud storage use '**sync**.getZipZapData();'
4. For data backup or store data to cloud storage use '**sync**.putZipZapData();'

6.2 Open-Source Integrations

We have integrated our SDK with some open-source android application and made them available in git after our integration [9]. Now, these apps have Backup/Restore feature along with seamless data sync across multiple devices of the same user.

Application	Local data	Usage
AndroFish	Shared Preference	A simple game, where user can guide fish to eat the other fish and grow bigger
Android-PreferencesManager	Shared Preference	Preferences Manager is an open-source application that allows you to seamlessly edit application's preferences
Android-pedometer	Local File	App for Android phones that counts your steps
Simple-Draw	Local File	A canvas you can draw on with different colors
LeeCo	Local File	LeeCo is an awesome app for (including unlock) problems, solutions, discuss(from leetcode) and comments
todo.txt-android	Local File	Official Todo.txt Android app for managing your todo.txt file stored in Dropbox

(continued)

(continued)

Application	Local data	Usage
Design game-pegboard	Local File	A simple design game with pins—to demonstrate the usage of ZipZapData SDK
OpenSudoku	Local DB	open-source Sudoku game for Android
CoCoin	Local DB	CoCoin, Multiview Accounting Application
Runnerup	Local DB	An open-source run tracker for Android
Simple-Notes	Local DB	A simple textfield for adding quick notes without ads

7 Conclusion

The main idea of the proposal is moving the responsibility of backup/restore data of an application to application itself rather than a separate stand-alone app on its own. This is working well and addresses the fundamental security model. This works well for real time app sync across the devices. With our integration on 10 + open-source applications, we observed backup/restore functionality can be added with few lines of code. App sync is a bit tricky as put Zip Zap Data () and get Zip Zap Data () needs to be called in proper life cycle of the corresponding activities. But we have seen it working well in 'OpenSudoku' game.

The current implementation of android app 'Zip Zap Data Sync', giving the feature of synchronize of device files and user's cloud storage is more like the – 'One click to Backup all' feature most user will like to have. Also, the storage choice of users cloud in encrypted form adds great value to the system by addressing the data privacy concerns.

We are planning to integrate more android apps with our SDK both open-source and commercial apps. Our development of SDK for iOS and Windows platform is in progress and will be available in github shortly. Another addition will be the flexibility of user to create multiple checkpoints for user data being preserved similar to memento work, which will let the user replay various instances of data with the application at any point of time.

References

1. https://en.wikipedia.org/wiki/Information_privacy
2. Gupta, V.H., Gopinath, K.: An extended verifiable secret redistribution protocol for archival systems. In: International Conference on Availability, Reliability and Security (ARES) (2006)
3. Song, C., Park, S., Kim, D., Kang, S.: Parity Cloud service: a privacy-protected personal data recovery service. In: IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications (2011)

4. Raje, M., Mukhopadhyay, D.: Algorithm for back-up and recovery of data stored on cloud along with authentication of the user. In: International Conference on Information Technology (ICIT) (2015)
5. Bhargavi, I., Padmaja, M., Ammayappan, K.: Securing BIG storage: present and future. In: Online International Conference on Green Engineering and Technologies (IC-GET) (2016)
6. Kelly, M., Nelson, M.L., Weigle, M.C.: A framework for aggregating private and public web archives. In: JCDL '18 Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries (2018)
7. Pocatilu, P., Boja, C., Ciurea, C.: Syncing Mobile Applications with Cloud Storage Services. Inform. Econom. **17**(2/2013) (2013)
8. http://www.techradar.com/how-to/phone-and-communications/how-to-backup-your-android-device-1298811
9. ZipZapData SDK and Sample Apps. https://github.com/zipzapdat/
10. Rothenberg, J.: Ensuring the longevity of digital documents. Sci. Am. **272**(1), 42–47 (1995)
11. Arunkumar, K., Devendran A.: Digital data preservation—a viable solution. In: Data Management, Analytics and Innovation, pp 129–141 (2018)
12. https://www.statista.com/statistics/241462/global-mobile-phone-website-traffic-share/
13. https://tools.ietf.org/html/rfc7089
14. Poursardar, F., Shipman, F.: What is part of that resource? User expectations for personal archiving. In: JCDL '17 Proceedings of the 17th ACM/IEEE on Joint Conference on Digital Libraries (2017)

A Systematic Mapping Study of Cloud Large-Scale Foundation—Big Data, IoT, and Real-Time Analytics

Isaac Odun-Ayo, Rowland Goddy-Worlu, Temidayo Abayomi-Zannu and Emanuel Grant

Abstract Cloud computing is a unique concept which makes analysis and data easy to manipulate using large-scale infrastructure available to Cloud service providers. However, it is sometimes rigorous to determine a topic for research in terms of Cloud. A systematic map allows the categorization of study in a particular field using an exclusive scheme enabling the identification of gaps for further research. In addition, a systematic mapping study can provide insight into the level of the research that is being conducted in any area of interest. The results generated from such a study are presented using a map. The method utilized in this study involved analysis using three categories which are research, topic, and contribution facets. Topics were obtained from the primary studies, while the research type such as evaluation and the contribution type such as tool were utilized in the analysis. The objective of this paper was to achieve a systematic mapping study of the Cloud large-scale foundation. This provided an insight into the frequency of work which has been carried out in this area of study. The results indicated that the highest publications were on IoT as it relates to model with 12.26%; there were more publications on data analytics as is relates to metric with 2.83%, more articles on big data in terms of tool, with 11.32%, method with 9.43% and more research carried out on data management in terms of process with 6.6%. This outcome will be valuable to the Cloud research community, service providers, and users alike.

I. Odun-Ayo (✉) · R. Goddy-Worlu · T. Abayomi-Zannu
Department of Computer and Information Sciences,
Covenant University, Ota, Nigeria
e-mail: isaac.odun-ayo@covenantuniversity.edu.ng

R. Goddy-Worlu
e-mail: rowland.goddy-worlu@stu.cu.edu.ng

T. Abayomi-Zannu
e-mail: temidayoabayomizannu@gmail.com

E. Grant
School of Electrical Engineering and Computer Science,
University of North Dakota, Grand Forks, USA
e-mail: emanuel.grant@engr.und.edu

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_24

Keywords Big data · Cloud computing · Internet of Things · Real-time analytics · Systematic mapping

1 Introduction

Cloud is popularly known as a distributed and parallel computing system which comprises a multitude of virtualized and interconnected computers which are provisioned dynamically and shown as at least, one unified computing resource established through negotiations among the service providers and clients which is in the form of a service level agreements (SLAs) [1]. Cloud computing provides storage and compute resources that appear inexhaustible. This is essential to the utilization of Cloud's broad-scale foundation for real-time analytics, IoT (Internet of Things), and big data. Several Cloud service types are known but three main ones standout, namely Software-as-a-Service (SaaS) which offers customized apps to its users across the Internet and the CSP (Cloud Service Provider) which takes the burden of installation and license from the user. Platform-as-a-Service (PaaS) enables the users to deploy and develop various apps on the CSP's infrastructure, while Infrastructure-as-a-Service (IaaS) allows the Cloud users to have control of some of the Cloud service provider's infrastructure and the consumer benefits from the massive compute and storage capacity available to the CSP. Cloud computing is turning out to be very valuable and the services being offered are continuously being improved on a day–to-day premise due to the strong indispensable applications and infrastructure [2, 3]. Cloud computing also has four deployment models and these are private, public, community, and hybrid Cloud. The private Cloud is hosted on-premises in an organization and could either be run by a third party or the organization itself, but in-house staff is utilized. Public Cloud are hosted by major Cloud service providers that possess the resources to build large data centers across various geographical locations. Community Cloud are usually hosted by institutions and organizations having common interests. Hybrid Cloud are used to take advantage of the different Cloud types. Organizations may want to retain core operations on-premises on a private Cloud at the same time, a lesser amount of essential data and services are migrated to the public Cloud. That being said, because of the multitenancy and virtualization procedure in Cloud, a lot of security issues have begun to surface [4, 5].

Big data deals with large volumes of data, IoT involves data transfers between "things", while real-time analytics relates to the processing of data in real time. The common denominator here is usually the processing of a large amount of data and based on the models and service types, Cloud computing has the infrastructure to support the processing and storage of large volumes of information. The big data paradigm often points to a vast and convoluted datasets which conventional processing does not have the capacity to capture, store, and analyze [6]. The essence is to obtain intelligence from such complex big data, which also requires more computing and storage power which could be administered by the Cloud [6].

The growing interest in IoT and big data requires an understanding of their definitions, challenges, and potentials [7]. Big data analytics is needed in the advancement in IoT services and processes due to the ever-increasing amount of data being generated daily [7]. The pace at which current data is being generated is astonishing and at this rate, it has surpassed the available capability to create the appropriate Cloud computing platform for the analysis of data which is a major challenge for researchers [8]. Again, one may be wondering about the connection between all these and the Cloud. This is simply summarized by Amazon Web Services (AWS). Analyzing a huge set of data requires a powerful compute capacity which can change in size depending on the volume of input data and the type of analyses [9]. Despite the fact that Cloud service providers are striving to offer reliable and efficient services on the Cloud, there is a major problem of trust [10]. The data workloads characteristics are impeccably appropriate to the pay-as-you-go model of the Cloud, whereby applications can easily be scaled down or up depending on the demand [9].

In embarking on a research or writing an article, a technical interest area must be considered by the researcher. This includes understanding the topic by reviewing relevant studies relating to that particular topic. It generally involves reviewing multiple journals, conference proceedings, books, etc. In addition, determining an area of interest may also require a lot of search on digital libraries, attending workshops, seminars, and conferences. Researchers can often stumble onto new and often unanticipated research ideas through long hours spent reviewing other people's research and through the process of conducting research as well. Also, researchers may become interested in particular research in a specific observed phenomenon serving as motivation to carry out a large-scale research in all fields and areas of study. To summarize, the researcher's curiosity about an observed phenomenon can usually provide a satisfactory impetus for selecting a research topic. Clearly, the way toward deciding a research topic is sometimes usually complicated. In the area of Cloud large-scale foundation, a large number of research have been undertaken especially as it relates to IoT, big data, and real-time analytics, hence it has become imperative to sum-up and provide an analysis of some of the work that has been done in this field. A systematic mapping study is the process of categorizing research by utilizing a structure and scheme and reporting the result using a map [11].

In this paper, a systematic mapping study presenting the frequency of studies of Cloud large-scale foundation is presented. Such information allows further insight into research carried out so far and areas requiring further attention. The systematic map was obtained by using three facets because of the areas being investigated and these three facets are as follows: the research facet which focuses on the types of research conducted and result within such areas; the contribution facet that deals with the mode of research, in terms of method and tool; and, the topic facets which extracts core topics in Cloud large-scale foundation. The purpose of this paper is to conduct a systematic mapping study of Cloud large-scale foundation for IoT, big data, and real-time analytics. The remaining paper is organized as follows:

Sect. 2 examines related work. Section 3 discusses the systematic mapping process. Section 4 presents the results and discussion. Section 5 is the conclusion with a suggestion for future work.

2 Related Work

In [12], the planning stage of a systematic mapping study was analyzed and it identified the software patterns as evident all through the requirement engineering phase of multiple projects, seeking for a better understanding of the parts being carried out by these patterns which is usually dependent on the fundamental parameters needed in the development procedure. A protocol was created for the study with the fundamental steps to replicate such a work in the research community in order to ascertain the validity of the research. The digital libraries that were utilized in their work were IEEExplore, ACM DL, Web of Science, and Scopus. The guidelines laid down in [11] were followed for this study.

In [13], they dwelled on the protocol's depiction for a systematic mapping study as it relates to DSL (domain-specific languages). Their study is diverted toward a better understanding of the domain-specific language research area with an emphasis on future directions and research trend. This study encompasses between July 2013 to October 2014, and it takes advantage of three rules for carrying out a systematic review which are conducting the review, planning, and reporting such.

In [14], the systematic mapping study is dependent on the examination of the utilization of idea maps within Computer Science and their study conveys the result which was centered on the evaluation and collection of previous research on concept maps within Computer Science. For this study, five different electronic databases were utilized. Manual approaches and backward snowballing were utilized in the search procedure. This study also showed a rich investigation and massive interest in concept maps because of teaching and learning support toward that path or area. The search strings of this study were applied on ScienceDirect, Scopus, ACM DL, Compedex, and IEEExplore.

Within the research work of [15], systematic mapping to inquire how game-related methods have been utilized in software engineering education and how these methods bolster explicit software engineering knowledge domains with the future directions and research gaps recognized. The essential investigations of the work tie down on the assessment of games, their factors, and use on software engineering education. Based on publications from 1974 to 2016, a sum of 156 essential investigations was recognized in this research and the mapping procedure was carried out in tandem with [11].

The research in [16] carried out a power system model mapping centered upon the overview of the power system models and applications being utilized by organizations in Europe, and identification and analysis of both their modeling gaps and features. 228 surveys were sent to power experts for information elicitation,

and 82 questionnaires were completed and the knowledge mapping was done accordingly.

In [17], a domain-specific languages systematic mapping was carried out with an essential enthusiasm for the kind of research, the type of contribution, and the focus area. The work features an inquiry from legitimate sources between 2006 to 2012 with the systematic mapping carried out dependent on highlighting the research questions, data extraction, search, classifying, and conducting the screening. The materials utilized in this research are experience papers, opinion papers, solution proposal, conceptual or philosophical papers, and validation of the research materials.

A systematic mapping was conducted in [18] on legal core ontologies and was based on [19] concepts. The work based its search more on both "legal concepts and theory". The studies selected were classified based on the contribution as reflected in tool, language, model, and method. Other steps include the identification of focus with a clear recommendation to use the two ontologies, identification of the used legal theories in legal core ontologies building process, and the analysis of every chosen research for cogent deductions about ontological and legal research.

In [11], a systematic mapping study in software engineering was conducted and is a foundation for many systematic mapping studies. It provides guidelines for the conduct of systematic mapping studies. In addition, a comparison of systematic reviews and systematic maps was also carried out. The work revealed that systematic maps and reviews are not the same in terms of goals, breadth, validity measures, and implications, because they employ different analysis methods.

The work in [20] provides an observational research in software Cloud-based testing based on systematic mapping. The procedure involves building a categorization scheme making use of both functional and nonfunctional testing methods, which were examined closely alongside the applications of their methods and peculiarities. They made use of 69 main analyses as discovered in 75 publications and just a small amount of the examination unites exact factual investigation with quantitative outcomes. Many of the analyses made use of a singular experiment for the assessment of their proposed solution and there has not been any research focused on systematic mapping of Cloud large-scale foundation for IoT, real-time analytics, and big data from the literature.

3 The Systematic Mapping Process

A systematic study provides an overview of a field of interest by depicting article frequencies in that area. This research was carried out by utilizing the formal steps on systematic mapping research in [11] and [19]. A systematic mapping research is a replicable process of eliciting and translating the available materials associated with a research objective [21]. There are definite procedures for designing a systematic map [11]. Usually, where the scope of the review is developed, the definition of

research questions is the initial step. An inquiry is normally conducted on each and every paper accessible in the proposed area of study and after the search, such papers are examined to discover those important and relevant to the study. Next is the classification scheme based on the use of keywording that is applied to the abstracts of the searched papers. Finally, there is a process of data extraction that prompts the making of the systematic map and each step is utilized in the making of the systematic map of Cloud large-scale foundations for IoT, big data, and real-time analytics. In the perspective of the paper choice criteria in terms of the requirements and research question, a sum of 112 publications were appropriate for inclusion based on the initial 1,716 papers from 2011 to 2018 and the list of selected main research papers are in the Appendix.

3.1 Definition Research Question

The significance of a systematic map is to offer a better understanding of the frequencies and types of research that has been done in an area under consideration. It might likewise be important to find out where such research articles were distributed. These issues determine the focus of the inquiries that are utilized for the research. In this particular paper, the research questions are

- RQ 1: What and how many articles cover the different areas in Cloud large-scale foundation for real-time analytics, IoT, and big data and how are they addressed?
- RQ 2: What evaluation and novelty do they constitute in particular, and what papers are being published in this field/area?

3.2 Conduct of Research for Primary Studies

The strategy of inquiry of primary studies is mostly accomplished by examining unique digital libraries. This is done through searches carried out on conferences and journal publications. To acquire the necessary papers needed for this systematic mapping paper, a search was performed on some scientific digital libraries that are accessible online and information from printed materials and books were not focused on in the search. The search was conducted on journal and conference

Table 1 Digital libraries used for the systematic mapping study

Digital libraries	Uniform resource locator
ACM	http://dl.acm.org/
IEEE	http://ieeexplore.ieee.org/xplore
Science direct	http://www.seciencedirect.com/
Springer	http://www.springerlink.com/

publications in four major digital libraries due to the high-impact factors of their articles. Table 1 shows the URL and the databases used in this research.

The outcome, population, comparison, and intervention in the focus of study influenced the design of the search string. The keywords utilized in the search string were gotten from every aspect structure of the title for this research on large-scale foundations for real-time analytics, big data, and IoT. The search string being utilized on major digital libraries are

(TITLE(Cloud) AND (TITLE("Large scale") OR KEY("Large scale")) ((TITLE ("Big Data") OR KEY(Big Data") OR (TITLE(IOT) OR TITLE ("Internet of things") OR KEY("Internet of things")) OR TITLE("Real time analysis"))) AND (ALIMIT-TO(BUBJAREA, "COMP").

The searches were carried out by making use of the above-customized search string on document metadata in order to make sure that the important articles were not omitted and appropriate databases significant to Computer Science and the Cloud were considered.

3.3 Screening of the Papers for Inclusion and Exclusion

The essence of a selection process is to discover and include each and every publication important to the review being carried out. It is a normal procedure to use exclusion and inclusion criteria to remove articles which are not important to real-time analytics, IoT, and big data in Cloud computing. The exclusion and inclusion criteria are also utilized to eliminate publications which are not important to the research questions. A few abstracts tend to point out the main focus without any further details, hence such abstracts were not considered in the review process. This study excluded papers on panel discussions, tutorials, summaries, and prefaces because they do not contain abstracts. Articles that discussed the main focus of the paper and some measure of secondary aspects were considered. The list of included papers is in the Appendix. The primary focus of this research is on big data, IoT, and real-time analytics, therefore, the exclusion and inclusion procedure was carried out as shown in Table 2.

Table 2 Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
The abstract explicitly mentions real-time analytics, big data, and IoT as it relates to Cloud computing. Furthermore, the abstracts has a secondary focus on large-scale foundation to a certain extent	The paper lies beyond the area of Cloud computing notably as it related to big data, IoT, and real-time analytics

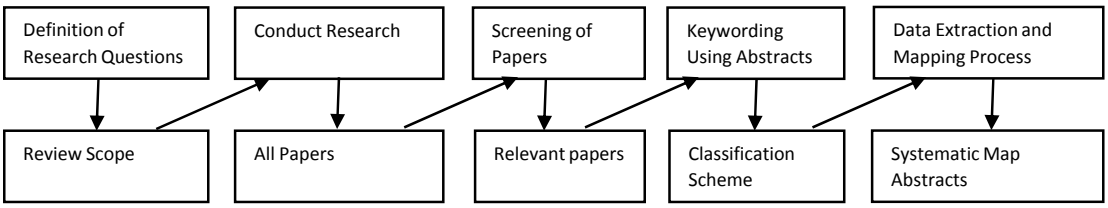


Fig. 1 The systematic mapping process [11]

3.4 Keywording of Abstracts

The keywording of abstracts is the process that leads to the design of a classification scheme. The systematic mapping process shown in Fig. 1 involves the entities listed below [11].

- Abstract
- Keywording
  - Article
  - Sort article into scheme
  - Update scheme
- Systematic map

Keywording is obviously required to lower the time needed to create a systematic map for real-time analytics, big data, and IoT in the Cloud. Also, keywording guarantees that the classification pattern covers every important study and the process includes examining the abstracts in order to extract ideas and keywords from different articles on the focus of study. Keywords from different articles are usually combined to provide a better understanding of the contribution and type of research. Subsequently, this was used in order to better comprehend the set of classifications needed for the study, but it was also crucial in some cases to also examine the introduction and conclusion of some publications so as to ensure reliable keywording. A cluster of keywords was used in order to know the types of facets to be used for creating the systematic map for this study and three main facets were eventually utilized for creating the map. The first facet is the topics category that utilizes the keywords that were directly related to the topic of study. Hence, the topic facet used keywords such as real-time, big data, IoT, orthogonal, and data analytics. The second facet focused on the type of contribution available in a research type. It considered the research contribution in terms of metric, model, method, tool, and process as it relates to the field of study under consideration. The third facet sees to the nature of study carried out.

3.5 Research Type Facet with Category and Description

The essence of the third facet is to determine what types of research were conducted and the applicability to the area of study. The style of the categorization of research approaches by [22] was utilized in this facet and has the following categories and description:

- Evaluation Research: Techniques adopted here have been implemented and evaluated. The outcome of such implementation also considered advantage and disadvantages.
- Validation Research: Processes utilized are novel, though are still to be applied.
- Solution Proposal: Research provides a unique answer to an issue and the advantages plus application of such solutions can be verified.
- Philosophical Papers: The method introduces a new way of examining a challenge with the use of concepts and frameworks.
- Opinion Papers: This kind of study does not adopt any known research methodology; it naturally expresses the researcher's opinion.
- Experience Papers: Relate the researcher's unique experience and explain the way things can be accomplished.

These categories were thought to be adequate and suitable for their utilization in this research. The papers included in this research were reviewed on this basis of the categories of research discussed above.

3.6 Data Extraction and Mapping of Study

During the keywording process, pertinent articles were arranged using a classification design. The following step was to extract data from the relevant publications and this procedure determines the nature of the classification scheme. During this process, new categories could be added, some categories could be eliminated, while some categories could be merged. The data extraction process for this study was done using Microsoft Excel tables which contain frequencies of publication based on different research types and contributions. The frequencies of publication are determined by combining either the contribution or topic tables or the research or topic type tables. The analysis focused on presenting the rates of publication as obtained from the Excel tables which help to identify which areas of the study were emphasized more. Furthermore, it facilitates the identification gaps in the publications that could stimulate further research.

By using the results of the Excel tables, bubble plots were created for the purpose of visual presentation of the frequencies of publications and the systematic map was produced using a two x–y scatter plot with bubbles to indicate the number of articles based on the classes under consideration. There were two quadrants in the systematic map that was created in this study as a result of the three facets

employed. Every quadrant presents a visual map which is usually centered at the junctions of either the research types or topics facet with the contribution category. This makes it uncomplicated to visualize multiple facets at the same time. Furthermore, summary statistics were entered into the bubbles to further enhance understanding of the systematic map and provide a quick summary of the study in the area of big data, IoT. and real-time analytics.

4 Results and Discussion

The primary objective of this research on a large-scale foundation for real-time analytics, big data, and IoT is to have a thematic analysis, carry out classification, and possibly identify the areas where the articles were published. From the analysis, gaps were noticeable from the map, thus identifying which topic areas and research types were lacking in papers. On the contrary, the results also point to areas that are sufficiently covered in terms of publications. Within this research, high-level classifications were utilized in order to access the papers mentioned, hence the reliable readings are obtained and used for developing the map.

4.1 The Contribution Category

The systematic map created from this study is in Fig. 2. On the x-axis of the left quadrant of Fig. 2 is the result of the contribution facet. The contribution facet indicates the type of inputs in the focus of study. The result showed that publications that discussed model in relation to big data, IoT and real-time analytics was 35.85% out of 106 papers examined in this category. Also, tool had 20.75%, metric had 5.66%, method had 25.47%, and process had 12.26%.

4.2 The Research Type Category

On the x-axis of the right quadrant of Fig. 2 is the type of research conducted on the Cloud that is related to big data, IoT and, real-time analytics. The result showed that publications that discussed solution research in relation to big data, IoT, and real-time analytics was 43.75% out of the 112 papers reviewed. Also, evaluation research had 33.93%, validation research had 4.46%, philosophical research had 7.14%, and experience research had 10.71%.

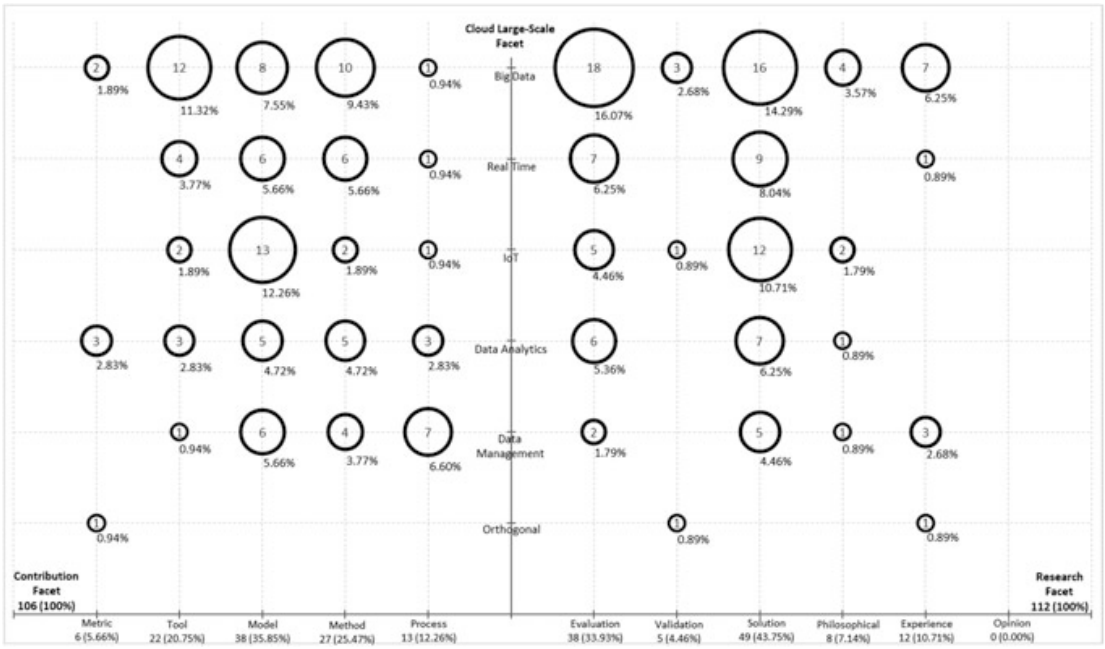


Fig. 2 Systematic map for real-time analytics, big data, and IoT on the Cloud

4.3 Topics and Contribution Facet

The topics used were extracted from the abstracts which related directly to the title of this research work. They are

1. Big data
2. Real-time analytics
3. IoT
4. Data Analytics
5. Data Management
6. Orthogonal

The left quadrant of Fig. 2 displays the relationship between the topics extracted and the contribution facet. Models contributed 38.85% of the papers reviewed, however, a breakdown showed that 7.55% was on big data, 5.66% was on real-time analytics, 12.26% was on IoT, and 4.72% was on data management and no orthogonal contribution. Other aspects of the contribution category as it relates to topics are shown in Fig. 2.

4.4 Topic and Research Facet

The right quadrant of Fig. 2 displays the relationship among the topics and research category and, for example, 33.93% of the papers reviewed in real-time analytics, big data, and IoT in the Cloud were related to evaluation research. The breakdown

indicated 16.07% was on big data, 6.25% was on real-time analytics, 4.4% was on IoT, 5.36% was on data analytics, and 1.79% was on data management. Other aspects of the topics category as it relates to research types are shown in Fig. 2.

4.5 Systematic Map for Big Data, Internet of Things, and Real-Time Analytics

The systematic map for real-time analytics, IoT, and big-data on the Cloud is shown in Fig. 2. The first quadrant is a two x–y scatter plot with bubbles at the intersection of the topic and contribution facet, while the left quadrant is the systematic map depicting a two x–y scatter plot showing the node of the topic and research facet. As discussed earlier, the results made it possible to identify which category of the study had more emphasis in terms of publications. It can be identified from the left quadrant of Fig. 2 that the highest publication was on IoT as it relates to model with 12.26%; there were more publications on data analytics as it relates to metric, (2.83%), more articles on big data in terms of tool, (11.32%), and method (9.43%). Furthermore, there were more research carried out on data management in terms of process with 6.6%. Similarly, on the right quadrant, the highest publication was on big data in terms of evaluation research. More articles also dealt with big data with respect to validation research (2.68%), solution research had 14.29%, philosophical research had 3.57%, and experience research had 6.25%.

On the other hand, to the best of our understanding, there were no any articles on metric in relation to real-time analytics, IoT, and data management on the left quadrant and there were no publications on validation research as it relates to real-time, data analytics and data management on the right quadrant. In addition, there were no publications on experience research in relation to IoT and data analytics. Furthermore, there were no articles on opinion research. Also, the least number of articles were in the study of model and process in the contribution category, while the lowest publications were on validation research, philosophical research, and experience research, in the research category. The visual appeal of the systematic map has been used to summarize the categories of publication making results available to researchers. The essence was to generate interest in the gaps, based on shortages in publications identified and then encourage further research. The value of a systematic map even without a consecutive systematic review cannot be overemphasized. Research gaps could be identified with ease, based on a shortage of publications as shown in frequencies of articles published. This research has created a systematic map that points to the areas lacking in studies in terms of Cloud large-scale foundation. The primary purpose of this study is that analysts at all levels and venture professionals can utilize this as a beginning stage to direct further investigations. This investigation created six classes, namely big data, real-time analytics, IoT, data analytics, data management, and orthogonal in connection to the focus of study. What's more, the six classes of study can be talked

about either regarding the instrument, method, model, process, and metric or in terms of validation, evaluation, philosophical, solution, and opinion research. These fields among others are consequently prescribed for future research. The rundown of included references will likewise help expecting researchers. The significant lessons learnt in this research is that research work is inexhaustible and continuing.

5 Conclusion

Cloud computing is a model that is becoming more relevant by the day. The Cloud enhances the operations of IoT and the processing of big data. Despite the volume of studies, it was identified from the results that so much is happening in terms of big data and IoT compared to real-time analytics. Systematic mapping is a veritable tool in helping to summarize and visually depict the frequencies of publications in an area of research. The mapping process was applied to big data, IoT, and real-time analytics with the attendant results displayed on the systematic map. The results from this research are based on the gaps identified in terms of model, tool, metric, method, and process in relation to real-time analytics, big data, and IoT. This research also identifies some gaps in the area of validation, evaluation, solution, opinion, and philosophical research on real-time analytics, IoT, and big data. To the best of the authors' understanding, there were no articles on metric in relation to real-time analytics, IoT, and data management. In addition, there were no publications on validation research as it relates to data analytics, and data management. Also, there were no publications on experience research in relation to IoT and data analytics. There were no articles on opinion research. Furthermore, the least number of articles were in the study of model and process in the contribution category, while the lowest publications were on validation research, philosophical research, and experience research, in the research category. It is normal that this outcome will fill in as a wide guide into topics that can be looked into on the region of real-time analytics, big data, and IoT. Further research could likewise be completed to approve this investigation or resolve conflicting issues.

Acknowledgements We acknowledge the support and sponsorship provided by Covenant University through the Centre for Research, Innovation and Discovery (CUCRID).

Appendix 1: List of Primary Studies

[1] Adam, O.Y., Lee, Y.C., Zomaya, A.Y. Constructing Performance-Predictable Clusters with Performance-Varying Resources of Cloud (2016) IEEE Transactions on Computers, 65 (9), art. no. 7362012, pp. 2709-2724.

[2] Agrawal, D., Das, S., El Abbadi, A. Big data and Cloud computing: Current state and future opportunities (2011) ACM International Conference Proceeding Series, pp. 530-533.

[3] Agrawal, H., Mathialagan, C.S., Goyal, Y., Chavali, N., Banik, P., Mohapatra, A., Osman, A., Batra, D. Cloudcv: Large-scale distributed computer vision as a Cloud service (2015) Mobile Cloud Visual Media Computing: From Interaction to Service, pp. 265-290.

[4] Al-Ayyoub, M., Jararweh, Y., Tawalbeh, L., Benkhelifa, E., Basalamah, A. Power Optimization of Large Scale Mobile Cloud Computing Systems (2015) Proceedings - 2015 International Conference on Future Internet of Things and Cloud, FiCloud 2015 and 2015 International Conference on Open and Big Data, OBD 2015, art. no. 7300885, pp. 670-674.

[5] Alfazi, A., Sheng, Q.Z., Zhang, W.E., Yao, L., Noor, T.H. Identification as a service: Large-scale Cloud service discovery over the world wide web (2016) Proceedings - 2016 IEEE International Congress on Big Data, BigData Congress 2016, art. no. 7584980, pp. 485-492.

[6] Al-Jaroodi, J., Mohamed, N., Jawhar, I., Mahmoud, S. CoTWare: A Cloud of Things Middleware (2017) Proceedings - IEEE 37th International Conference on Distributed Computing Systems Workshops, ICDCSW 2017, art. no. 7979819, pp. 214-219.

[7] Al-Quraan, M., Al-Ayyoub, M., Jararweh, Y., Tawalbeh, L., Benkhelifa, E. Power optimization of large scale mobile Cloud system using cooperative Cloudlets (2016) Proceedings - 2016 4th International Conference on Future Internet of Things and Cloud Workshops, W-FiCloud 2016, art. no. 7592697, pp. 34-38.

[8] Alsmirat, M.A., Jararweh, Y., Obaidat, I., Gupta, B.B. Internet of surveillance: a Cloud supported large-scale wireless surveillance system (2017) Journal of Supercomputing, 73 (3), pp. 973-992.

[9] Angrisani, L., Ianniello, G., Stellato, A. Cloud based system for measurement data management in large scale electronic production (2014) 2014 Euro Med Telco Conference - From Network Infrastructures to Network Fabric: Revolution at the Edges, EMTC 2014, art. no. 6996651,.

[10] Apolonia, N., Sedar, R., Freitag, F., Navarro, L. Leveraging low-power devices for Cloud services in community networks (2015) Proceedings - 2015 International Conference on Future Internet of Things and Cloud, FiCloud 2015 and 2015 International Conference on Open and Big Data, OBD 2015, art. no. 7300840, pp. 363-370.

[11] Auger, A., Exposito, E., Lochin, E. Sensor observation streams within Cloud-based IoT platforms: Challenges and directions (2017) Proceedings of the 2017 20th Conference on Innovations in Cloud, Internet and Networks, ICIN 2017, art. no. 7899407, pp. 177-184.

[12] Bachiega, J., Reis, M.A.S., De Araujo, A.P.F., Holanda, M. Cost optimization on public Cloud provider for big geospatial data: A case study using open street map (2017) CLOSER 2017 -Proceedings of the 7th

International Conference on Cloud Computing and Services Science, pp. 54-62.

[13] Bojan, V.-C., Raducu, I.-G., Pop, F., Mocanu, M., Cristea, V. Cloud-based service for time series analysis and visualisation in Farm Management System (2015) Proceedings - 2015 IEEE 11th International Conference on Intelligent Computer Communication and Processing, ICCP 2015, art. no. 7312697, pp. 425-432.

[14] Bosse, S. From the Internet-of-Things to Sensor Cloud - Unified Distributed Computing in Heterogeneous Environments with Smart and Mobile Multi-Agent Systems (2015) Smart Systems Integration 2015-9th International Conference and Exhibition on Integration Issues of Miniaturized Systems: MEMS, NEMS, ICs and Electronic Components, SSI 2015, pp. 297-305.

[15] Cao, Y., Sun, D. Migrating large-scale air traffic modeling to the Cloud (2015) Journal of Aerospace Information Systems, 12 (2), pp. 257-266.

[16] Chang, B.-J., Lee, Y.-W., Liang, Y.-H. Reward-based Markov chain analysis adaptive global resource management for inter-Cloud computing (2018) Future Generation Computer Systems, 79, pp. 588-603.

[17] Chang, V. A cybernetics Social Cloud (2017) Journal of Systems and Software, 124, pp. 195-211.

[18] Chen, H., Guo, W. Real-time task scheduling algorithm for Cloud computing based on particle swarm optimization (2015) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9106, pp. 141-152.

[19] Chen, T.-Y., Wei, H.-W., Leu, J.-S., Shih, W.-K. EDZL scheduling for large-scale cyber service on real-time Cloud (2011) Proceedings - 2011 IEEE International Conference on Service-Oriented Computing and Applications, SOCA 2011, art. no. 6166234,.

[20] Clemente-Castelló, F.J., Nicolae, B., Mayo, R., Fernández, J.C., Rafique, M.M. On exploiting data locality for iterative MapReduce applications in hybrid Cloud (2016) Proceedings - 3rd IEEE/ACM International Conference on Big Data Computing, Applications and Technologies, BDCAT 2016, pp. 118-122.

[21] Costan, A., Dobre, C. 1st Workshop on Big Data Management in Cloud - BDMC2012 (2013) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 7640 LNCS, pp. 1-2.

[22] Cuzzocrea, A., Fortino, G., Rana, O. Managing data and processes in Cloud-enabled large-scale sensor networks: State-of-the-art and future research directions (2013) Proceedings - 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing, CCGrid 2013, art. no. 6546142, pp. 583-588.

[23] Da Silva, M.A.A., Sadovykh, A., Bagnato, A., Brosse, E. Taming the complexity of big data multi-Cloud applications with models (2014) CEUR Workshop Proceedings, 1234, pp. 1-12.

[24] Das, A.K., Koppa, P.K., Goswami, S., Platania, R., Park, S.-J. Large-scale parallel genome assembler over Cloud computing environment (2017) Journal of Bioinformatics and Computational Biology, 15 (3), art. no. 1740003,.

[25] Demidov, D. A systematic approach to describing the source code of a Cloud platform with assured security (2017) Proceedings - 2017 5th International Conference on Future Internet of Things and Cloud Workshops, W-FiCloud 2017, 2017-January, pp. 31-36.

[26] Dey, S., Chakraborty, A., Naskar, S., Misra, P. Smart city surveillance: Leveraging benefits of Cloud data stores (2012) Proceedings - Conference on Local Computer Networks, LCN, art. no. 6424076, pp. 868-876.

[27] Dhar, P., Gupta, P. Intelligent parking Cloud services based on IoT using MQTT protocol (2017) International Conference on Automatic Control and Dynamic Optimization Techniques, ICACDOT 2016, art. no. 7877546, pp. 30-34.

[28] Dzik, J., Palladinos, N., Rontogiannis, K., Tsarpalis, E., Vathis, N. MBrace: Cloud computing with monads (2013) Proceedings of the 7th Workshop on Programming Languages and Operating Systems, PLOS 2013 - In Conjunction with the 24th ACM Symposium on Operating Systems Principles, SOSP 2013, .

[29] Feller, E., Ramakrishnan, L., Morin, C. Performance and energy efficiency of big data applications in Cloud environments: A Hadoop case study (2015) Journal of Parallel and Distributed Computing, 79-80, pp. 80-89.

[30] Fernández, A., del Río, S., López, V., Bawakid, A., del Jesus, M.J., Benítez, J.M., Herrera, F. Big Data with Cloud Computing: An insight on the computing environment, MapReduce, and programming frameworks (2014) Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 4 (5), pp. 380-409.

[31] Fortino, G., Russo, W. Towards a Cloud-assisted and agent-oriented architecture for the Internet of Things (2013) CEUR Workshop Proceedings, 1099, pp. 97-103.

[32] Gebremeskel, G.B., Chai, Y., Yang, Z. The paradigm of big data for augmenting internet of vehicle into the intelligent Cloud computing systems (2014) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8662 LNCS, pp. 247-261.

[33] George, J., Chen, C.-A., Stoleru, R., Xie, G.G., Sookoor, T., Bruno, D. Hadoop MapReduce for tactical Cloud (2014) 2014 IEEE 3rd International Conference on Cloud Networking, CloudNet 2014, art. no. 6969015, pp. 320-326.

[34] Gomes, H.M., Carvalho, J.M.D., Veloso, L.R., Teixeira, A.G., Jr., Filho, T.B.D.O., Araujo, A.M.C.D., Zhang, T., Trarbach, L., Machado, F. MapReduce vocabulary tree: An approach for large scale image indexing and search in the Cloud (2016) Proceedings - 2016 IEEE 2nd International

Conference on Multimedia Big Data, BigMM 2016, art. no. 7545016, pp. 170-173.

[35] Hajibaba, M., Gorgin, S. A review on modern distributed computing paradigms: Cloud computing, jungle computing and fog computing (2014) Journal of Computing and Information Technology, 22 (2), pp. 69-84.

[36] Han, H., Lee, Y.C., Choi, S., Yeom, H.Y., Zomaya, A.Y. Cloud-aware processing of MapReduce-based OLAP applications (2013) Conferences in Research and Practice in Information Technology Series, 140, pp. 31-38.

[37] He, S., Cheng, B., Wang, H., Huang, Y., Chen, J. Proactive personalized services through fog-Cloud computing in large-scale IoT-based healthcare application (2017) China Communications, 14 (11), art. no. 8233646, pp. 1-16.

[38] Huang, Q., Li, Z., Xia, J., Jiang, Y., Xu, C., Liu, K., Yu, M., Yang, C. Accelerating geocomputation with Cloud computing (2013) Modern accelerator technologies for geographic information science, 9781461487456, pp. 41-51.

[39] Huang, T.-C., Shieh, C.-K., Huang, S.-W., Chiu, C.-M., Liang, T.-Y. Automatic self-suspended task for a mapreduce system on Cloud computing (2014) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 8643, pp. 257-268.

[40] Huo, Z., Mukherjee, M., Shu, L., Chen, Y., Zhou, Z. Cloud-based Data-intensive Framework towards fault diagnosis in large-scale petrochemical plants (2016) 2016 International Wireless Communications and Mobile Computing Conference, IWCMC 2016, art. no. 7577209, pp. 1080-1085.

[41] Ji, C., Li, Y., Qiu, W., Awada, U., Li, K. Big data processing in Cloud computing environments (2012) Proceedings of the 2012 International Symposium on Pervasive Systems, Algorithms, and Networks, I-SPAN 2012, art. no. 6428800, pp. 17-23.

[42] Jinzhou, Y., Jin, H., Kai, Z., Zhijun, W. Discussion on private Cloud PaaS construction of large scale enterprise (2016) Proceedings of 2016 IEEE International Conference on Cloud Computing and Big Data Analysis, ICCCBDA 2016, art. no. 7529570, pp. 273-278.

[43] Kang, S., Veeravalli, B., Aung, K.M.M., Jin, C. An efficient scheme to ensure data availability for a Cloud service provider (2015) Proceedings - 2014 IEEE International Conference on Big Data, IEEE Big Data 2014, art. no. 7004378, pp. 15-20.

[44] Kanwal, S., Lonie, A., Sinnott, R.O., Anderson, C. Experiences in implementing large-scale biomedical workflows on the Cloud: Challenges in transitioning to the clinical domain (2015) CEUR Workshop Proceedings, 1468, .

[45] Karlsson, J., Torreño, O., Ramet, D., Klambauer, G., Cano, M., Trelles, O. Enabling large-scale bioinformatics data analysis with Cloud computing (2012) Proceedings of the 2012 10th IEEE International Symposium on Parallel and Distributed Processing with Applications, ISPA 2012, art. no. 6280355, pp. 640-645.

[46] Kaur, B. Optimizing VM provisioning of mapreduce tasks on public Cloud (2016) ACM International Conference Proceeding Series, 12-13-August-2016, art. no. a79, .

[47] Kaur, R., Chana, I., Bhattacharya, J. Data deduplication techniques for efficient Cloud storage management: a systematic review (2018) Journal of Supercomputing, 74 (5), pp. 2035-2085.

[48] Kebande, V., Venter, H.S. A functional architecture for Cloud forensic readiness large-scale potential digital evidence analysis (2015) European Conference on Information Warfare and Security, ECCWS, 2015-January, pp. 373-382.

[49] Kitanouma, T., Nii, E., Adachi, N., Takizawa, Y. SmartFinder: Cloud-based self organizing localization for mobile smart devices in large-scale indoor facility (2017) GIoTS 2017 - Global Internet of Things Summit, Proceedings, art. no. 8016245.

[50] Kos, A., Umek, A., Tomaic, S. Comparison of Smartphone Sensors Performance Using Participatory Sensing and Cloud Application (2016) Proceedings - 2015 International Conference on Identification, Information, and Knowledge in the Internet of Things, IIKI 2015, art. no. 7428349, pp. 181-184.

[51] Król, D., Kitowski, J. Towards adaptable data farming in Cloud (2015) Proceedings - 4th IEEE International Conference on Big Data and Cloud Computing, BDCloud 2014 with the 7th IEEE International Conference on Social Computing and Networking, SocialCom 2014 and the 4th International Conference on Sustainable Computing and Communications, SustainCom 2014, art. no. 7034805, pp. 283-284.

[52] Li, P., Guo, S. Load balancing for privacy-preserving access to big data in Cloud (2014) Proceedings - IEEE INFOCOM, art. no. 6849286, pp. 524-528.

[53] Li, T., Wang, K., Zhao, D., Qiao, K., Sadooghi, I., Zhou, X., Raicu, I. A flexible QoS fortified distributed key-value storage system for the Cloud (2015) Proceedings - 2015 IEEE International Conference on Big Data, IEEE Big Data 2015, art. no. 7363794, pp. 515-522.

[54] Lin, J., Yin, J., Cai, Z., Liu, Q., Li, K., Leung, V.C.M. A secure and practical mechanism for outsourcing ELMs in Cloud computing (2013) IEEE Intelligent Systems, 28 (6), pp. 35-38.

[55] Liu, L., Liu, L., Fu, X., Huang, Q., Zhang, X., Zhang, Y. A Cloud-based framework for large-scale traditional Chinese medical record retrieval (2018) Journal of Biomedical Informatics, 77, pp. 21-33.

[56] Liu, X.-F., Zhan, Z.-H., Lin, J.-H., Zhang, J. Parallel differential evolution based on distributed Cloud computing resources for power electronic circuit optimization (2016) GECCO 2016 Companion - Proceedings of the 2016 Genetic and Evolutionary Computation Conference, pp. 117-118.

[57] Liu, Z., Hu, J., Li, Y., Huang, Y. Toward virtual dataspaces for material scientific data Cloud (2016) Concurrency Computation, 28 (6), pp. 1737-1750.

[58] Liu, Z., Hu, C., Li, Y., Hu, J. DSDC: A domain scientific data Cloud based on virtual dataspaces (2012) Proceedings of the 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops, IPDPSW 2012, art. no. 6270579, pp. 2176-2182.

[59] Lolos, K., Konstantinou, I., Kantere, V., Koziris, N. Elastic management of Cloud applications using adaptive reinforcement learning (2018) Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017, 2018-January, pp. 203-212.

[60] Loreti, D., Ciampolini, A. SHYAM: A system for autonomic management of virtual clusters in hybrid Cloud (2016) Communications in Computer and Information Science, 567, pp. 363-373.

[61] Loreti, D., Ciampolini, A. MapReduce over the Hybrid Cloud: A Novel Infrastructure Management Policy (2015) Proceedings - 2015 IEEE/ACM 8th International Conference on Utility and Cloud Computing, UCC 2015, art. no. 7431408, pp. 174-178.

[62] Luo, W., Zhang, H. Visual analysis of large-scale LiDAR point Cloud (2015) Proceedings - 2015 IEEE International Conference on Big Data, IEEE Big Data 2015, art. no. 7364044, pp. 2487-2492.

[63] Luszczek, P., Kurzak, J., Yamazaki, I., Keffer, D., Dongarra, J. Scaling point set registration in 3D across thread counts on multicore and hardware accelerator platforms through autotuning for large scale analysis of scientific point Cloud (2018) Proceedings - 2017 IEEE International Conference on Big Data, Big Data 2017, 2018-January, pp. 2893-2902.

[64] Marwan, M., Kartit, A., Ouahmane, H. Security in Cloud-based medical image processing: Requirements and approaches (2017) ACM International Conference Proceeding Series, Part F129474, art. no. 6.

[65] Mohrehkesh, S., Fedorov, A., Vishwanatha, A.B., Drakopoulos, F., Kikinis, R., Chrisochoides, N. Large Scale Cloud-Based Deformable Registration for Image Guided Therapy (2016) Proceedings - 2016 IEEE 1st International Conference on Connected Health: Applications, Systems and Engineering Technologies, CHASE 2016, art. no. 7545815, pp. 67-72.

[66] Moses, M.B.S., Ambika, K. A hybrid scheme foranonymous authentication of data storage in Cloud (2016) Proceedings of the 2015 International Conference on Green Computing and Internet of Things, ICGCIoT 2015, art. no. 7380489, pp. 360-364.

[67] Mouradian, C., Yangui, S., Glitho, R.H. Robots as-a-service in Cloud computing: Search and rescue in large-scale disasters case study (2018) CCNC 2018 - 2018 15th IEEE Annual Consumer Communications and Networking Conference, 2018-January, pp. 1-7.

[68] Moustafa, N., Creech, G., Sitnikova, E., Keshk, M. Collaborative anomaly detection framework for handling big data of Cloud computing (2017) 2017 Military Communications and Information Systems Conference, MilCIS 2017 - Proceedings, 2017-December, pp. 1-6.

[69] Mseddi, A., Salahuddin, M.A., Zhani, M.F., Elbiaze, H., Glitho, R.H. On optimizing replica migration in distributed Cloud storage systems (2015) 2015 IEEE 4th International Conference on Cloud Networking, CloudNet 2015, art. no. 7335304, pp. 191-197. Nastic, S., Vogler, M., Inzinger, C., Truong, H.-L., Dustdar, S.

[70] RtGovOps: A runtime framework for governance in large-scale software-defined IoT Cloud systems (2015) Proceedings - 2015 3rd IEEE International Conference on Mobile Cloud Computing, Services, and Engineering, MobileCloud 2015, art. no. 7130866, pp. 24-33.

[71] Nunez, D., Agudo, I., Lopez, J. Delegated access for hadoop clusters in the Cloud (2015) Proceedings of the International Conference on Cloud Computing Technology and Science, CloudCom, 2015-February (February), art. no. 7037691, pp. 374-379.

[72] Ochoa, L., González-Rojas, O., Verano, M., Castro, H. Searching for optimal configurations within large-scale models: A Cloud computing domain (2016) Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9975 LNCS, pp. 65-75.

[73] Ogunyadeka, A., Younas, M., Zhu, H., Aldea, A. A Multi-key Transactions Model for NoSQL Cloud Database Systems (2016) Proceedings - 2016 IEEE 2nd International Conference on Big Data Computing Service and Applications, BigDataService 2016, art. no. 7474331, pp. 24-27.

[74] Ohnaga, H., Aida, K., Abdul-Rahman, O. Performance of hadoop application on hybrid Cloud (2016) Proceedings - 2015 International Conference on Cloud Computing Research and Innovation, ICCCRI 2015, art. no. 7421903, pp. 130-138.

[75] Pakdel, R., Herbert, J. A Cloud-based data analysis framework for object recognition (2015) CLOSER 2015-5th International Conference on Cloud Computing and Services Science, Proceedings, pp. 79-86.

[76] Panneerselvam, J., Liu, L., Antonopoulos, N. Characterisation of hidden periodicity in large-scale Cloud datacentre environments (2018) Proceedings - 2017 IEEE International Conference on Internet of Things, IEEE Green Computing and Communications, IEEE Cyber, Physical and Social Computing, IEEE Smart Data, iThings-GreenCom-CPSCom-SmartData 2017, 2018-January, pp. 496-503.

[77] Perera, C., Talagala, D.S., Liu, C.H., Estrella, J.C. Energy-efficient location and activity-aware on-demand mobile distributed sensing platform for sensing as a service in iot Cloud (2015) IEEE Transactions on Computational Social Systems, 2 (4), art. no. 7397993, pp. 171-181.

[78] Petrolo, R., Mitton, N., Soldatos, J., Hauswirth, M., Schiele, G. Integrating wireless sensor networks within a city Cloud (2014) 2014 11th Annual IEEE International Conference on Sensing, Communication, and Networking Workshops, SECON Workshops 2014, art. no. 6979700, pp. 24-27.

[79] Pham, X.-Q., Man, N.D., Tri, N.D.T., Thai, N.Q., Huh, E.-N. A cost- and performance-effective approach for task scheduling based on collaboration between Cloud and fog computing (2017) International Journal of Distributed Sensor Networks, 13 (11).

[80] Prassanna, J., Ajit Jadhav, P., Neelanarayanan, V. Toward an analysis of load balancing algorithms to enhance efficient management of Cloud data centres (2016) Smart Innovation, Systems and Technologies, 49, pp. 143-155.

[81] Rani, B.K., Babu, A.V. Scheduling of Big Data application workflows in Cloud and inter-Cloud environments (2015) Proceedings - 2015 IEEE International Conference on Big Data, IEEE Big Data 2015, art. no. 7364103, pp. 2862-2864.

[82] Rastogi, G., Sushil, R. Analytical literature survey on existing load balancing schemes in Cloud computing (2016) Proceedings of the 2015 International Conference on Green Computing and Internet of Things, ICGCIoT 2015, art. no. 7380705, pp. 1506-1510.

[83] Rea, S., Aslam, M.S., Pesch, D. Serviceware-A service based management approach for WSN Cloud infrastructures (2013) 2013 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2013, art. no. 6529470, pp. 133-138.

[84] Requa, M., Vaughan, G., David, J., Cotton, B. Using Cloud bursting to count trees and shrubs in Sub-Saharan Africa (2016) Proceedings - 2016 IEEE International Conference on Big Data, Big Data 2016, art. no. 7840947, pp. 2960-2963.

[85] Roopaei, M., Rad, P., Jamshidi, M. Deep learning control for complex and large scale Cloud systems (2017) Intelligent Automation and Soft Computing, 23 (3), pp. 389-391.

[86] Rossigneux, F., Lefèvre, L., Gelas, J.-P., De Assunção, M.D. A generic and extensible framework for monitoring energy consumption of openstack Cloud (2015) Proceedings - 4th IEEE International Conference on Big Data and Cloud Computing, BDCloud 2014 with the 7th IEEE International Conference on Social Computing and Networking, SocialCom 2014 and the 4th International Conference on Sustainable Computing and Communications, SustainCom 2014, art. no. 7034862, pp. 696-702.

[87] Shao, Y., Luo, Y., Hu, X., Xue, Y., Xiang, Y., Yin, K. FLAX: A flexible architecture for large scale Cloud fabric (2015) Proceedings - 2015 IEEE International Conference on Smart City, SmartCity 2015, Held Jointly with 8th IEEE International Conference on Social Computing and Networking, SocialCom 2015, 5th IEEE International Conference on Sustainable Computing and Communications, SustainCom 2015, 2015 International Conference on Big Data Intelligence and Computing, DataCom 2015, 5th International Symposium on Cloud and Service Computing, SC2 2015, art. no. 7463881, pp. 1151-1154.

[88] Sheshasaayee, A., Megala, R. A study on resource provisioning approaches in autonomic Cloud computing (2017) Proceedings of the International Conference on IoT in Social, Mobile, Analytics and Cloud, I-SMAC 2017, art. no. 8058325, pp. 141-144.

[89] Spichkova, M., Thomas, I.E., Schmidt, H.W., Yusuf, I.I., Drumm, D.W., Androulakis, S., Opletal, G., Russo, S.P. Scalable and fault-tolerant Cloud computations: Modelling and implementation (2016) Proceedings of the International Conference on Parallel and Distributed Systems - ICPADS, 2016-January, art. no. 7384320, pp. 396-404.

[90] Spichkova, M., Schmidt, H.W., Thomas, I.E., Yusuf, I.I., Androulakis, S., Meyer, G.R. Managing usability and reliability aspects in Cloud computing (2016) ENASE 2016 - Proceedings of the 11th International Conference on Evaluation of Novel Software Approaches to Software Engineering, pp. 288-295.

[91] Sundharakumar, K.B., Dhivya, S., Mohanavalli, S., Vinob Chander, R. Cloud based fuzzy healthcare system (2015) Procedia Computer Science, 50, pp. 143-148.

[92] Taherkordi, A., Eliassen, F. Poster abstract: Data-centric IoT services provisioning in Fog-Cloud computing systems (2017) Proceedings - 2017 IEEE/ACM 2nd International Conference on Internet-of-Things Design and Implementation, IoTDI 2017 (part of CPS Week), pp. 317-318.

[93] Taherkordi, A., Eliassen, F. Scalable modeling of Cloud-based IoT services for smart cities (2016) 2016 IEEE International Conference on Pervasive Computing and Communication Workshops, PerCom Workshops 2016, art. no. 7457098, .

[94] Talia, D., Trunfio, P., Marozzo, F. Data Analysis in the Cloud: Models, Techniques and Applications (2015) Data Analysis in the Cloud: Models, Techniques and Applications, pp. 1-138.

[95] Tang, H., Li, Y., Jia, T., Wu, Z. Evaluating performance of rescheduling strategies in Cloud system (2016) Proceedings - 15th IEEE International Conference on Trust, Security and Privacy in Computing and Communications, 10th IEEE International Conference on Big Data Science and Engineering and 14th IEEE International Symposium on Parallel and Distributed Processing with Applications, IEEE TrustCom/BigDataSE/ISPA 2016, art. no. 7847125, pp. 1559-1566.

[96] Tian, W., Xue, R., Dong, X., Wang, H. An approach to design and implement RFID middleware system over Cloud computing (2013) International Journal of Distributed Sensor Networks, 2013, art. no. 980962, .

[97] Tunc, C., Hariri, S., Montero, F.D.L.P., Fargo, F., Satam, P., Al-Nashif, Y. Teaching and Training Cybersecurity as a Cloud Service (2015) Proceedings - 2015 International Conference on Cloud and Autonomic Computing, ICCAC 2015, art. no. 7312173, pp. 302-308.

[98] Vaquero, L.M., Celorio, A., Cuadrado, F., Cuevas, R. Deploying large-scale datasets on-demand in the Cloud: Treats and tricks on data distribution (2015) IEEE Transactions on Cloud Computing, 3 (2), art. no. 6910293, pp. 132-144.

[99] Vilaplana, J., Solsona, F., Teixidó, I., Mateo, J., Usié, A., Torres, N., Comas, J., Alves, R. MetReS: A metabolic reconstruction database for Cloud computing (2014) Proceedings - 2014 International Conference on Intelligent Networking and Collaborative Systems, IEEE INCoS 2014, art. no. 7057165, pp. 653-658.

[100] Wang, G.-L., Han, Y.-B., Zhang, Z.-M., Zhu, M.-L. Cloud-based integration and service of streaming data (2017) Jisuanji Xuebao/Chinese Journal of Computers, 40 (1), pp. 107-125.

[101] Wang, T., Yao, S., Xu, Z., Jia, S., Xu, Q. A data placement strategy for big data based on DCC in Cloud computing systems (2015) Proceedings - 2015 IEEE International Conference on Smart City, SmartCity 2015, Held Jointly with 8th IEEE International Conference on Social Computing and Networking, SocialCom 2015, 5th IEEE International Conference on Sustainable Computing and Communications, SustainCom 2015, 2015 International Conference on Big Data Intelligence and Computing, DataCom 2015, 5th International Symposium on Cloud and Service Computing, SC2 2015, art. no. 7463793, pp. 623-630.

[102] Woodworth, J., Salehi, M.A., Raghavan, V. S3C: An architecture for space-efficient semantic search over encrypted data in the Cloud (2016) Proceedings - 2016 IEEE International Conference on Big Data, Big Data 2016, art. no. 7841040, pp. 3722-3731.

[103] Xu, Y., Helal, A. Scalable Cloud-Sensor Architecture for the Internet of Things (2016) IEEE Internet of Things Journal, 3 (3), art. no. 7155463, pp. 285-298.

[104] Yang, C., Liu, C., Zhang, X., Nepal, S., Chen, J. A time efficient approach for detecting errors in big sensor data on Cloud (2015) IEEE Transactions on Parallel and Distributed Systems, 26 (2), art. no. 6714550, pp. 329-339.

[105] Yingchi, M., Ziyang, X., Ping, P., Longbao, W. Delay-Aware Associate Tasks Scheduling in the Cloud Computing (2015) Proceedings - 2015 IEEE 5th International Conference on Big Data and Cloud Computing, BDCloud 2015, art. no. 7310724, pp. 104-109.

[106] Yu, T., Wang, X., Jin, J., McIsaac, K. Cloud-orchestrated physical topology discovery of large-scale IoT systems using UAVs (2018) IEEE Transactions on Industrial Informatics, 14 (5), pp. 2261-2270.
[107] Yuan, D., Jin, J., Grundy, J., Yang, Y. A framework for convergence of Cloud services and Internet of things (2015) Proceedings of the 2015 IEEE 19th International Conference on Computer Supported Cooperative Work in Design, CSCWD 2015, art. no. 7230984, pp. 349-354.
[108] Zhang, J., Li, T., Pan, Y. PLAR: Parallel large-scale attribute reduction on Cloud systems (2014) Parallel and Distributed Computing, Applications and Technologies, PDCAT Proceedings, art. no. 6904253, pp. 184-191.
[109] Zhang, Y.F., Zhang, G., Yang, T., Wang, J.Q., Sun, S.D. Service encapsulation and virtualization access method for Cloud manufacturing machine (2014) Jisuanji Jicheng Zhizao Xitong/Computer Integrated Manufacturing Systems, CIMS, 20 (8), pp. 2029-2037.
[110] Zhou, X., Tang, N., Kuang, Y. A universal framework for flexible Cloud computing (2016) Proceedings of 2016 IEEE International Conference on Big Data Analysis, ICBDA 2016, art. no. 7509841, .
[111] Zhuang, Z., Guo, C. OCPA: An algorithm for fast and effective virtual machine placement and assignment in large scale Cloud environments (2013) Proceedings - 2013 International Conference on Cloud Computing and Big Data, CLOUDCOM-ASIA 2013, art. no. 6821001, pp. 254-259.
[112] Zimmermann, A., Pretz, M., Zimmermann, G., Firesmith, D.G., Petrov, I., El-Sheikh, E. Towards service-oriented enterprise architectures for big data applications in the Cloud (2013) Proceedings - IEEE International Enterprise Distributed Object Computing Workshop, EDOC, art. no. 6690543, pp. 130-135.

References

1. Buyya, R., Broberg, J., Goscinski, A.: Cloud Computing Principles and Paradigms, pp. 4–10. Wiley, New York (2011)
2. Odun-Ayo, I., Ananya, M., Agono, F., Goddy-Worlu, R.: Cloud computing architecture: a critical analysis. In: IEEE Proceedings of the 2018 18th International Conference on Computational Science and Its Applications (ICCSA 2018), pp. 1–7 (2018). https://doi.org/10.1109/iccsa.2018.8439638
3. Odun-Ayo, I., Odede, B., Ahuja, R.: Cloud applications management-issues and developments. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (LNCS), vol. 10963, pp. 683–694. Springer, Berlin, Germany (2018)
4. Odun-Ayo, I., Misra, S., Abayomi-Alli, O., Ajayi, O.: Cloud multi-tenancy: issues and developments. In: UCC '17 Companion. Companion Proceedings of the10th International Conference on Utility and Cloud Computing, pp. 209–214 (2017)
5. Odun-Ayo, I., Misra, S., Omoregbe, N., Onibere, E., Bulama, Y., Damasevičius, R.: Cloud-based security driven human resource management system. Front. Artif. Intell. Appl. **295**, 96–106 (2017). https://doi.org/10.3233/978-1-61499-773-3-96

6. Stephanopoulos, G., Mavromoustakis, C.X., Mastorakis, G.S., Paragiostakis, S., Pallis, E., Batalia, J.M.: Big Data and cloud computing: a survey of the state of the art and research challenges. In: Mavromoustakis, C., Mastorakis, G., Dobre, C. (eds.) Advances in Mobile Computing and Big Data in SG Era, Studies in Big Data, vol. 22. Springer, Berlin (2017)
7. Mohammadi, M., Al-fuqaha, A., Sorour, S., Guizani, M. (2017) Deep learning for IoT big data and streaming analytics: a survey. arXiv:1712.0430IvI [cs.NI]
8. Hashem, A., Yaqoob, I., Anuar, N.M., Mokhtar, S., Gani, A., Khan, S.U.: The rise of big data on cloud computing: review and open research issues. Inf. Syst. **47**, 98–115 (2014)
9. Amazon Web Services: Big data analytics option on AWS (2016)
10. Odun-Ayo, I., Omoregbe, N., Odusami, M., Ajayi, O.: Cloud ownership and reliability - issues and developments. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (LNCS), vol. 10658, pp. 231–240. Springer, Berlin, Germany
11. Petersen, K., Feldt, R., Mujtaba, S., Mattsson, M.: Systematic Mapping Studies in Software Engineering. In: EASE'08 Proceedings of the 12th international conference on Evaluation and Assessment in Software Engineering, Italy, pp. 68–77, 26–27 June 2008
12. Barros-Justo, J.L., Cravero-Leal, A.L., Benitti, F.B. Capilla-Sevilla, R.: Systematic mapping protocol: the impact of using software patterns during requirements engineering activities in real-world settings. Cornell University Library (2017). arXiv:1701.05747v1 [cs.SE]
13. Kosar, T., Bohra, S., Mernik, M.A.: Protocol of a systematic mapping study for domain-specific languages. J. Inf. Softw. Technol. **21**(3), 77–91 (2016)
14. Santos, V., Souza, E.F., Felizardo, K.R., Vijaykumar, N.L.: Analyzing the use of concept maps in computer science: a systematic mapping study. Inform. Educ. **16**(2), 257–288 (2017). https://doi.org/10.15388/infedu.2017.13
15. Souza, M., Veado, L., Moreira, R.T., Figueiredo, E., Costa, H.: A systematic mapping study on game-related methods for software engineering education. Inf. Softw. Technol. **95**, 201–218 (2018)
16. Fernandez-Blanco, C.R., Careri, F., Kavvadias, K., Hidalgo Gonzalez, I., Zucker, A., Peteves, E.: Systematic Mapping of Power System Models: Expert Survey, EUR 28875 EN. Publications Office of the European Union, Luxembourg (2017). https://doi.org/10.2760/422399. ISBN 978-92-79-76462-2
17. Mernik, M.: Domain-specific languages: a systematic mapping study. In: International Conference on Current Trends in Theory and Practice of Informatics, Lecture Notes in Computer Science, vol 10139, pp. 464–472. Springer, Berlin, Germany (2017)
18. Griffo, C., Almeida, J.P.A., Guizzardi, G.: A systematic mapping of the literature on legal core ontologies. In: Brazilian Conference on Ontologies, ONTOBRAS 15, CEUR Workshop Proceedings, 1442 (2015)
19. Kitchenham, B., Charters, S.: Guidelines for performing systematic literature review in software engineering. Version 2. 2007-01 (2007)
20. Ahmad, A., Brereton, P., Andras, P.: A systematic mapping study of empirical studies on software cloud testing methods. In: IEEE International Conference on Software Quality, Reliability and Security Companion, pp. 555–562 (2017)
21. Muhammed, A.C., Muhammed, A.B.: A Systematic Mapping Study Of Software Architectures For Cloud Based Systems, IT University Technical Report Series, IT University of Copenhagen (2014)
22. Wieringa, R., Maiden, N.A., Mead, N.R., Rolland, C.: Requirement engineering paper classification and evaluation criteria. A proposal and a discussion. Requir. Eng. **11**(1), 102–107 (2006)

Studies on Radar Imageries of Thundercloud by Image Processing Technique

Sonia Bhattacharya and Himadri Bhattacharyya Chakrabarty

Abstract Severe atmospheric event can cause huge damage to civilization. Severe thunderstorm is one of those weather events. Analysis of cloud imageries can be used to forecast severe thunderstorm. Convective clouds are one of the main reasons for the formation of severe thunderstorm. Analysis of such cloud imageries by image processing can be used to predict severe thunderstorm. Analysis of RGB values of pixel of cloud imageries can be used to show the formation of severe thunderstorm. Histogram analysis of such cloud imageries can also be used to predict severe thunderstorm. In this study analysis of RGB values of pixels and histograms of cloud imageries has been used to now cast severe thunderstorm with a lead time of 6 to 8 h. This lead time is necessary to save life and property from huge damages.

Keywords Convective cloud · Histogram · Image processing · Rader imageries · RGB values · Severe thunderstorm

This is here by certified that this paper has not been submitted, accepted or published anywhere and the work done in the manuscript is original.

S. Bhattacharya (✉)
CWTT, Department of Computer Science, Panihati Mahavidyalaya,
Sodepur, Kolkata 700110, India
e-mail: bhattacharyasonia@rocketmail.com

H. B. Chakrabarty
Principal, JRM, University of Calcutta, Kolkata, India
e-mail: himi0201@gmail.com

Head of the Department & PG Coordinator, Department of Computer Science,
Surendranath College (On lien), University of Calcutta, Kolkata, India

UGC sponsored Visiting Professor, Radio Physics and Electronics,
University of Calcutta, Kolkata, India

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_25

1 Introduction

Convective clouds are formed vertically upwards as the result of the instability of atmosphere. In India, the coastal areas (like the west coast, Orissa, Andhra Pradesh, and West Bengal) as well as north east have been affected many times by thunderstorms and heavy rainfall. Natural calamities are unpreventable and incur huge losses both in terms of life and property. Analysis of convective cloud imageries can give correct prediction of such severe weather event. During the last decade, new classes of cloud resolution mesoscale models have been hurriedly developed. They are able to simulate the life cycle of each convective clouds [1]. Some papers [2–4] concerned the propagation and regeneration of cells within a multi cell storm. The three months of March–April–May (MAM) are pre-monsoon season in India. Severe thunderstorms occur mostly in this time period. Four months of June–July–August–September are identified as monsoon time in India. Cyclonic weather features are evolved generally at the post-monsoon season due to formation of convective cloud during October-November months. The observations showed that the mesoscale phenomenon is the beginning of a severe convection before the monsoon and its effect on the climate, [15]. It is often apparent on satellite photos that convective conditions are occurring simultaneously over extensive areas and are the result of the vertical structure of an air mass and the characteristics of the actual dominant large-scale weather systems, [5]. A large number of research works have been done to analyze clouds based on satellite images [5]. Since the number of droplets of cloudy liquid and ice crystals increases in the cloud, reflectivity value increases, this has a direct relation with the rain [6]. Color is an important feature for image representation [5]. Color moments, color histogram and coherent color vector are the important components of the extraction of color characteristics [7]. Any satellite picture of convective cloud fields has a great variety of cloud patterns and cloud sizes, [8]. The size distribution of convective clouds might give hints of the precipitation observed at ground stations [8]. The main aim of this research work is to study color density of the pixel which changes with formation of cloud. Red color in cloud imagery represents water content of the cloud. This study will reveal that color density of pixel increases from image to image as the cloud devolved with respect to time. In this work the imageries of a thunder cloud have been studied stage by stage such as developing stage/Cumulus stage, mature stage, and Dissipating stage.

2 Data

Here in this study ten images of a thundercloud having different stages from time to time are considered for analysis. All these images are taken over Kolkata (22.3°N/88.3°E) on April 10, 2005 from 07:12:04 to 14:12:04. The cloud imageries were obtained from the observations through Doppler Radar of Regional Meteorological Center, India Meteorological Dept., Kolkata.

3 Methodology

Image Processing

Here in this study ten GIF images have been considered for weather forecast. All these imageries are TRUE color image. Here each pixel has a particular color; that color being described by the amount of red, green and blue in it. Since the total number of bits needed for each pixel is 24, these images are also called bit color images. An image of this type can be considered as a stack of three matrices; representing the values of red, green, and blue for each pixel. This means that for each pixel there are three values [14]. Here in this study impixel (I) function has been used which returns the value of pixels in image I. The pixel values of two consecutive images have been obtained using impixel function which gives two three columnar matrices. A comparison has been done between these two matrices of two consecutive imageries which revealed that there is a noticeable increase between their values. There are a total of nine comparisons between these ten images.

4 Observations on Cloud Imageries

As the first three images showed that formation of convective cloud, in which mainly three types of colors are observed, these are Blue, Red and Yellow. Red color represents the water content of the cloud. In the first image there is no cloud activity is observed, in the second image cloud has been formed over Bankura and Purulia and became denser in the third image moving towards Kolkata (Figs. 1, 2, 3, 4, 5 and 6).

In the fourth, fifth and sixth images it can be observed that density of red color pixels has been increased noticeably and the cloud has been moved more towards Kolkata. It can be observed that from seventh image the density of red colored pixels were decreasing as it moves over Kolkata. Since red colored pixel denotes water content of the cloud, reduction of red color pixel density indicates towards the precipitation (Figs. 7 and 8).

It can be clearly observed that in ninth and tenth image red colored pixel were noticeably reduced leaving yellow colored pixels, this signifies that precipitation has been done (Figs. 9 and 10).

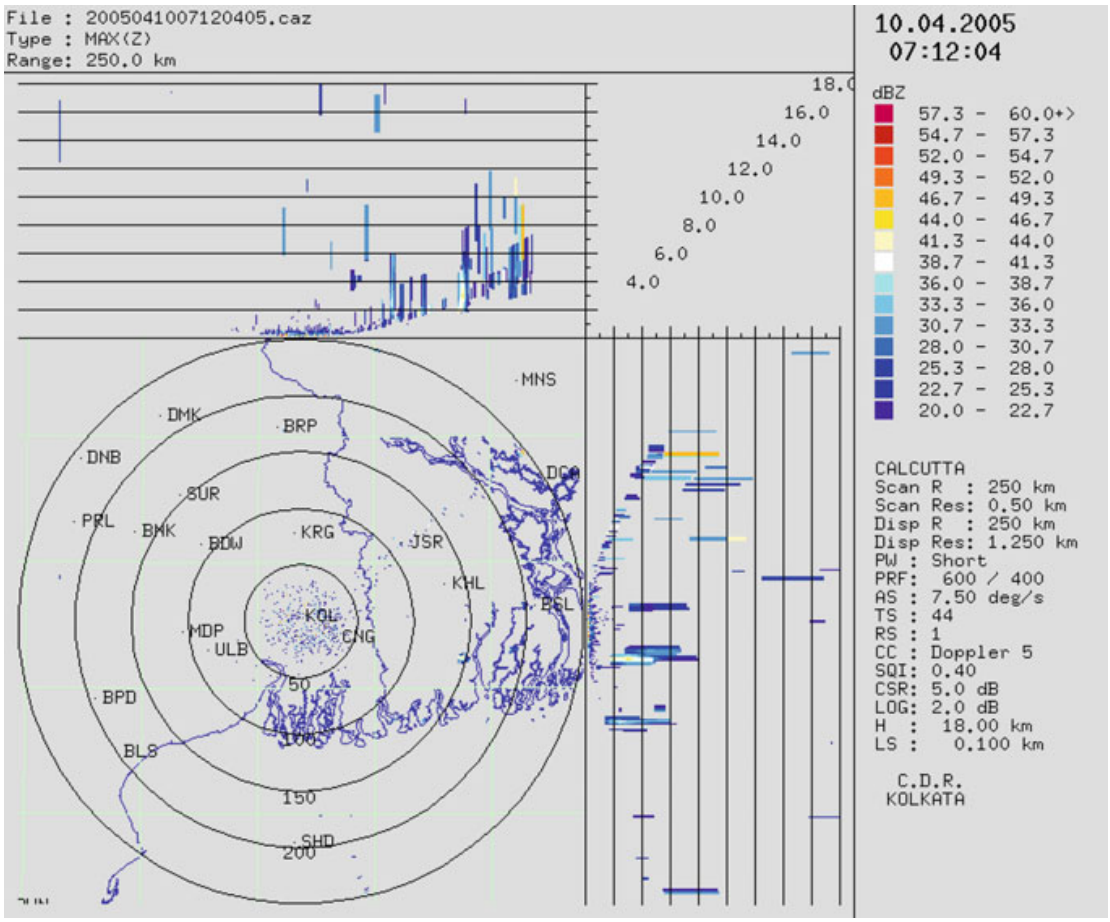


Fig. 1 First image of cloud formation

5 Result

Using impixel function RGB value of each pixel has been obtained and a comparison between the RGB values of the pixels of two consecutive images yield the following result:

Table 1 show that the density of RGB values increases from image to image and has maximized between sixth and seventh images. We have already observed that red color pixel density were most in seventh image. These values are reducing for the next remaining images, though values increase slightly in the last image. Comparing the last two images, it has been found that a few more red colored pixels are present in the tenth image.

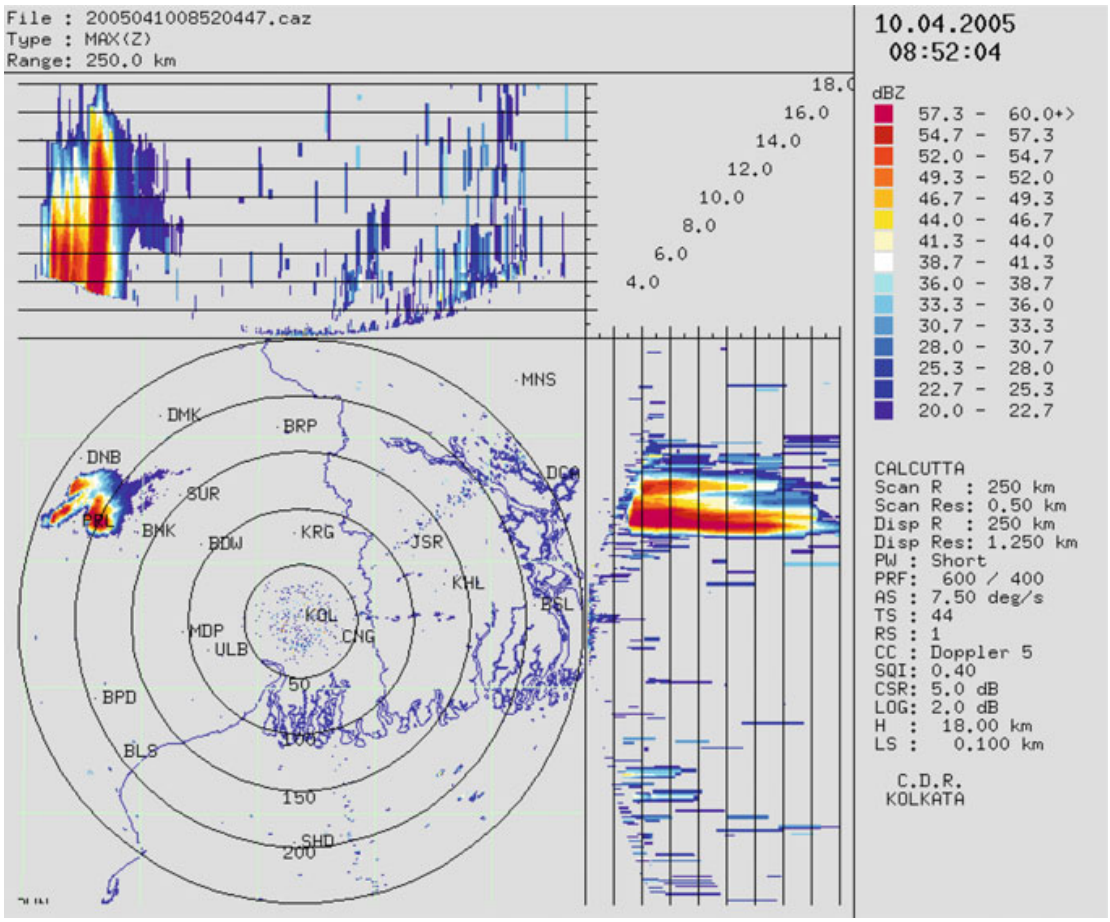


Fig. 2 Second image of cloud formation

6 Analysis of Histograms

It can be further observed that the histograms of these images show that the intensity has increased with formation of the cloud. The imhist(I) function displays plot of the histogram. The histogram shows the distribution of pixel values. The study reveals that frequency of pixels has been increased from one image to next consecutive image. It has been reached most in the seventh image. The seventh image shows maximum presence of red colored pixels indicating towards the intensification of cloud. Precipitation is indicated by increase of yellow colored pixels in image the effect of which can be seen from eighth imagery. The distribution of pixel values has been increased from image to image as cloud has developed with respect to time (Figs. 11, 12, 13, 14, 15, 16, 17, 18, 19 and 20).

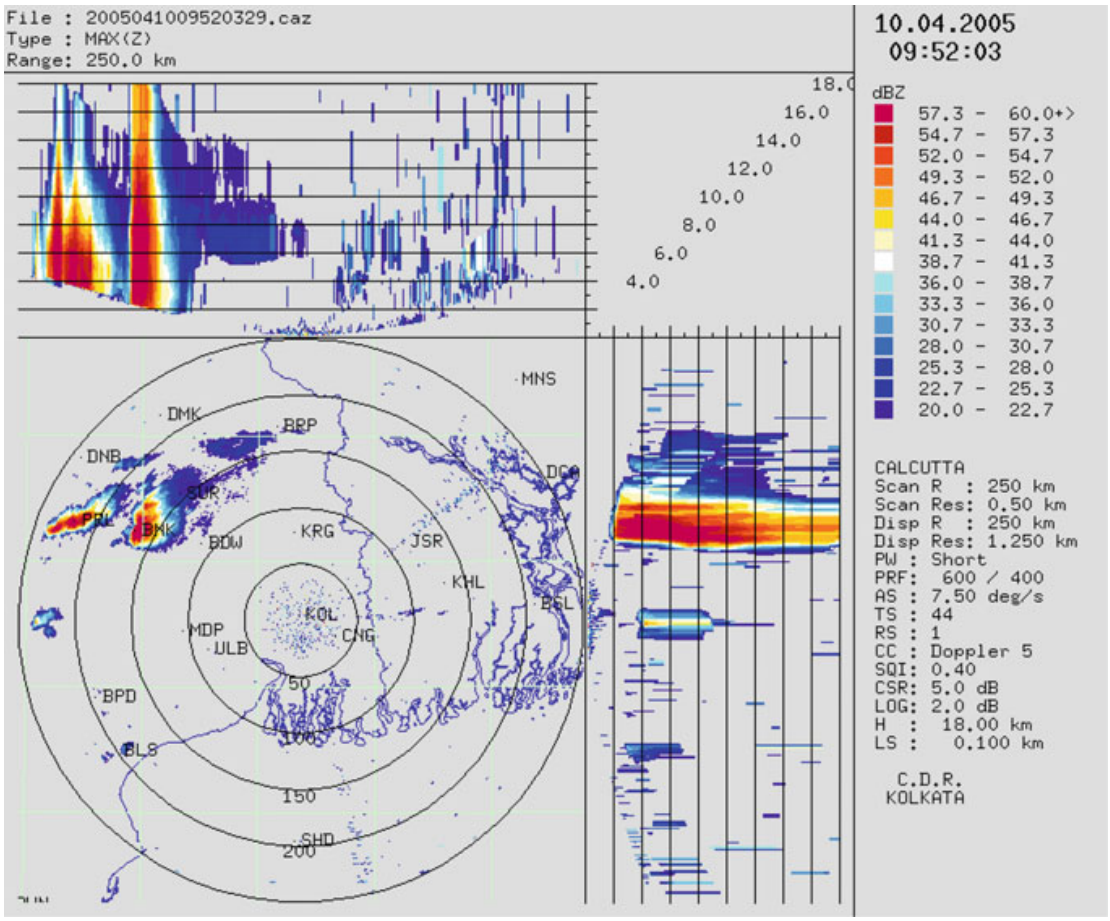


Fig. 3 Third image of cloud formation

7 Discussion

The pixel comparison of clear sky images is used for the determination of colors [9]. It is very important to monitor the analysis of clouds images to forecast when the thunderstorm along with the precipitation occurs. It can also be analyzed how much precipitation can occur by studying cloud images and its duration of rainfall. One can guess the severity of the natural disaster tentatively 6 h before by this type of cloud analysis. People can become alert with a lead time of 6 to 8 h. There are several meteorological satellites that provides large-scale observation of clouds through the day [10, 11, 12]. The patterns of image pixels are used for the classification of images [13].

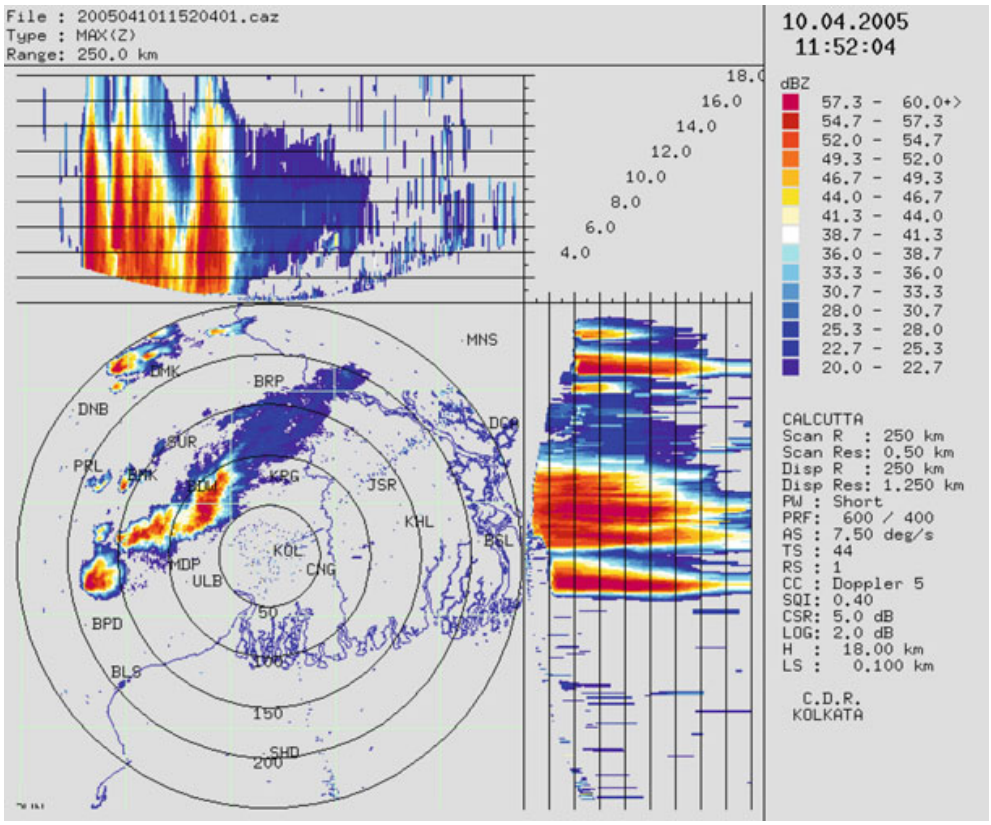


Fig. 4 Fourth image of cloud formation

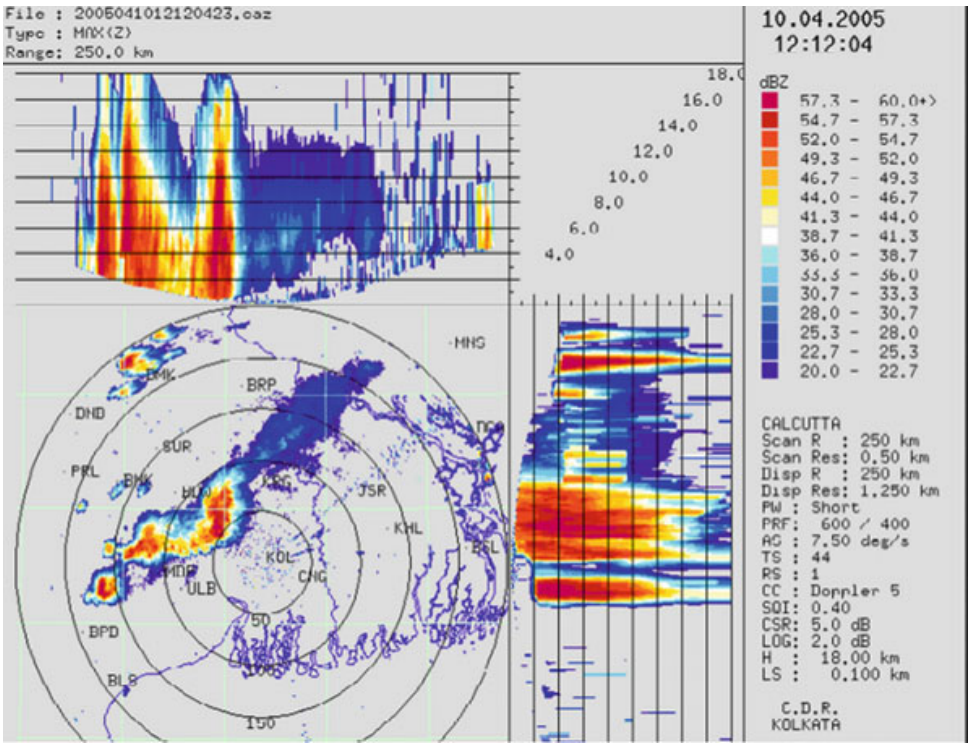


Fig. 5 Fifth image of cloud formation

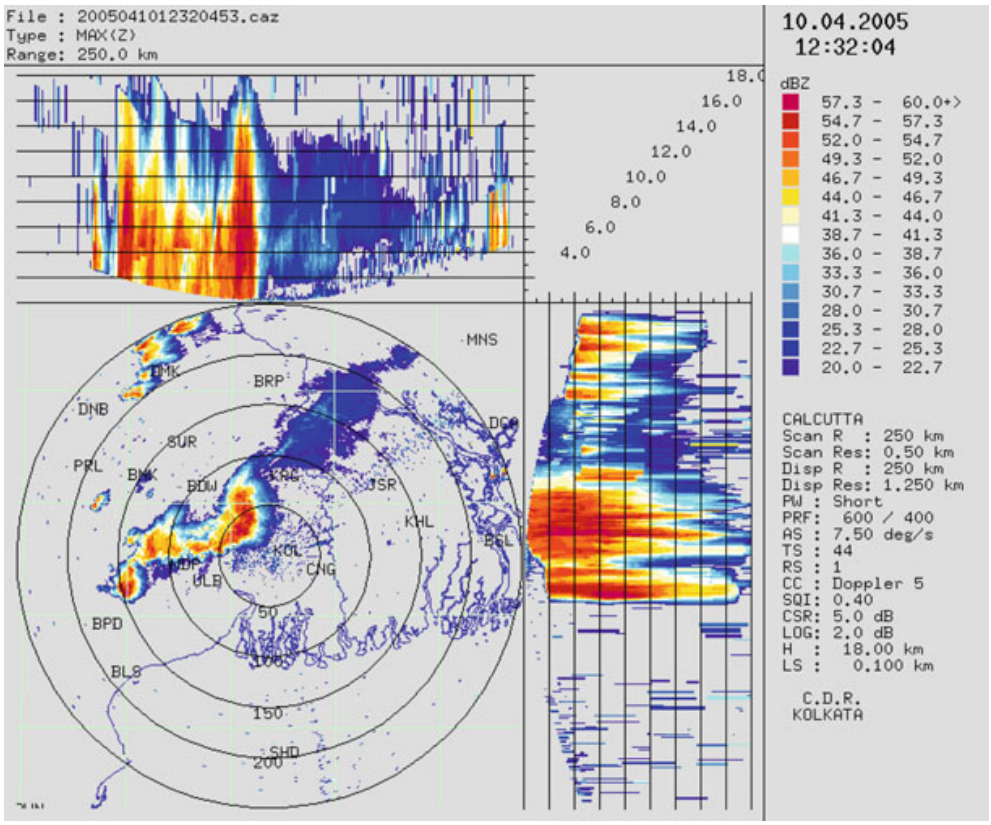


Fig. 6 Sixth image of cloud formation

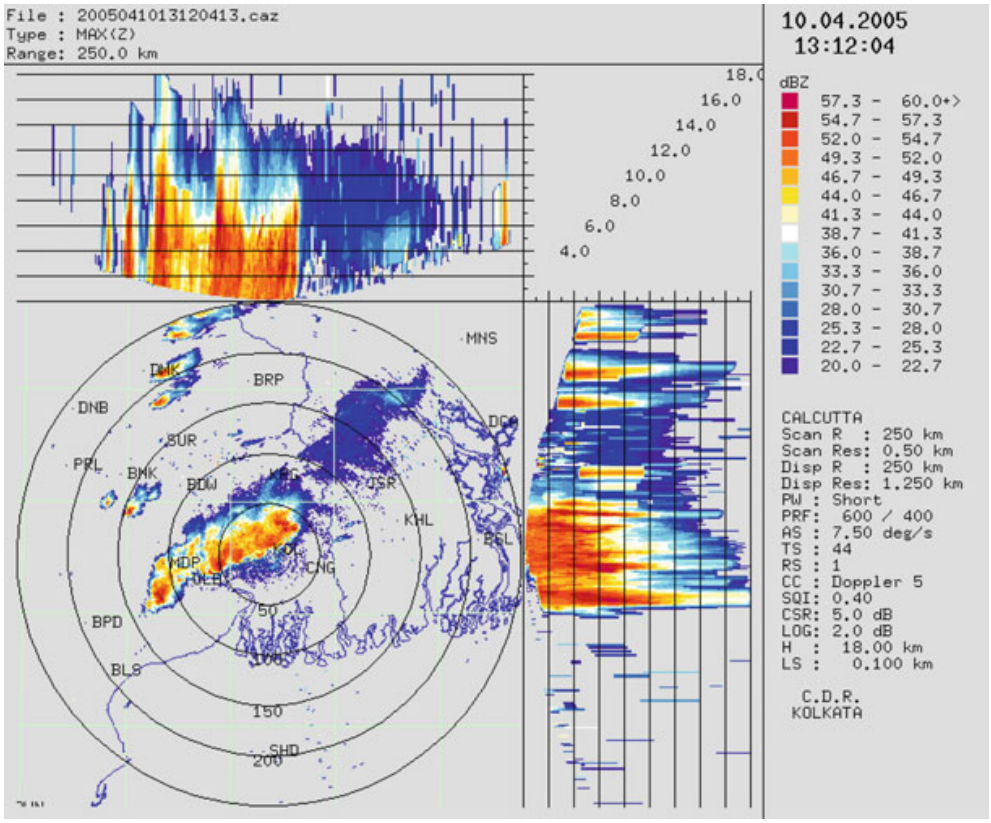


Fig. 7 Seventh image of cloud formation

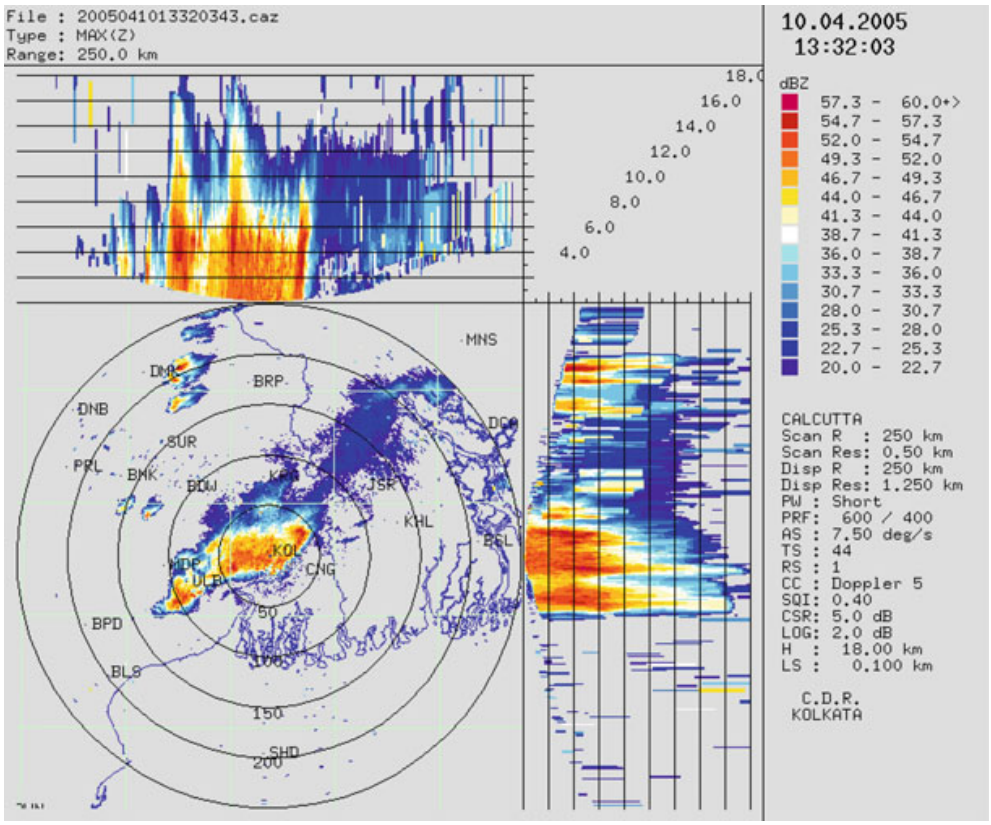


Fig. 8 Eighth image of cloud formation

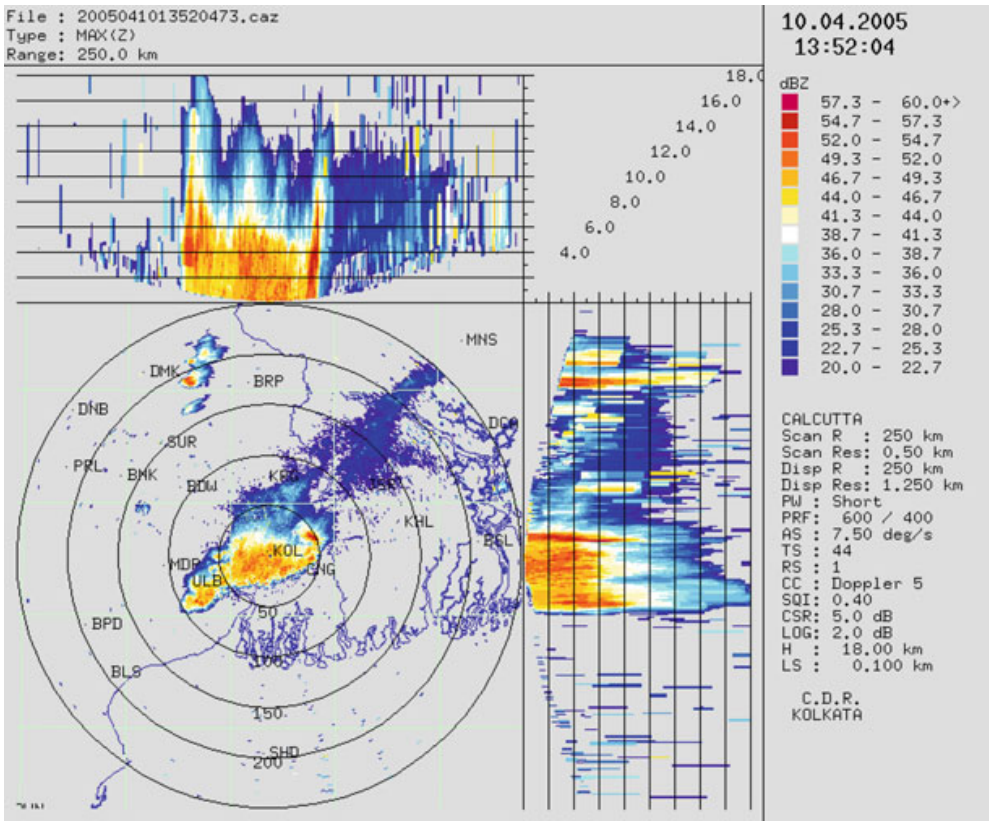


Fig. 9 Ninth image of cloud formation

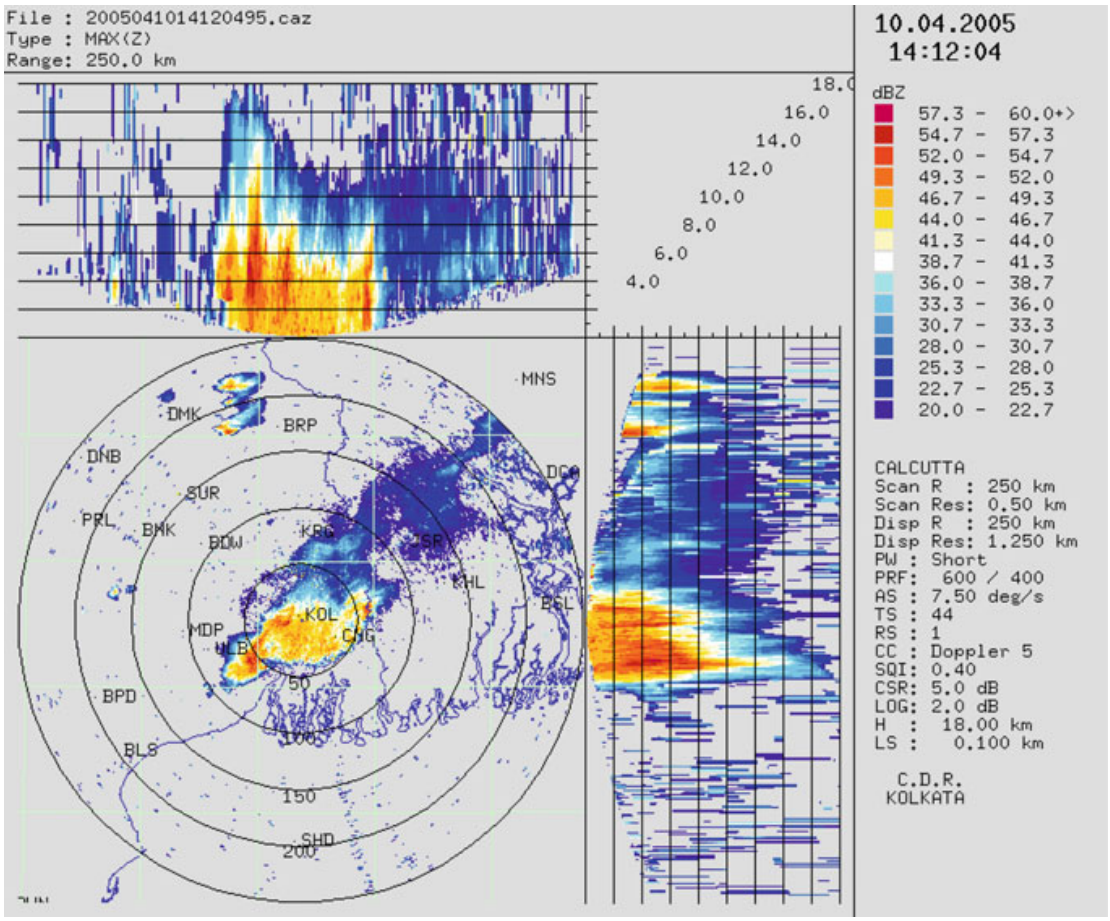


Fig. 10 Tenth image of cloud formation

Table 1 Compared value between two consecutive images

Increase in pixel value	Consecutive images
218	First-second image
240	Second-third image
829	Third-fourth image
1336	Fourth-fifth image
2126	Fifth-sixth image
3472	Sixth-seventh image
2539	Seventh-eighth image
1837	Eighth-ninth image
1845	Ninth-tenth image

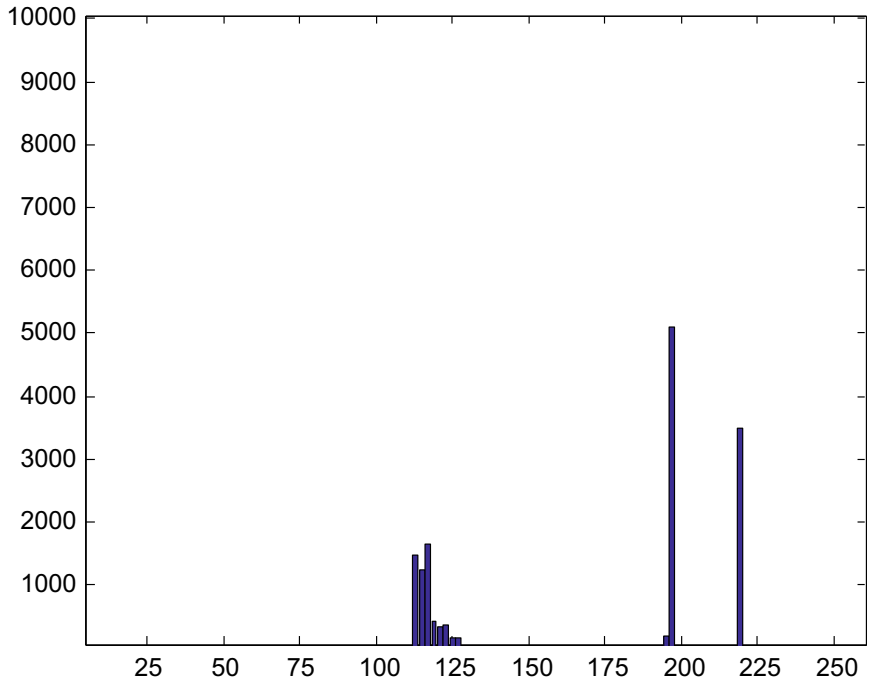


Fig. 11 Histogram of first image

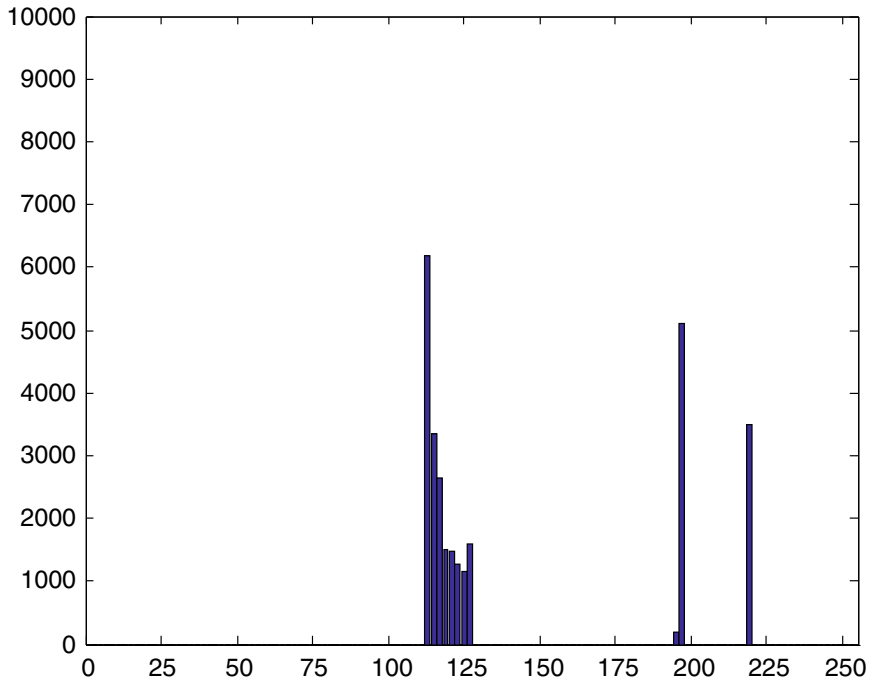


Fig. 12 Histogram of second image

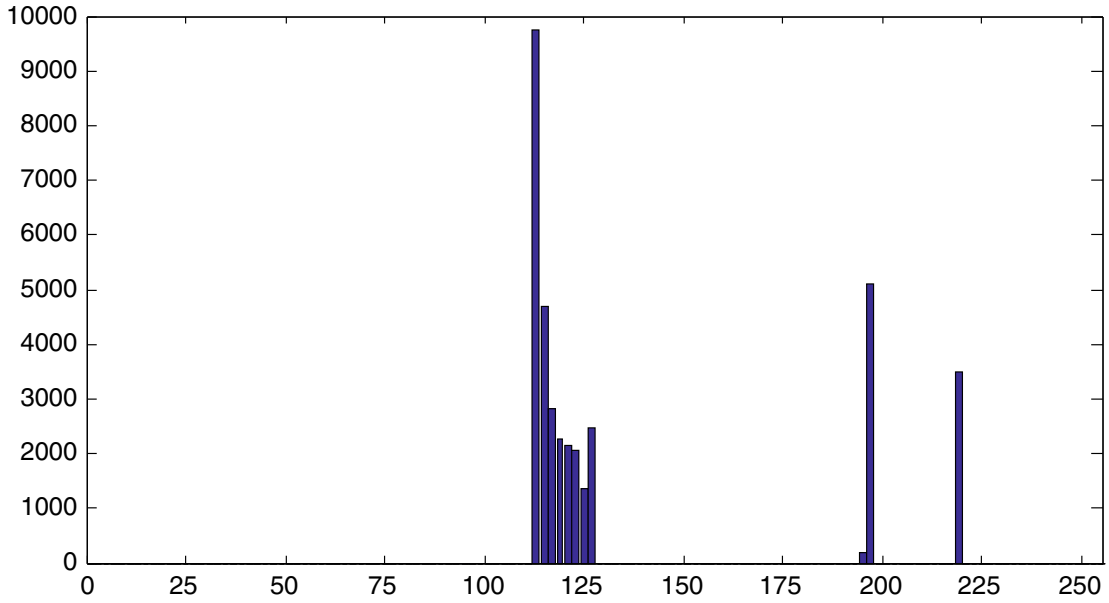


Fig. 13 Histogram of third image

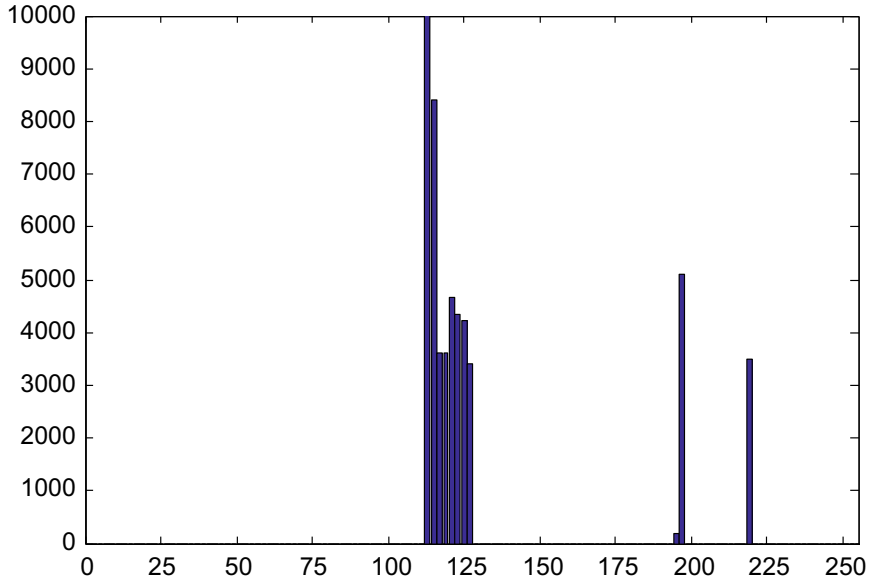


Fig. 14 Histogram of fourth image

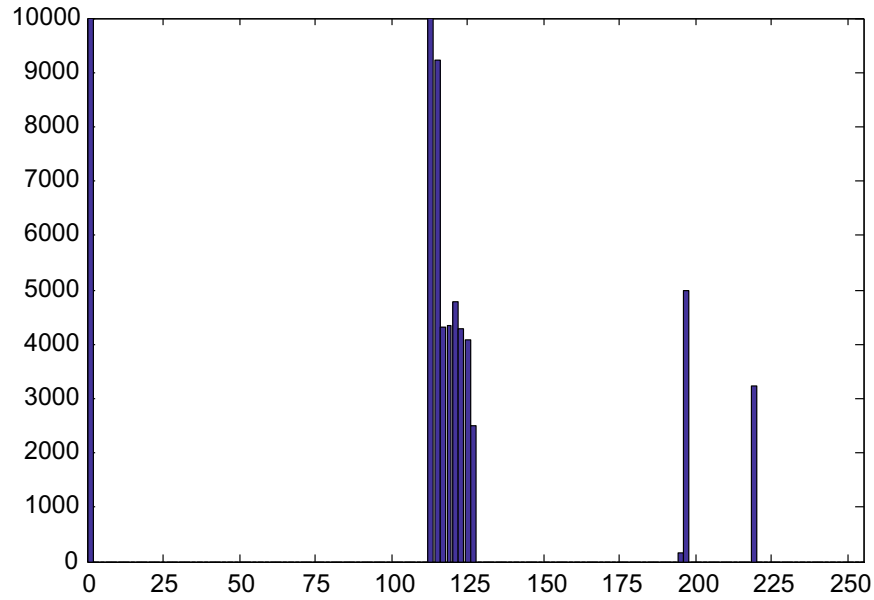


Fig. 15 Histogram of fifth image

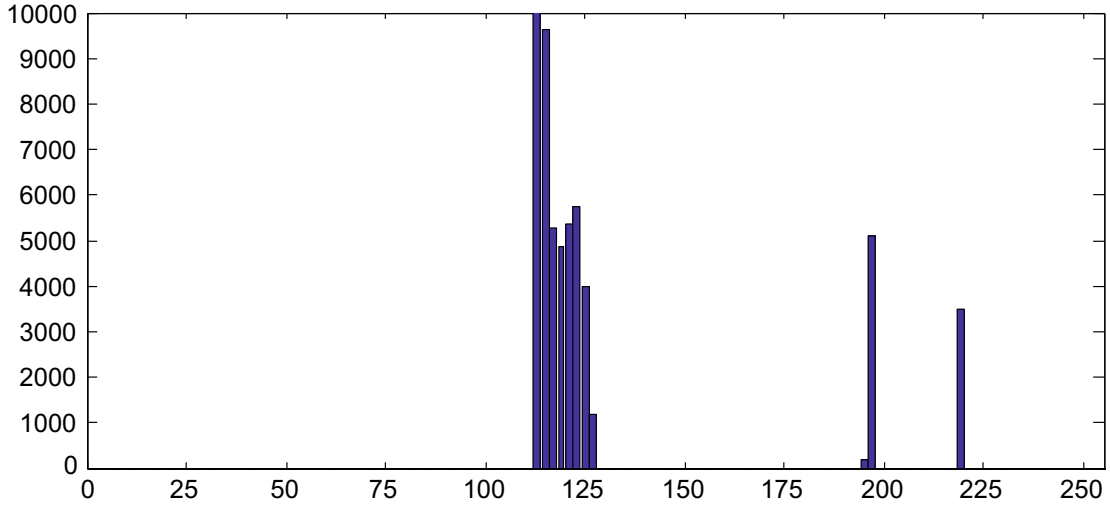


Fig. 16 Histogram of sixth image

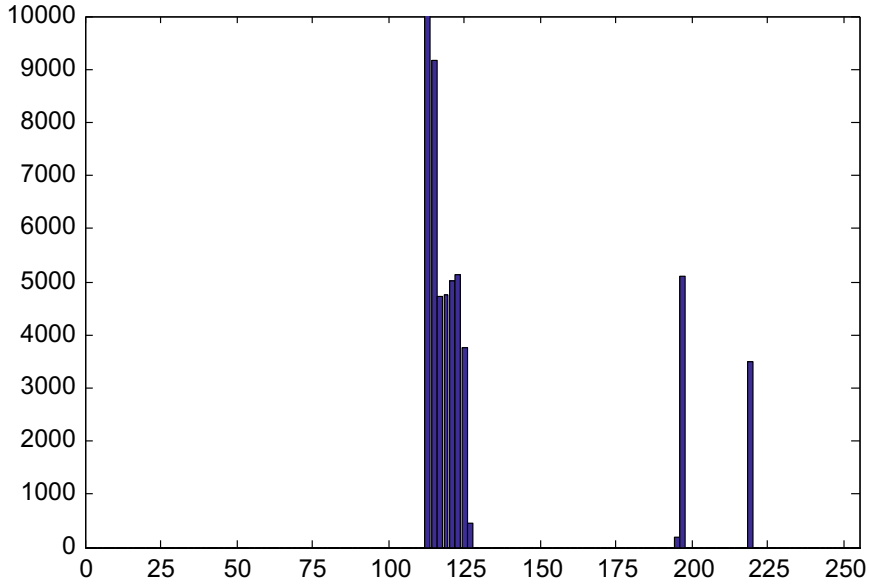


Fig. 17 Histogram of seventh image

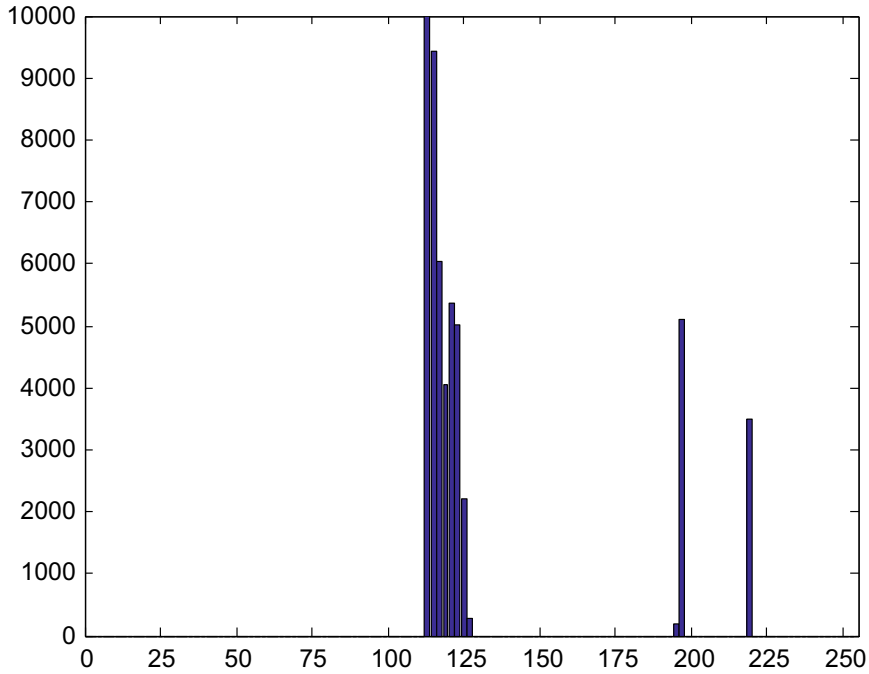


Fig. 18 Histogram of eighth image

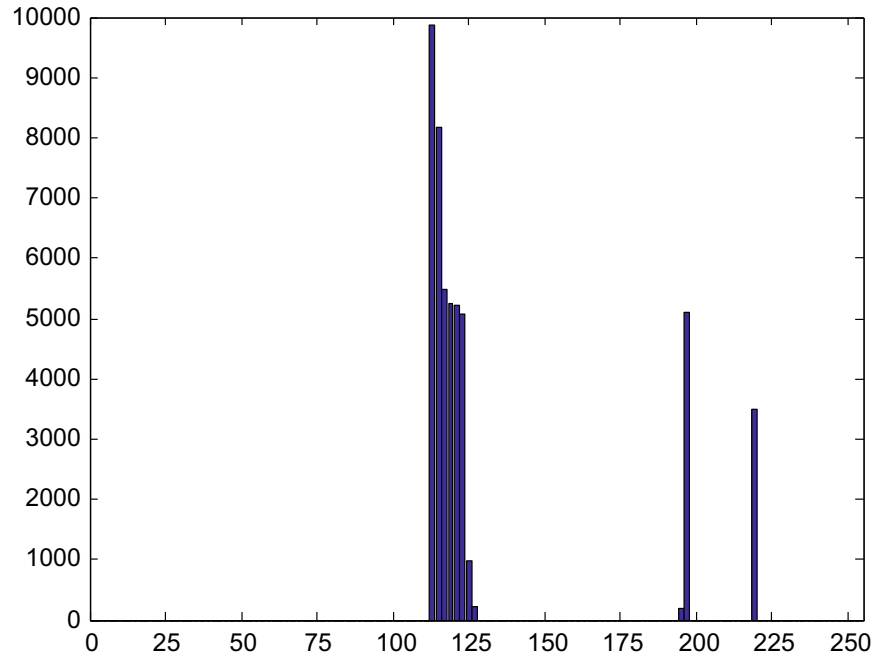


Fig. 19 Histogram of ninth image

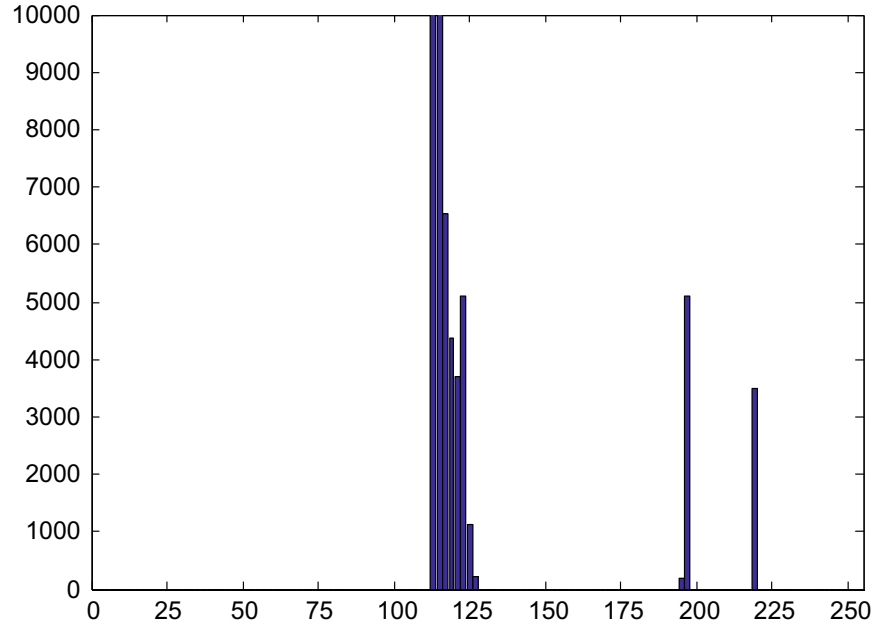


Fig. 20 Histogram of tenth image

References

1. Mladjen, C., Janc, D., Vujovic, D., Vuckovic, V.: The effects of a river valley on an isolated cumulonimbus cloud development. Atmos. Res. **66**, 123–139 (2003)
2. Lin, Y.L., Joyce, L.E.: A further study of mechanisms of cell regeneration, development and propagation within a two-dimensional multicell storm. J. Atmos. Sci. **58**, 2957–2988 (2001)
3. Lin, Y.-L., Deal, R.L., Kulie, M.S.: Mechanisms of cell regeneration, propagation, and development within two-dimensional multicell storms. J. Atmos. Sci. **55**, 1867–1886 (1998)
4. Fovell, R.G., Tan, P.-H.: Why multicell storms oscillate. In: 18th Conference on Severe Local Storms. American Meteorological Society, San Francisco, CA, pp. 186–189. Preprints (1996)
5. Anil Kumar, P., Anuradha, B., Arunachalam, M.S.: Extraction of time series convective cloud profile from doppler weather radar MAX (Z) product using a novel image processing technique. Int. J. Adv. Eng. Res. Dev. **4**(7), 2348–4470 (2017)
6. Gil, J.Y., Kimmel, R.: Efficient dilation, erosion, opening, and closing algorithms. IEEE Trans. Pattern Anal. Mach. Intell. **24**(12), 1606–1617 (2002)
7. Skamarock, W.C., Klemp, J.B., Dudhia, J., Gill, D.O., Barker, D.M., Wang, W., Powers, J.G.: A description of the advanced research WRF, Version 2. NCAR Tech. Note NCAR/TN-4681STR, p. 88 (2008)
8. Gryschka, M., Witha, B., Etling, D.: Scale analysis of convective clouds. Meteorol. Z. **17**(6), 785–791 (2008)
9. Souza-Echer, M.P., Pereira, E.B., Bins, L.S., Andrade, M.A.R.: A simple method for the assessment of the cloud cover state in high-latitude regions by a ground-based digital camera. J. Atmos. Ocean. Technol. **23**(3), 437–447 (2006)
10. Hutchison, K.D., Hardy, K.R., Gao, B.C.: Improved detection of optically thin cirrus clouds in nighttime multispectral meteorological satellite using total integrated water vapor information. J. Appl. Meteorol. **34**, 1161–1168 (1995)
11. Jolivet, D., Feijt, A.J.: Cloud thermodynamic phase and particle size estimation using the 0.67 and 1.6 mm channels from meteorological satellites. Atmos. Chem. Phys. Discuss. **3**, 4461–4488 (2003)
12. Glantz, P.: Satellite retrieved cloud optical thickness sensitive to surface wind speed in the subarcticmarine boundary layer. Environ. Res. Lett. **5**, 034002 (2010). https://doi.org/10.1088/1748-9326/5/3/034002
13. Ghonima, M.S., Urquhart, B., Chow, C.W., Shields, J.E., Cazorla, A., Kleiss, J.: A Method for Cloud Detection and Opacity Classification Based on Ground Based Sky Imagery, pp. 4535–4569. Copernicus Publications, Göttingen (2012)
14. McAndrew, A.: An Introduction to Digital Image Processing with Matlab Notes for SCM2511 Image Processing 1. School of Computer Science and Mathematics Victoria University of Technology, Footscray (2004)
15. Himadri Chakrabarty, C.A., Murthy, S.Bhattacharya, Gupta, A.D.: Application of artificial neural network to predict squall-thunderstorms using RAWIND data. Int. J. Sci. Eng. Res. **4**(5), 1313–1318 (2013). (ISSN 2229-5518)

Artificial Intelligence and Data Analysis

PURAN: Word Prediction System for Punjabi Language News

Gurjot Singh Mahi and Amandeep Verma

Abstract This paper presents an outline of the PURAN: A state-of-the-art word prediction system for Punjabi language news. Word prediction systems are used to increase the user text composition rate while typing the text. Brief background of the various approaches utilized in the development of word prediction systems, while discussing the various factors affecting the development of such systems is provided. This paper also elaborates the word prediction system architecture in detail. The system performance was tested on Keystroke saving, Hit ratio, Average rank and Average keystrokes benchmark metrics. The paper demonstrates that the PURAN has achieved highest Hit ratio in Regional news genre followed by National news genre by making lowest average keystrokes in the said categories of news. The results show that system has achieved 88.38% Average Hit ratio with 51.42% Average keystroke saving for $N = 10$.

Keywords Word completion · Word prediction · Punjabi text composition

1 Introduction

Word prediction system enables the user to complete the partially entered string known as the prefix, by providing the selection among list of possible words. The word prediction system populates the list of possible words from the provided database for the given prefix. The word prediction systems instinctually help the user with automatic completion of the word. Thus word prediction system in turns helps to increase the text composition rate. Several software applications like PROFET [1], FASTY [2], PAL [24] were proposed to predict the word in English and other European languages. Text composition in Indian languages is relatively

G. S. Mahi (✉) · A. Verma
Department of Computer Science, Punjabi University, Patiala, Punjab, India
e-mail: gurjotmahi28@gmail.com

A. Verma
e-mail: vaman71@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_26

more tedious as compared to other languages due to numerous concerns. Large character database, different vowels symbols, and complex language syntax makes text composition difficult in Indian context [4]. The studies like hIndiA [4] and [29] were performed for other Indian languages—Hindi and Bengali.

This paper intends to present a word prediction system—PURAN, designed to predict the words in News items for Punjabi language. PURAN was designed to increase the text composition rate for Punjabi news composition for 5 news genres namely Business, International, National, Regional, and Sports. PURAN system architecture is divided into two phases—Corpus creation and statistical inference phase and word prediction phase. The PURAN system has performed well on all benchmark metrics when tested in simulated environment.

The paper is distributed into eight sections. Section 2 provides the background about the state of art in word prediction systems and various factors affecting its performance. Section 3 presents the proposed system architecture used to design PURAN word predictor. System architecture is elaborated using a working example in Sect. 4. Performance metrics used to test the system architecture are discussed in Sect. 5. Testing dataset and system configuration is given in Sect. 6. Explored results and discussion is mentioned in Sect. 7. Section 8 concludes the study.

2 Background

Many studies [1–5] have been carried out to develop specialized software applications known as word prediction systems to increase the users' text composition rate. The idea of the development of a word prediction system was started way back in 1968 when [9] published the technique to reduce the keystrokes for completing a word [10]. Most of the word prediction systems were developed to help the people with motor and speech disabilities to overcome the communication problems [11]. Garay-Vitoria and Abascal [11] also states that word selection based text-input interfaces helps the people having poor control over the limbs to communicate in regular conversations. The success of word prediction system depends on various factors, i.e., prediction methodology, prediction speed, dictionary structure, user interface used for demonstrating the results, number of suggestions in user interface, etc. Some of the main factors affecting word prediction are discussed as follows.

2.1 Methodology

The success of word prediction system heavily depends on the methodology used for the design of the system. Although, diverse strategies are used for designing prediction systems, but collectively predictors methodology can be divided into two main categories, (1) Statistical Predictor and (2) Syntactic Predictor [4].

Statistical Predictor. The statistical predictors rely on the n-gram language model for language information in statistical word prediction system [4]. Most of the word prediction systems designed were based on frequency calculation, where distinctive word unigram occurrences were taken into account. Garay-Vitoria and Abascal [12] suggested the method of using the frequencies for word prediction, in which a dictionary-based approach is used. Dictionary is composed of words (unique unigrams) with respective frequencies. Each time a word is accepted from the proposed list of words the frequency of the respective word is updated in the frequency. Similar frequency-based approach was used by [13–16] for the development of word prediction system.

Frequency-based completion system concentrates on the unigram language model for prediction of the word. Fazly [17] gave the predictor which was trained on very large English language corpus using unigram and bigram language model. Unigram language model only concentrate on the current word being typed and bigram model approximated the probability of one word in the past for word prediction. Also, trigram (which takes the two previous words into account for prediction) is the baseline language model used for the word prediction systems [4, 18]. The main disadvantage of trigram model is that it requires large training data and demand huge memory space to store the processed corpus [11]. The present study used the frequency distribution method for the development of PURAN word prediction system.

Syntactic Predictor. Rule based syntactic structures of a natural language are followed for the design of syntactic word prediction systems. Frequency remains an important factor in addition to the other approaches during design of syntactic word prediction system, like [2] used the frequency table of words with syntactic knowledge i.e. Part-of-Speech (PoS) for more precision to design the prediction system—FASTY, for people suffering from motor and speech disability, enabling them to type 4 European languages, i.e., German, French, Dutch, and Swedish more precisely. Syntactic predictors like FASTY takes the probability of the grammatically correct Part-of-Speech (PoS) information into account for making any natural language word prediction. Another prediction system named New Profet [1] used lexicons for word completion for individuals suffering from linguistic impairments. [1] developed an efficient lexicon development algorithm, which helped in the creation of word lexicons from Stockholm-Umeå Corpus.[1] Furthermore, this algorithm facilitates the creation of new lexica from untagged or PoS tagged corpus.

In addition to the lexicon- and PoS-based syntactic word predictors, [19] developed a morphologically based word predictor for the Swedish language. An algorithm developed by the [19] allows users to select words by providing the base form of the word in a two-step process. Another syntactic prediction system like [20] and [21] provides a word completion or prediction system for Inflected languages. Syntactic predictors make the prediction using semantic knowledge about the natural language for which the word prediction system is being designed.

[1] https://spraakbanken.gu.se/eng/resources/suc.

2.2 Dictionary

An arrangement of the words with their frequencies or probability values requires a data structure for enhanced and firm results. Garay-Vitoria and Abascal [11, 22, 23] provide a detailed architecture about importance and usage of dictionaries in the word prediction systems. Dictionary is a general concept that is used to manage the keys and its values in an efficient manner. Keys in a dictionary data structure are implied as words and values are referred as words frequencies or probabilities. Word predictors designed using bigram or trigram language model require large storage capacity, which results in large dictionary size and complexity [22]. A single dictionary structure used for lexicon makes the access easy, but with the increase in number of dictionaries the problem complexity and access time for lexicon also increases [11, 22]. The lexicons are usually stored in the tree data structure, which helps in faster search but also have a complex organization structure.

Garay-Vitoria and Abascal [11] states that adapting dictionaries is an important factor while considering the natural language word prediction system. A single dictionary containing words and their respective frequencies are easy to maintain and update, as these frequencies can be easily updated in accordance to their usage in the real time.

2.3 User Interface

Newell et al. [24] gave detailed arguments about the importance of user interface in the selection of word from the suggestion list. The user interface provides an interactive space for a list of suggestions to appear from which user selects the most appropriate word that matches the user-specified prefix. Cognitive cost and suggestions demonstration are one of the factors related to the user interface. Point-of-gaze[2] plays an important role in increasing the production rate of a word predictor, as ease to locate the needed word in the list directly affects the hit ratio of the word prediction [24]. The design of user interface unswervingly affects the eye and hand movement during the selection of a word from suggestion list [4, 11, 24, 25]. Minimum eye and hand movement should be supported by the user interface [4]. Newell et al. [24], Swiffin et al. [15] suggests that vertical list of suggestions require less effort to select the word from the suggestion list. Swiffin et al. [15] argues that keeping the user interface immediately below the given prefix help in keeping minimum head and eye movement. Garay-Vitoria and Abascal [22], Koester and Levine [26] suggests that user interface suggestions can be listed in alphabetical order for better results. Garay-vitoria and Abascal [22] also mentions that triangular matrix provides the best average case access time to select the word

[2]"Eye tracking", https://en.wikipedia.org/wiki/Eye_tracking.

from the suggestion list, but difficult screen presentation makes them hard to use. Due to the extensive importance of word prediction system, user interface perspective is the uttermost important factor to look into.

2.4 Number of Suggestions

Another significant factor affecting the deliverance of words prediction system is the number of suggestions appearing for the user to select from. With the use of more suggestions in the suggestion list, a predictor can avail higher hit ratio, but will also increase the cognitive load on the user [11]. Garay et al. [21] points that suggestion list with ten suggestions provide stability between keystrokes and cognitive cost.

3 System Architecture

To perform the prediction of the word - ω, the system architecture is distributed into two phases, (1) Corpus creation and statistical inference phase and (2) word prediction phase. The phase of corpus creation and statistical calculation is performed only once whereas, phase 2 is initiated whenever the user enters any new prefix or update the prefix with a new set of characters.

3.1 Corpus Creation and Statistical Inference

The initial phase of the system architecture is divided into four diverse steps, (1) Extraction of Punjabi news corpus, (2) Preprocessing of the corpus, (3) Calculation of frequency distribution and (4) Formation of words Trie and frequency dictionary. Figure 1 illustrates the Corpus creation and statistical inference phase of system architecture.

Extraction of Punjabi news corpus. Corpora or Corpus[3] or Text corpus is referred to as a structured set of text, which is used to formulate some hypothesis about language features or perform statistical analysis on the language data. India is country of 1,324,171,354 people.[4] Such a large population is usually seen as a huge potential reader base by Indian news industry. In 2013–14, 99660 registered newspapers and periodicals were in circulation in India [6] with the reader base of 450586 [7] in various Indian languages. The findings by [8] suggest that there are more than 300

[3]"Text corpus", https://en.wikipedia.org/wiki/Text_corpus.

[4]World Population Prospects: The 2017 Revision". United Nations Department of Economic and Social Affairs, Population Division.

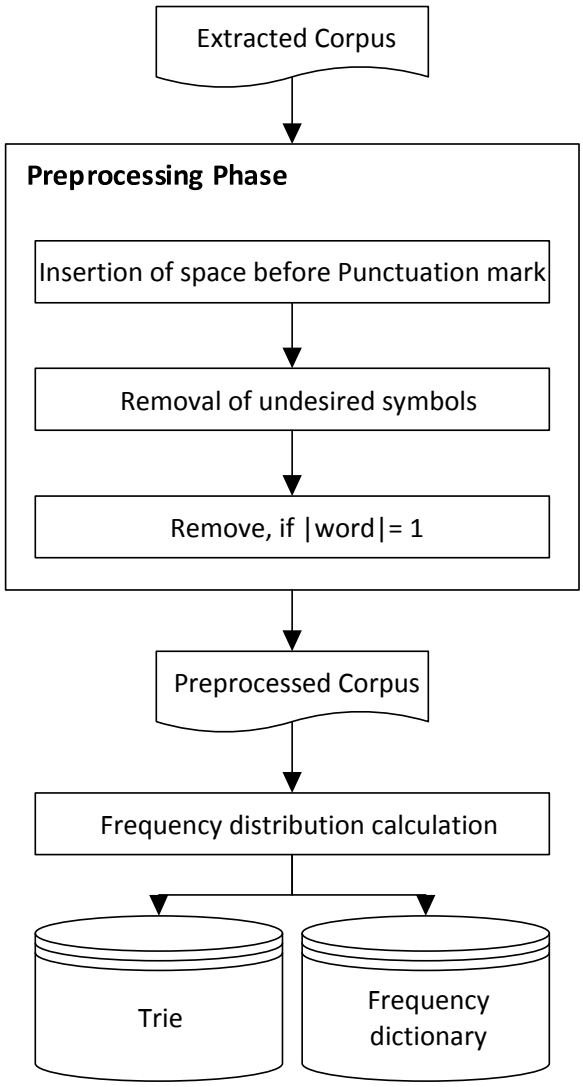


Fig. 1 Corpus creation and statistical inference

million internet users across India. Advancement of technology and low data rates has widely increased the viewership of e-paper and online newspapers across readers in India. Indian newspapers have also reinvented themselves to stay relevant in changing time and have started publishing news online in local Indian languages. One of the most popular Indian languages is Punjabi[5] language. Punjabi (ਪੰਜਾਬੀ [pañjābī]) is an Indo-Aryan language with more than 100 million speakers across the world. Punjabi is widely spoken in Punjab state of India and Pakistan, with speakers distributed across the world and is written using two distinct scripts—*Gurmukhi* and *Shahmukhi*. 1212 Newspapers and periodicals are published in Punjabi language (specifically using Gurmukhi script) in India [6] with enormous reader's base.

The Punjabi language news content extracted from Ajit,[6] a prominent Punjabi news website was used for developing the news corpus. The extracted corpus contains five genres of news, i.e., Business, International, National, Regional, and

[5]"Punjabi Language", https://en.wikipedia.org/wiki/Punjabi_language.

[6]"Ajit", https://www.ajitjalandhar.com/.

Table 1 Summary of extracted corpus statistics

Statistics	Before preprocessing	After preprocessing
Total words	1408041	1357443
Total characters	8795714	7075784
Average length of words	4.252	4.212
Unique words	65551	48521
Mode of words	4	4

Sports. This extracted corpus is further sent to preprocessing step for additional refinement.

Preprocessing of the corpus. The extracted corpus contains many undesired symbols and words, which does not hold any semantic meaning in the word prediction system. This makes the corpus to be preprocessed before any statistical inference could be obtained from the corpus. The preprocessing phase is divided into three stages, (1) Insertion of Space before a punctuation mark, (2) Removal of undesired symbols, and (3) Removal of words with length equal to one.

Insertion of Space before a punctuation mark. The preprocessing is instigated by inserting a space before Punjabi language punctuation mark—"।" [ɖan̪ ɖɑː]. In the extracted corpus, many sentences are ended by joining "।" [ɖan̪ ɖɑː] with no space before the preceding word, which create an ambiguousness in the corpus. This makes many words a diverse entity, even if similar words are present in the corpus. For example, in ਮੇਰੀ ਮਾਂ ਬੋਲੀ ਪੰਜਾਬੀ ਹੈ। [meːriː mɑːŋ boːliː pañjābī hɛːɖan̪ ɖɑː], punctuation mark "।" [ɖan̪ ɖɑː] is joined with the word "ਹੈ" [hɛː], which make a word "ਹੈ।" [hɛːɖan̪ ɖɑː] incorrect in accordance to the Punjabi language semantics, as "ਹੈ" [hɛː] is acknowledged as a complete word. Inserting one-character space before punctuation mark makes the words recognizable and accessible for processing in the later stage of statistical inference. For example, in the sentence—ਮੇਰੀ ਮਾਂ ਬੋਲੀ ਪੰਜਾਬੀ ਹੈ । [meːriː mɑːŋ boːliː pañjābī hɛː ɖan̪ ɖɑː], in which "ਹੈ" [hɛː] is now renewed into an individual word in the corpus and is available for further processing.

Removal of undesired symbols. Another stage in the preprocessing step comprises of the elimination of undesired symbols like @,!,(,), etc., as such symbols do not hold any semantic value during the statistical inference stage. After removal of undesired symbols, the processed corpus is sent to stage 3 for final processing.

Removal of words with length equal to one. All single length words were removed from the list, as such words are merely characters and do not hold any significant value in the corpus.

After completing all the stages of preprocessing on the extracted corpus, the remaining corpus is stated as the preprocessed corpus. Statistical analysis was conducted on preprocessed corpus and summary of the calculated statistics is shown in Table 1.

This preliminary preprocessing step can be stated more formally as,

Let c denotes the extracted corpus, i.e., unprocessed Bag of Words (BoW). The preprocessing phase on c can be stated as

$$\text{Preprocess(c)} = \acute{c} \tag{1}$$

where $\acute{c}$ is the processed corpus with N variables. This processed corpus ć is forwarded to step 2 to make statistical inference.

Calculating frequency distribution. Subsequently, after the preprocessing step, another significant step in the system architecture is to make the statistical inference. For this Frequency distribution is calculated from the preprocessed corpus ($\acute{c}$). The frequency distribution is an instantaneous description of the dataset which states the frequency of the discrete variables in the processed corpus. This step can be stated as,

Let n denotes the number of discrete variables x in processed corpus $\acute{c}$. Then $\text{VAR}_{\acute{c}}$ is denoted as the set of discrete variables as

$$\text{VAR}_{\acute{c}} = \{x_0, x_1, x_2, \ldots, x_n\} \tag{2}$$

The frequency of ith discrete variable in the corpus $\acute{c}$ can be given by

$$f_0 + f_1 + f_2 + \ldots + f_n = N \tag{3}$$

which furthermore can be given in form of a frequency set $\text{FREQ}_{\acute{c}}$,

$$\text{FREQ}_{\acute{c}} = \{f_0, f_1, f_2, \ldots, f_n\} \tag{4}$$

where, f_i is the frequency for *ith* discrete variable for $i = 0, 1, 2, \ldots, n$ in $\text{VAR}_{\acute{c}}$.

Set $\text{VAR}_{\acute{c}}$ and $\text{FREQ}_{\acute{c}}$ are further utilized in step 3 of a system architecture for the generation of word Trie and Frequency dictionary.

Formation of words Trie and frequency dictionary. Set $\text{VAR}_{\acute{c}}$ and $\text{FREQ}_{\acute{c}}$ are utilized in the final step for the formation of word Trie and frequency dictionary. Frequency dictionary named DICT_{PUN} is modeled in form of "<word>": <frequency>, in which <word> corresponds to the distinct variable or word in the set $\text{VAR}_{\acute{c}}$ and <frequency> corresponds to its relative frequencies which is extracted from set $\text{FREQ}_{\acute{c}}$. DICT_{PUN} is utilized to extract the frequencies of distinct words in O(1) time. Extracted distinct variables in a set $\text{VAR}_{\acute{c}}$ are further saved in the form of word Trie data structure named $(\text{TRIE}_{\text{PUN}})$ for an efficient words organization and extraction. The algorithm given in [27] is applied for the development of words Trie data structure $(\text{TRIE}_{\text{PUN}})$. The TRIE_{PUN} and DICT_{PUN} are further utilized in phase 2 of system architecture for generation of suggestion list and ranking of words.

3.2 Word Prediction

After phase 1 of system architecture, another phase is to predict the user intended word starting with specified prefix, i.e., string of character/s. This phase is termed as Phase 2. This phase is divided into three steps, (1) Initialization of prefix,

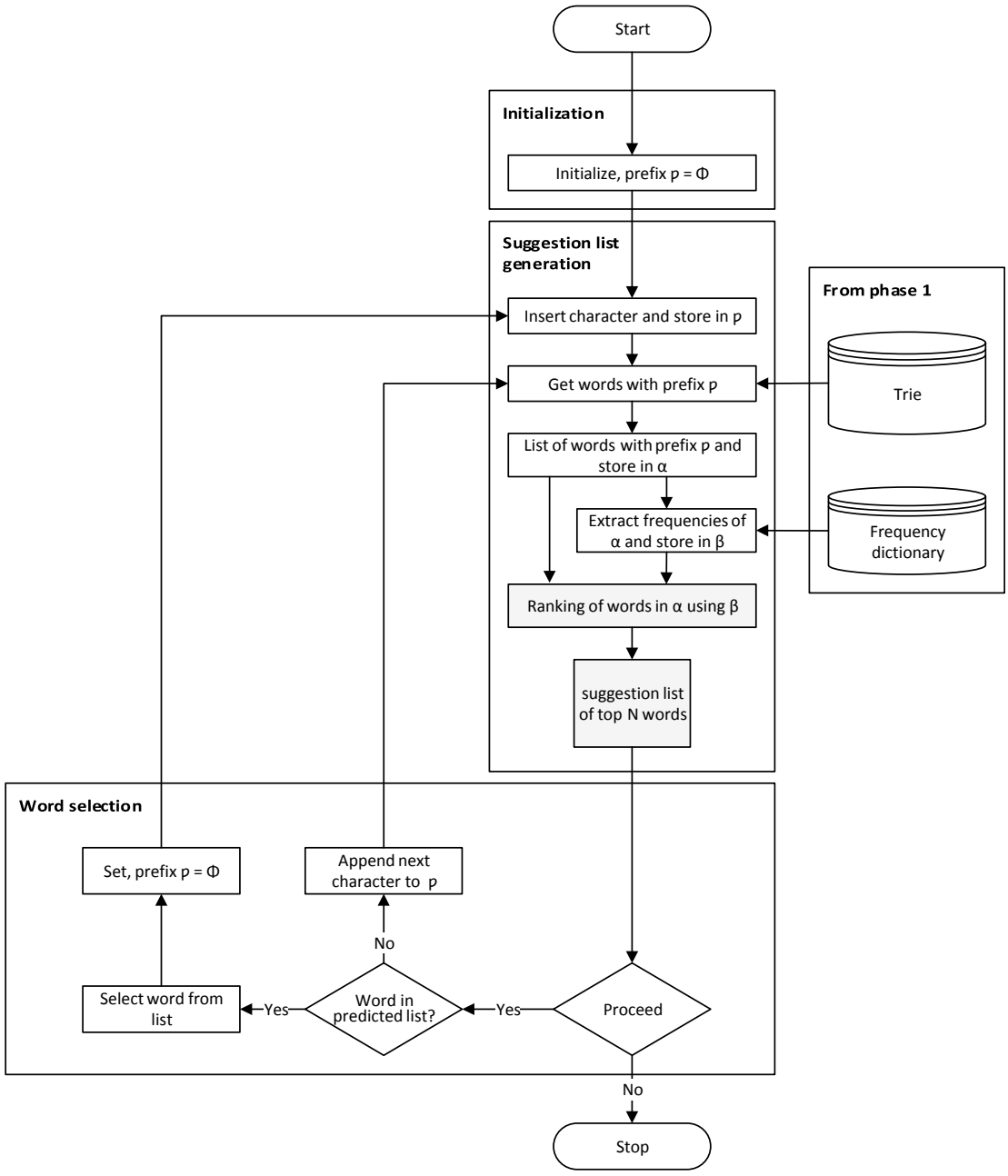


Fig. 2 Word prediction

(2) Suggestion list generation, and (3) Word selection. Figure 2 illustrates the final word prediction phase of system architecture.

Initialization of prefix. The initial phase in this step starts by initializing the prefix with null. If p denotes the user entered prefix, then this step can be stated as,

$$p = \Phi \tag{5}$$

This prefix (p) is further utilized in step 2 of the word prediction phase.

Suggestion list generation. TRIE_{PUN} and DICT_{PUN} generated in step 4 of phase 1 is utilized in this step. Subsequently, user-intended first character is stored in prefix $-p$ and is used to extract all the similar words in list $-\alpha$, from TRIE_{PUN}, which more formally is known as set $\text{VAR}_{\acute{c}}$ in Eq. 2. Furthermore, list of words in α are used to extract frequencies in another list $-\beta$, using frequency dictionary DICT_{PUN}, which more formally is stated as set $\text{FREQ}_{\acute{c}}$ in Eq. 4. The words in α are ranked against the extracted frequencies in β using the Heap Sort algorithm [27], from higher frequency to lower frequency. Algorithm 1 is used to accomplish this task. This α list of top N suggestions is listed for user selection. This system architecture was tested for 3, 5, 10, 15, 20, and 25 values of N in Sect. 7 of this article.

ALGORITHM 1: List Generation Algorithm

INPUT: Prefix(p) of the word to be predicted and dictionary (punjabi_dict) of Punjabi words as keys corresponding to their frequencies as values

OUTPUT: Returns the list of predicted words(α) to be generated

1. α = TRIE(p) // List of all the words to be generated with prefix p
2. β = [] // Empty list for frequencies
3. **for** each $\tilde{\omega} \in \alpha$ **do**
   β.append(punjabi_dict[$\tilde{\omega}$]) // extract frequency β for corresponding word $\tilde{\omega}$
4. HEAPSORT(α,β) // Perform heapsort on α on the basis of β
5. Return α // Return the words in order form of frequencies

The problem of suggestion list generalized more formally by [10] as follows, Let $L_{\text{PRED}}(p)$ denotes the set of extracted words from TRIE_{PUN} such that

$$L_{\text{PRED}}(p) \in \text{VAR}_{\acute{c}} \tag{6}$$

$L_{\text{PRED}}(p)$ is seen as a set of k suggestions from set $\text{VAR}_{\acute{c}}$. The foremost job of said system architecture is to rank the predicted suggestions in set $L_{\text{PRED}}(p)$ to set $R_{\text{PRED}}(p)$ from higher to lower, such that $|R_{\text{PRED}}(p)| = N$, where $\text{N} > 0$, represents the top suggestions of set $L_{\text{PRED}}(p)$ to be included in set $R_{\text{PRED}}(p)$, for which cost function $\text{COST}(R_{\text{PRED}}(p))$ for ranking set $R_{\text{PRED}}(p)$ is formulated as

$$\text{COST}(R_{\text{PRED}}(p)) = \sum_{\omega \epsilon R_{\text{PRED}}(p) \subset L_{\text{PRED}}(p)} \text{cost}(\omega) \tag{7}$$

where $\text{cost}(\omega)$ represent the cost of suggesting ω.

The purpose of step 2 of word prediction phase is to rank the words extracted in set $L_{\text{PRED}}(p)$ with minimum rank in the set $R_{\text{PRED}}(p)$ which is purposely known as the cost of selecting the word $-\omega$.

Word selection. After step 2, the last step is to select the user intended word $-\omega$ from N words in the list α. If the intended word $-\omega$ is found the list, the word is

selected from the list and prefix is again initialized to null, i.e., $p = \Phi$. If user intended word is not found in the list, next character is appended to prefix $-p$.

Step 2 and 3 of phase 2 is iterated till user intended word $-\omega$ is not found for the updated value of p or till system does not exit. The maximum number of iterations made for predicting single word depends upon the length of word needed to be predicted, i.e., if $|\omega| = l$, then the number of iterations needed would be l.

4 Working Example

The intended system functioning is described using Table 2. It is assumed that the proposed system wants to predict the user intended word (ω)"ਸੋਨਾ" [so:nɑ:]. The system initiate by entering initial prefix p_0 to be "ਸ" [s]. The system extracts the words list (α_0) starting with p_0 from $TRIE_{PUN}$ as ['ਸਿੰਘ', 'ਸੀ', 'ਸਰਕਾਰ', 'ਸ਼ੁਰੀ', 'ਸਨ'] and list of corresponding list of frequencies β_0 from $DICT_{PUN}$ as [54442, 7132, 5418, 5100, 3982] from higher to lower. For every extracted word x_i in α_0, the corresponding frequency f_i is extracted in β_0 for every $i = 0, 1, 2, \ldots, n$. This initial process is termed as iteration 0.

Table 2 System architecture working example

User intended word (ω)		ਸੋਨਾ
Total characters in ω		4
Iteration 0	Prefix (p_0)	ਸ
	Total characters (c_0)	1
	Extracted words from $TRIE_{PUN}(\alpha_0)$	['ਸਿੰਘ', 'ਸੀ', 'ਸਰਕਾਰ', 'ਸ਼ੁਰੀ', 'ਸਨ']
	IPA Format (ipa_0)	['sɪn̪g^h', 'si:', 'sarkɑ:r', 'ʂ^{r}i:', 'sn']
	Extracted frequencies from $DICT_{PUN}(\beta_0)$	[54442, 7132, 5418, 5100, 3982]
Iteration 1	Prefix (p_1)	ਸੋ
	Total characters (c_1)	2
	Extracted words from $TRIE_{PUN}(\alpha_1)$	['ਸੋਢੀ', 'ਸੋਚ', 'ਸੋਨੂੰ', 'ਸੋਧ', 'ਸੋਨੀ']
	IPA format (ipa_1)	['so:ɖ^{h}i:', 'so:c', 'so:nu:̃', 'so:d^h', 'so:ni:']
	Extracted frequencies from $DICT_{PUN}(\beta_1)$	[198, 126, 116, 112, 112]
Iteration 2	Prefix (p_2)	ਸੋਨ
	Total characters (c_2)	3
	Extracted words from $TRIE_{PUN}(\alpha_2)$	['ਸੋਨੂੰ', 'ਸੋਨੀ', 'ਸੋਨੀਆ', 'ਸੋਨੇ', 'ਸੋਨਾ']
	IPA format (ipa_2)	['so:nu:̃', 'so:ni:', 'so:ni:ɑ:', 'so:ne:', 'so:nɑ:']
	Extracted frequencies from $DICT_{PUN}(\beta_2)$	[116, 112, 76, 48, 44]

If the intended word "ਸੋਨਾ" [soːnɑː] is not found in α_0, next character is appended to the prefix p_0 and is termed as p_1. Again system process is repeated for p_1, the said process is termed to be iteration 1. In iteration 1, α_1 and β_1 is extracted as ['ਸੋਢੀ', 'ਸੋਚ', 'ਸੋਨੂੰ', 'ਸੋਧ', 'ਸੋਨੀ'] and [198, 126, 116, 112, 112]. Again system process is repeated as the intended word is not found in α_1. Iteration 2 is performed by appending next characters to p_1 and terming it as p_2. The predicted system is again tested to extract the user intended word using three characters in the prefix p_2. α_2 and β_2 list is extracted to be, ['ਸੋਨੂੰ', 'ਸੋਨੀ', 'ਸੋਨੀਆ', 'ਸੋਨੇ', 'ਸੋਨਾ'] and [116, 112, 76, 48, 44] for iteration 2.

In iteration 2 the user-intended word is found at rank 5 in list α_2 by making three keystrokes. One keystroke is saved using the described system architecture in Sect. 3.

5 Performance Metrics

For the assessment of proposed system architecture, testing metrics proposed by [11] are used. The metrics used for the assessment are discussed as follows:

Hit ratio. It is a metric that is used to describe the reliability of the word prediction system. Higher the hit ratio, higher will be the credibility of the prediction system to predict the correct word. It is defined by formula 8,

$$\text{Hit ratio} = \frac{\text{number of times words is predicted}}{\text{total number of written words}} \tag{8}$$

in which, number of times words are predicted is the calculation of the words prediction before the length actual word.

Keystroke saving. Hit ratio is not only enough to measure the actual saving of keystrokes as the predictor can guess words after entering all the characters. So, to measure the keystroke saving, the following formula 9 is used,

$$\text{Keystroke saving} = 1 - \frac{\text{number of keystrokes made}}{\text{total length of word}} \tag{9}$$

which is defined as the number of keystrokes saved during the course of prediction of the word. It can be defined as follows if the word comprises of n characters, and predictor predicts the word by making m strokes, the number of keystrokes saved can be defined as 1—(m/n). It is depicted using the following Fig. 3.

If total keystrokes m is equal to n, then the resulting response would be 0, as no keystrokes are saved, as the resultant word is completed by making n strokes only.

Average rank. As the vertical list is used to demonstrate the predicted for user selection. The Average rank calculation mechanism is utilized to compute the average rank of the predicted words in the vertical list of suggestions. This helps in

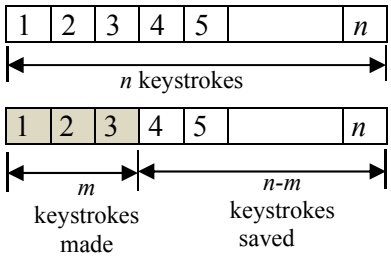


Fig. 3 Keystroke saving mechanism

calculating at which rank the selected word appears in the list. Formula 10 is used for calculation of average rank,

$$\text{Average rank} = \frac{\text{ranks total}}{\text{number of words}} \tag{10}$$

If during the prediction of the word is not found in the list, rank is set to be 0.

Average keystrokes. The average keystroke calculation mechanism is utilized to recognize the average number of keystrokes entered for each word during the prediction testing. Formula 11 is used for the calculation of average keystrokes,

$$\text{Average keystroke} = \frac{\text{total keystrokes made}}{\text{number of words}} \tag{11}$$

6 Testing Dataset and System Configuration

One of the prominent newspaper of Punjabi–Jagbani[7] was selected to test the performance of system architecture. Jagbani News crawler[8] was developed to extract the Punjabi news in accordance to the news genres. As the training corpus consists of five major genres of news, i.e. Business, International, National, Regional, and Sports, similar genres of news were extracted from Jagbani website to test the performance of the system. The statistics of the news items in each genre are given in Table 3.

The system architecture was implemented on the Windows 10 Operating system, Intel Core i5-6200U CPU 2.40 GHz with 8 GB RAM. The system is designed using Python programming language. NLTK and URLLIB packages are used for the implementation of said system and design of Web crawlers. The screenshot of the designed system is shown in Fig. 4.

[7] http://jagbani.punjabkesari.in/ .

[8] "Jagbani News crawler", https://github.com/GurjotSinghMahi/jagbani_website_crawler.

Table 3 Classification of news in testing corpus

Genre	News items
Business	22
International	65
National	48
Regional	115
Sports	31

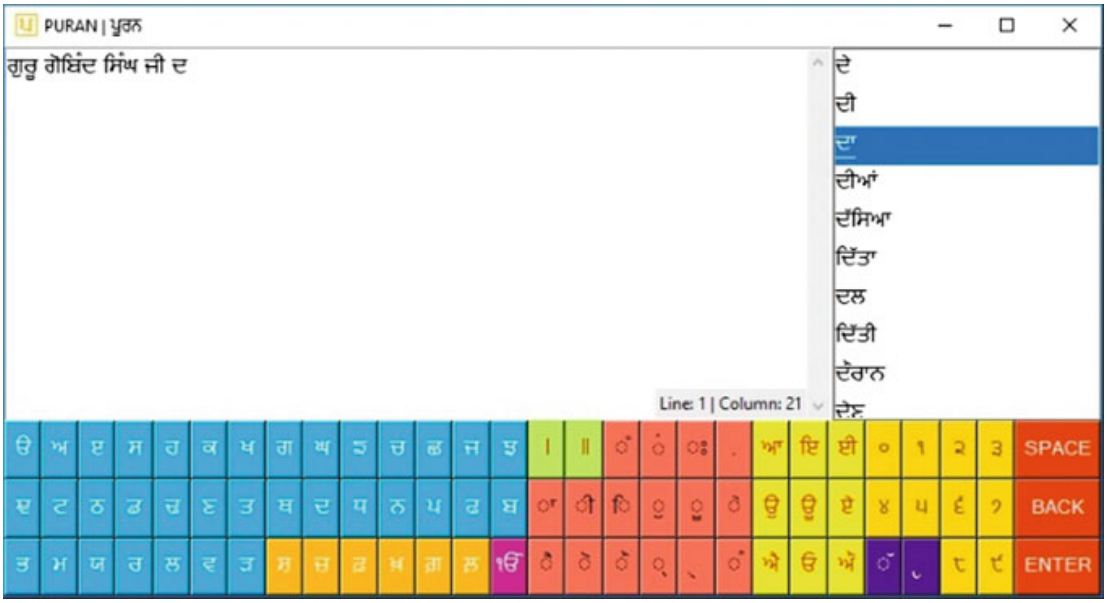


Fig. 4 Punjabi news words prediction system—PURAN

7 Results and Discussion

The performance of PURAN system was rigorously tested in the simulated environment on the basis of metrics mentioned from Eqs. (8) to (11). The extracted news items were passed to the PURAN news words prediction system for evaluation in accordance with the genres. The resultant system is observed by setting the number of words in suggestion list, i.e., N to be 3, 5, 10, 15, 20, and 25. The value of N shows the number of predicted words in accordance to their ranks, to choose from during performance analysis. Table 4 demonstrates the calculated metrics values for the proposed system architecture.

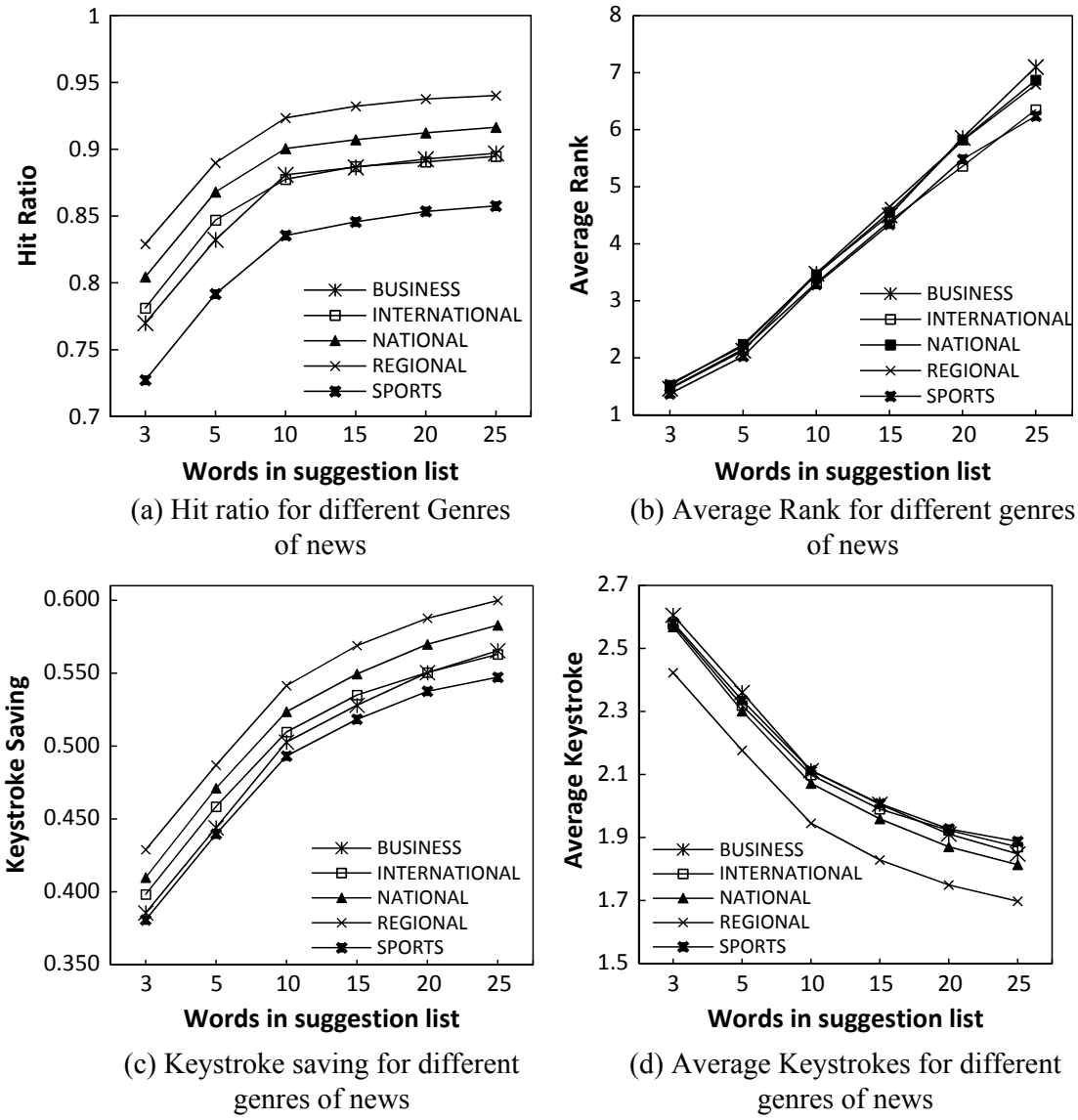
The observed scores are plotted using a line chart. Figure 5 shows PURAN performance for different genres of news in a simulated environment. Figure 5a, depicts the Hit ratio for different Genres of news, Fig. 5b illustrated the Average Rank for different genres of news, Fig. 5c depicts the Keystroke saving for different genres of news and Fig. 5d depicts the Average Keystrokes for different genres of news in a simulated environment.

Table 4 Calculated metrics for proposed system architecture

Metric	Genre	Words in the suggestion list					
		3	5	10	15	20	25
Hit ratio	Business	0.770	0.832	0.881	0.887	0.893	0.897
	International	0.781	0.847	0.878	0.887	0.891	0.895
	National	**0.805**	**0.868**	**0.901**	**0.907**	**0.912**	**0.916**
	Regional	**0.829**	**0.890**	**0.924**	**0.932**	**0.938**	**0.940**
	Sports	0.727	0.792	0.835	0.846	0.854	0.858
Average rank	Business	**1.459**	**2.125**	3.475	4.498	5.856	7.097
	International	1.464	2.151	**3.315**	**4.394**	**5.358**	**6.349**
	National	1.519	2.238	3.456	4.554	5.821	6.869
	Regional	1.544	2.207	3.481	4.649	5.811	6.788
	Sports	**1.373**	**2.027**	**3.290**	**4.342**	**5.480**	**6.238**
Keystroke saving	Business	0.386	0.444	0.503	0.528	0.550	0.565
	International	0.398	0.458	0.510	0.535	0.551	0.563
	National	**0.410**	**0.471**	**0.524**	**0.549**	**0.570**	**0.583**
	Regional	**0.429**	**0.487**	**0.541**	**0.569**	**0.588**	**0.600**
	Sports	0.381	0.440	0.493	0.518	0.538	0.547
Average keystrokes	Business	2.606	2.360	2.112	2.006	1.911	1.848
	International	2.577	2.319	2.099	1.990	1.923	1.870
	National	**2.569**	**2.301**	**2.072**	**1.959**	**1.871**	**1.814**
	Regional	**2.423**	**2.177**	**1.945**	**1.829**	**1.750**	**1.698**
	Sports	2.582	2.336	2.113	2.008	1.928	1.888

The bold text highlight the top values in the particular range

As the above results clearly demonstrate that the PURAN word predictor has performed very well on all the benchmark metrics. From Fig. 5a and d, it can be clearly derived that, PURAN has achieved highest Hit ratio in Regional news genre followed by National news genre by making lowest average keystrokes for predicting the words. These results clearly demonstrate that words can be easily predicted using PURAN words predictor for several mentioned news genres with high hit ratio and low keystrokes in Regional and National news genres. Figure 5c demonstrates that PURAN has given outstanding results for keystroke saving for Regional news genre and trailed by National genre, which is above the benchmark of 49.1%, 47%, and 40.5% keystroke saving given by [1, 16, 28], when N is taken as 10, 15, 20, and 25. Subsequently, it can be clearly stated that with the increase in the number of words in suggestion list the average keystroke saving also increases. Figure 5b states that with the increase in a number of suggestions in a vertical list of predicted words the average rank also increases.



(a) Hit ratio for different Genres of news

(b) Average Rank for different genres of news

(c) Keystroke saving for different genres of news

(d) Average Keystrokes for different genres of news

Fig. 5 PURAN performance for different genres of news in a simulated environment

8 Conclusion

Most the word prediction systems earlier designed were for the English and other European languages. This work is an effort to accomplish Punjabi Language with its first prediction system, particularly in News category. This paper presents the reader with the word prediction background with various factors affecting the word predictor accuracy, i.e., predicting methodology, a number of suggestions, user interface, and other features. The overall applied architecture used for suggesting the user the most appropriate word based on the prefix provided by the user was demonstrated. The system performance is tested on the various benchmark metrics like *Keystroke saving*, *Hit ratio*, *Average rank,* and *Average keystrokes* for rigorous review of the proposed system to examine its credibility in Punjabi news category.

PURAN prediction system works well in the categories of Regional and National news genres followed by International, Business, and Sports genres.

Garay et al. [21] states that no more than 10 suggestions must be presented to the user in suggestion list considering more physical movement, i.e., more hand and eye movement during the selection of the desired word. The authors also endorse the idea of [21] and take only 10 words for Final system architecture. The system has achieved 88.38% Average Hit ratio with 51.42% Average keystroke saving for $N = 10$. It is meaningful to mention while concluding that PURAN word predictor can be easily used to write articles for the Punjabi newspapers with an ease and also with high hit ratio and keystroke saving.

Acknowledgements The Authors would like to acknowledge the contribution of Arvinder Singh Kang from UrbanLogiq for providing the Punjabi news corpus for this research. We thank Varun Manchanda, Faculty, RGNIYD Regional Centre, Chandigarh for proofreading the draft.

References

1. Carlberger, A., Carlberger, J., Magnuson, T., Hunnicutt, M.S., Palazuelos-cagigas, S.E., Navarro, S.A.: Profet, a new generation of word prediction: an evaluation study. In: Natural Language Processing for Communication Aids, pp 23–28 (1997)
2. Matiasek, J., Baroni, M., Trost, H.: FASTY—a multi-lingual approach to text prediction. In: International Conference on Computers for Handicapped Persons, pp 243–250. Springer, Berlin (2002)
3. Dunlop, M.D., Crossan, A.: Predictive text entry methods for mobile phones. Pers. Technol. **4** (2–3), 134–143 (2000)
4. Sharma, M.K., Samanta, D.: Word prediction system for text entry in Hindi. ACM Trans. Asian Lang. Inf. Process. **13**(2), 1–29 (2014)
5. Arnold, K.C., Gajos, K.Z., Kalai, A.T.: On suggesting phrases vs. predicting words for mobile text composition. In: Proceedings of the 29th Annual Symposium on User Interface Software and Technology—UIST'16, pp 603–608 (2016)
6. Number of registered newspapers and periodicals by language and periodicity. http://www.mospi.gov.in/sites/default/files/statistical_year_book_india_2015/Table-36.1.xlsx
7. Claimed circulation of registered newspapers. http://www.mospi.gov.in/sites/default/files/statistical_year_book_india_2015/Table-36.4.xlsx
8. 11th annual report 2014–15 (2015). http://www.iamai.in/sites/default/files/annual_report/AnnualReport2014-15.pdf. Accessed: 06 May 2018
9. Longuet-Higgins, H.C., Ortony, A.: The adaptive memorization of sequences. In: Proceedings of the 3rd Annual Machine Intelligence Workshop, pp 311–322 (1968)
10. Cai, F., de Rijke, M.: A survey of query auto completion in information retrieval. Found. Trends® Inf. Retr. **10**(4), 273–363 (2016)
11. Garay-Vitoria, N., Abascal, J.: Text prediction systems: a survey. Univ. Access Inf. Soc. **4**(3), 188–203 (2006)
12. Garay-Vitoria, N., González-Abascal, J.: Using statistical and syntactic information in word prediction for input speed enhancement. In: Information Systems Design and Hypermedia, pp 223–230 (1994)
13. Heckathorne, C.W., Childress, D.S.: Applying anticipatory text selection in a writing aid for people with severe motor impairment. IEEE Micro **3**(3), 17–23 (1983)

14. Hunnicutt, S.: Input and output alternatives in word prediction. STL-QPSR **2**(3), 015–029 (1987)
15. Swiffin, A.L., Arnott, J.L., Pickering, J.A., Newell, A.F.: Adaptive and predictive techniques in a communication prosthesis. Augment. Altern. Commun. **3**(4), 181–191 (1987)
16. Venkatagiri, H.: Efficiency of lexical prediction as a communication acceleration technique. Augment. Altern. Commun. **9**(3), 161–167 (1993)
17. Fazly, A.: The use of syntax in word completion utilities. M.Tech. Thesis, Department of Computer Science, University of Toronto (2002)
18. Jurafsky, D., Martin, J.H.: Speech and Language Processing. Prentice Hall, New Jersey (2000)
19. Bertenstam, J., Hunnicutt, S.: Adding morphology to a word predictor. In: The European Context for Assistive Technology (Proceedings of the 2nd Tide Congress), pp 312–315 (1995)
20. Garay-Vitoria, N., Abascal, J.G.: Word prediction for inflected languages application to basque language. In: Natural Language Processing for Communication Aids, pp 29–36 (1997)
21. Garay, N., Abascal, J., Gardeazabal, L.: Evaluation of prediction methods applied to an inflected language. In: International Conference on Text, Speech and Dialogue, pp 397–403 (2002)
22. Garay-vitoria, N., Abascal, J.: A comparison of prediction techniques to enhance the communication rate. In: ERCIM Workshop on User Interfaces for All, pp 400–417 (2004)
23. Garay-Vitoria, N., Abascal, J.: User interface factors related to word-prediction systems. In: Proceedings of the 7th International Conference on Work with Computing Systems WWCS2004, pp 77–82 (2004)
24. Newell, A., Arnott, J., Booth, L., Beattie, W., Brophy, B., Ricketts, I.: Effect of the 'PAL' word prediction system on the quality and quantity of text generation. Augment. Altern. Commun. **8**(4), 304–311 (1992)
25. MacKenzie, I.S., Tanaka-Ishii, K.: Text entry using a small number of buttons. In: MacKenzie, I.S., Tanaka-Ishii, K. (eds.) Text Entry Systems: Mobility, Accessibility, Universality, pp. 105–121. Morgan Kaufmann, San Francisco, CA (2007)
26. Koester, H.H., Levine, S.: Model simulations of user performance with word prediction. Augment. Altern. Commun. **14**(1), 25–36 (1998)
27. Brass, P.: Advanced Data Structures, vol. 1. Cambridge University Press, Cambridge (2008)
28. Lesher, G.W., Moulton, B.J., Higginbotham, D.J.: Techniques for augmenting scanning communication. AAC Augment. Altern. Commun. **14**(2), 81–101 (1998)
29. Ghosh, S., Sarcar, S., Samanta, D.: Designing an efficient virtual keyboard for text composition in Bengali. In: Proceedings of the 3rd International Conference on Human Computer Interaction, April, pp 84–87 (2011)

Implementation of hDE-HTS Optimized T2FPID Controller in Solar-Thermal System

Binod Shaw, Jyoti Ranjan Nayak and Rajkumar Sahu

Abstract In this paper, a fair approach is interpreted to validate the novelty of Type-2 Fuzzy PID (T2FPID) controller over Type-1 or conventional Fuzzy PID (FPID) and PI controller as secondary frequency controller and Heat Transfer Search (HTS) algorithm is adopted to extract the optimum gains of the controllers. T2FPID controller has a beautiful property to handle large uncertainties of the system with extra degree of freedom. T2FPID controller is implemented in a two area interconnected thermal-PV power system to enhance the system performance. ITAE is adopted as objective function of the system to lessen the undershoot, overshoot, and settling time of frequency and tie-line power deviation. A novel hDE-HTS algorithm is adopted to enhance the system performance by searching the relevant pair of gains of controller. This analysis is executed by implementing a step signal (load disturbance) of magnitude 0.1 in area-2 to study the transiency of the system. The novelty of this work is to implement hDE-HTS optimized T2FPID controller to enhance the Solar-thermal system responses (Frequency and tie-line power deviations).

Keywords Automatic generation control (AGC) · Differential evolution (DE) · Fuzzy PID controller (FPID) · Heat transfer search (HTS) · PI controller · Type-2 fuzzy PID controller (T2FPID)

B. Shaw · J. R. Nayak (✉) · R. Sahu
Department of Electrical Engineering, NIT, Raipur, Chhattisgarh 492010, India
e-mail: bapi.jyoti.2@gmail.com

B. Shaw
e-mail: binodshaw2000@gmail.com

R. Sahu
e-mail: rajkumarsao@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_27

1 Introduction

In recent power system, interconnection of areas is an eminent approach to share the load demand along with to maintain stability and reliability of the system. This helps to supply reliable, stable, and economical power to the consumers. In interconnected power system, load deviation in one area causes frequency and power deviations in all interconnected areas. The small deviations are diminished by primary controller of generators. But for large deviations, secondary controller is highly essential to diminish the deviations [1]. Proportional (P), Integral (I), Derivative (D), PID, fuzzy, etc., are used as secondary controller to enhance the capability to handle the load fluctuations in the power system. The elementary purposes of the AGC are to

i. Settle the frequency deviation as soon as possible (i.e., $\Delta f = 0$).
ii. Settle the tie-line power deviation to its scheduled value (i.e., $\Delta P_{\text{tie}} = 0$).
iii. Minimize settling time, undershoot, and overshoot and of the system after any load fluctuation.

Renewable energy is an imperative alternative to fulfill the enormous expanding load demand. Solar system is the most preferable energy system due to its simple implementation. The combination of renewable energy systems enhances the complexity of the power system. In complex power system, intelligent controllers with enormous competence to handle uncertainties should prefer.

AGC is the most important strategy in the power system which concerns the reliable and stable power generation. Various secondary controllers have been implemented as AGC to enhance the system capability to handle the load fluctuation. Ibrahim et al. [2] have portrayed a review of AGC with different schemes to analyze the novelity of controllers. Abraham et al. [3] have implemented superconducting magnetic energy storage (SMES) to portray the effect of storage device on AGC in a two area interconnected hydro-thermal system. Singh and Sen [4] have successfully designed an interconnected power system with thermal units by using PID controller optimized by GA. Dash et al. [5] has introduced cascade controller titled as PI-PD controller optimized by FPA in four area interconnected reheat thermal power plants. Many researchers have enhanced the PID controller by reconstruct the PID controllers entitled as 2DOF PID, 3DOF PID, cascade PD-PID, etc. in [6–8] respectively. Even for better performance the fractional order PID controller is enforced as AGC in [9] and [12]. From last few decades, fuzzy logic controller is generally preferred for various applications. Fuzzy logic controller (FLC) has convinced as a very robust and intelligent controller used as AGC in [13, 14]. The robustness of FPID controller is enhanced with the assemblage of advantages from both PID and FLC controller. FPID controller optimized by different powerful algorithms have depicted in [15–19]. The degree of freedom of FLC is enhanced by providing Foot Of Uncertainty (FOU) entitled as T2FLC [20]. T2FPID controller is adopted to enhance the AGC performance in [21].

In present work, grid connected PV system is enforced in area-1 and reheat thermal plant is enforced in area-2. From last few decades, FPID controller has dominated over conventional controllers. T2FPID controller is an adequate controller over FPID controller to handle nonlinear and uncertain system. HTS algorithm is adopted to extract the controller gains. The superiority of hDE-HTS algorithm is validated over HTS to tune T2FPID controller.

2 System Investigated

In this work, two area interconnected power system is implemented. In area-1 PV system and in area-2 reheat thermal system are resided in the power system network as illustrated in Fig. 1 [10, 11]. The power system parameters are portrayed in Appendix 1. To verify the transiency of the system a 10% load change is enforced in area-2 which propagates an error in each area entitled as Area Control Error (ACE). The basic objective of AGC is to lessen the ACE which concedes deviations in tie-line power and frequency deviation as characterized in Eq. (1).

$$\text{ACE} = B\Delta f + \Delta P_{\text{tie}} \tag{1}$$

where B_1, and B_2 are the bias factors of frequency. The deviations of frequency with respect to nominal value in area-1 and area-2 are Δf_1 and Δf_2 respectively. The deviation of power in tie-line is ΔP_{tie} and is characterized in Eq. (2).

$$\Delta P_{\text{tie}} = \frac{2\pi T_{12}}{s}(\Delta f_1 - \Delta f_2) \tag{2}$$

PI, FPID, and T2FPID controllers are executed in both the areas individually to scrutinize the controller potency to enhance the system performance. Intelligent T2FPID controller is observed as superior controller over other conventional controllers due to its capability to handle nonlinear and uncertainties. The structure of T2FPID controller is as portrayed in Fig. 2. The objective function for this system by considering tie-line power deviation and frequency deviation is characterized in Eq. (3)

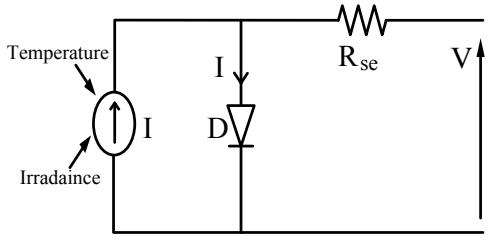


Fig. 1 Equivalent circuit solar system [11]

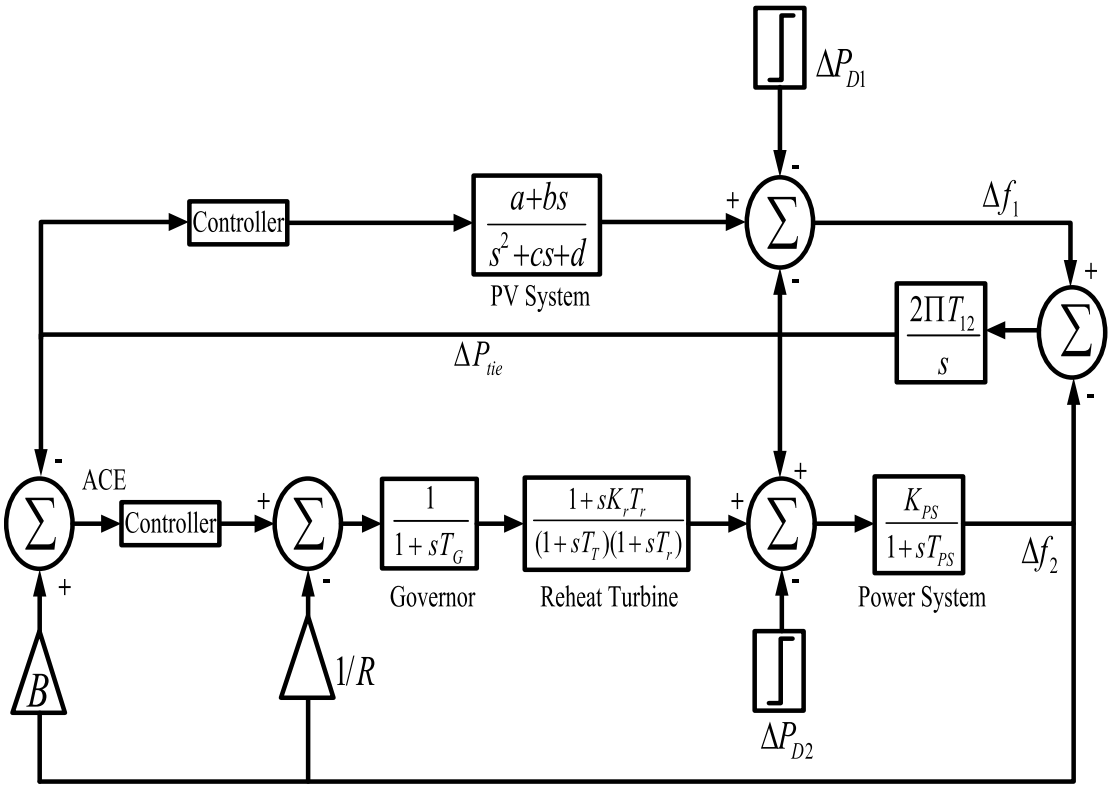


Fig. 2 Transfer function model of solar-reheat thermal power system [10] and [11]

$$J = \text{ITAE} = \int_0^T t(\Delta f_1 + \Delta f_2 + \Delta P_{\text{tie}}) \tag{3}$$

Subject to:

$$-2 \le K_i \le 2 \quad \text{i} = 1, 2, \ldots, \text{n}$$

PV System

The PV system is basically depends upon two important factors, i.e., irradiance and temperature. The PV system can be represented with a current source, diode and series resistance as portrayed in Fig. 1.

In present work, PV system is adopted with constant irradiance (1000 w/m^2) and temperature (27 °C). The transfer functions by concerning MPPT, inverter, and filter as characterized in Eq. (4).

$$G_{\text{PV}} = \frac{-18s + 900}{s^2 + 100s + 50} \tag{4}$$

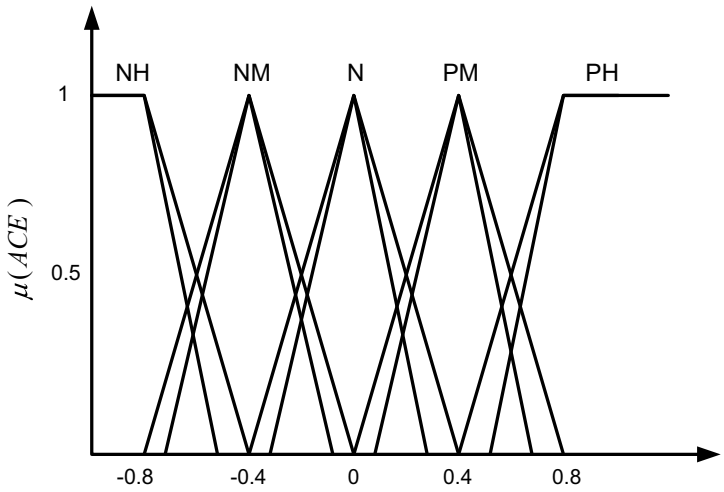


Fig. 3 MFs of Type-2 FLC

3 Type-2 Fuzzy PID Controller

Design of intelligent controllers is an imperative approach to provide reliable and stable power in a highly complex power system. T1FPID controller has substantiated as a better controller over classical controllers. In this work, T2FPID controller is adopted to contribute the supremacy of T2FPID controller over FPID and PID controller. The framework of T2FPID controller is portrayed in Fig. 5. The supremacy of T2FLC is better than T1FLC because of the FOU provided by the controller. FOU is the boundary provided by the two type-1 membership function (MF) entitled as Lower mf (LMF) and Upper mf (UMF) as portrayed in Fig. 4 [20]. The degree of freedom is enhanced by this FOU and also enhances the capability to handle the uncertain information. The MF adopted for T2FPID controller is illustrated in Fig. 3. In this work, five interval type-2 MFs are adopted entitled as negative huge (NH), negative mild (NM), Null (N), positive mild (PM), and positive huge (PH). The rule base is tabulated in Table 1.

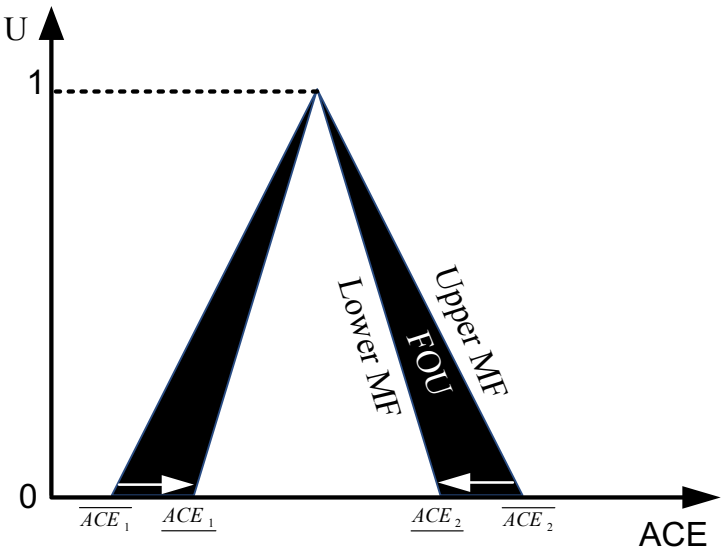


Fig. 4 Membership function of interval type-2

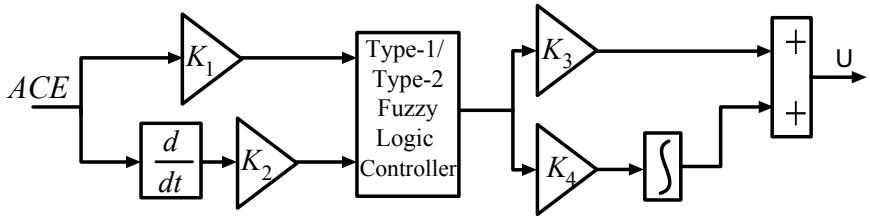


Fig. 5 Type-2 FPID controller structure

Table 1 Rule base

$\dot{e}$ / e	NH	NM	N	PM	PH
NH	NH	NH	NM	NM	N
NL	NH	NM	NM	N	PM
Z	NM	NM	N	PM	PM
PL	NM	N	PM	PM	PH
PH	N	PM	PM	PH	PH

4 hDE-HTS Algorithm

The basic purpose of using optimization technique is to design optimum controller which is best-suited for the system. In this work, a novel algorithm entitled as HTS is adopted to contribute highly convenient gain parameters within specified band to minimize "*J*" as illustrated in Eq. (3). Patel et al. [22] have introduced HTS algorithm stimulated from the law of thermodynamics and heat transfer. HTS algorithm is derived from the movement of molecules during three modes of heat transfer, i.e., conduction, convection, and radiation. HTS is a population-based iterative computational technique. The molecules, temperature and energy level are proportionate to the population, design variables and solution respectively. DE algorithm is very influential to enhance the diversity factor of the algorithm with a possibility to trap into local optima [23]. The combination of benefits of both algorithms is proficient enough as compare to individual one. The steps for hDE-HTS algorithm are as:

1. Initialize random population, i.e., $[X]_{\mathrm{NP}\times D}$.
2. Evaluate the fitness of the population, i.e., $f(X)$.
3. Initialize the conduction factor (CEF), convection factor (COF) and radiation factor (RF).
4. Donor vector (V_i) is defined as in Eq. (5).

$$V_i = X_{i,r_1} + F(X_{i,r_2} - X_{i,r_3}) \tag{5}$$

where X_{i,r_1}, X_{i,r_1} and X_{i,r_1} are three randomly taken vectors from X_i.

5. Offspring vector (U_i) used for crossover operation is characterized as in Eq. (6).

$$\begin{array}{ll} U_i = V_i & \text{If rand} \leq \text{cr} \\ = X_i & \text{Otherwise} \end{array} \tag{6}$$

6. Selection of target vector is defined as in Eq. (7).

$$\begin{array}{ll} X_i = U_i & \text{If } f(U_i) \leq f(X_i) \\ = X_i & \text{Otherwise} \end{array} \tag{7}$$

7. The solution of DE algorithm is considered as initial vector for HTS algorithm.
8. Initialize a random value R [0 1].
9. Conduction phase (if $R \leq 0.33333$)
   Update the particle position as characterized in Eqs. (8) and (9).

$$X_i^{g+1} = \begin{cases} X_k + (-R^2 * X_k), & \text{if } f(X_i^g) > f\left(X_k^g\right) \\ X_i + (-R^2 * X_i), & \text{if } f(X_i^g) < f\left(X_k^g\right) \end{cases} \quad \text{if } f(X_i^g) \leq f\left(X_{\text{best}}^g\right)/\text{CEF} \tag{8}$$

$$X_i^{g+1} = \begin{cases} X_k + (-r * X_k), & \text{if } f(X_i^g) > f\left(X_k^g\right) \\ X_i + (-r * X_i), & \text{if } f(X_i^g) < f\left(X_k^g\right) \end{cases} \quad \text{if } f(X_i^g) \geq f\left(X_{\text{best}}^g\right)/\text{CEF} \tag{9}$$

   where r, g and k are the random number, generation and random solution.
10. Convection phase (if $0.3333 \leq R \leq 0.6666$)
    Update the particle position as characterized in Eq. (10).

$$X_i^{g+1} = X_i^g + R * (X_{\text{best}} - TCF * X_i) \tag{10}$$

    Where TCF is the temperature change factor.

$$TCF = \begin{cases} abs(R - r), \text{if } f(X_i^g) \leq f\left(X_{\text{best}}^g\right)/COF \\ round(1 + r), \text{if } f(X_i^g) \leq f\left(X_{\text{best}}^g\right)/COF \end{cases}$$

11. Radiation phase ($R \geq 0.6666$)
    Update the particle position as characterized in Eqs. (11) and (12).

$$X_i^{g+1} = \begin{cases} X_i^g + R(X_k - X_i), \text{if } f(X_i^g) > f\left(X_k^g\right) \\ X_i^g + R(X_i - X_k), \text{if } f(X_i^g) < f\left(X_k^g\right) \end{cases}, if\, f(X_i^g) \leq f\left(X_{\text{best}}^g\right)/RF \tag{11}$$

$$X_i^{g+1} = \begin{cases} X_i^g + r(X_k - X_i), \text{if } f(X_i^g) > f\left(X_k^g\right) \\ X_i^g + r(X_i - X_k), \text{if } f(X_i^g) < f\left(X_k^g\right) \end{cases}, if\, f(X_i^g) \geq f\left(X_{\text{best}}^g\right)/RF \tag{12}$$

12. The best fitness particles updated for next generation by replacing the worst particles.
13. Repeat steps 4–12 until the termination criteria satisfied.
14. Evaluate the optimum solution.

The parameters of hDE-HTS algorithm are illustrated in Appendix 2.

5 Result and Discussion

HTS algorithm is adopted to design optimum PID, FPID, and T2FPID controller. K_1, K_2, K_3 and K_4 are the design variables to be tuned by HTS. A novel hDE-HTS algorithm is opted to tune T2FPID controller to validate the superiority of the hybrid controller. Both HTS and hDE-HTS algorithm are executed individually for different controllers. Each controller is optimized with 50 numbers of population and 100 generations. The basic purpose of execution of HTS algorithm is to lessen ITAE (Integral Time Absolute Error) as cited in Eq. (3). The optimal gain values of all controllers extracted by using both algorithms are tabulated in Table 2.

HTS algorithm optimized PI is validated as better controller than FA optimized PI controller [11]. In the same fashion, supremacy of HTS optimized FPID controller is substantiated by contrasting with ICA-optimized FPID controller [10].

The response parameters by conceding the settling time, undershoot and overshoot of both frequency and tie-line power deviation are tabulated in Table 3. In Table 3, a fair observation is illustrated to demonstrate the supremacy of T2FPID controller and hDE-HTS algorithm. The response parameters and ITAE of the T2FPID controller are minimum.

The deviations in frequency of both areas and tie-line power are portrayed in Figs. 6, 7, and 8 respectively.

Figures 6, 7, 8 and Table 3 precisely represent the superiority of the T2FPID controller. T2FPID controller minimizes the transiency (undershoot, overshoot and settling time) of the system as correlate to PI and FPID controllers.

Table 2 HTS optimized gain parameters

Controllers	Area-1				Area-2			
	K_1	K_2	K_3	K_4	K_1	K_2	K_3	K_4
HTS PI	−1.1861	−0.4544			−1.9474	−1.7060		
HTS FPID	−1.3502	−1.8061	1.7636	1.0369	2.0000	0.2045	−2.0000	−2.0000
HTS T2FPID	−1.0440	−2.0000	2.0000	1.5284	1.3001	1.1986	−2.0000	−1.6456
hDE-HTS T2FPID	−1.2922	−1.7420	1.6962	0.9909	2.0000	0.2031	−2.0011	−2.0000

Table 3 Response parameters such as undershoot (U_{sh}), overshoot (O_{sh}) and settling time (T_s) of the system

Controllers	Undershoot			Overshoot			Settling time			ITAE
	ΔF_1	ΔF_2	ΔP_{tie}	ΔF_1	ΔF_2	ΔP_{tie}	ΔF_1	ΔF_2	ΔP_{tie}	
PI FA [11]	−0.3109	−0.2308	−0.0460	0.1495	0.115	0.0459	25.76	26.91	20.58	5.464
PI HTS	−0.2064	−0.1641	−0.0357	0.0554	0.0859	0.0338	17.19	17.06	10.14	2.312
FPID ICA [10]	−0.1699	−0.1446	−0.0051	0.0207	0.0192	0.0117	16.34	16.35	6.40	1.595
FPID HTS	−0.1255	−0.1192	−0.0004	0	0.0043	0.0029	12.97	11.21	1.86	1.36
T2FPID HTS	−0.0418	−0.0411	−0.0003	0	0.0074	0.0008	12.35	11.16	0.789	1.04
T2FPID hDE-HTS	−0.01387	−0.0142	−0.0001	0	0.0007	0.0006	12.14	10.85	0.76	0.85

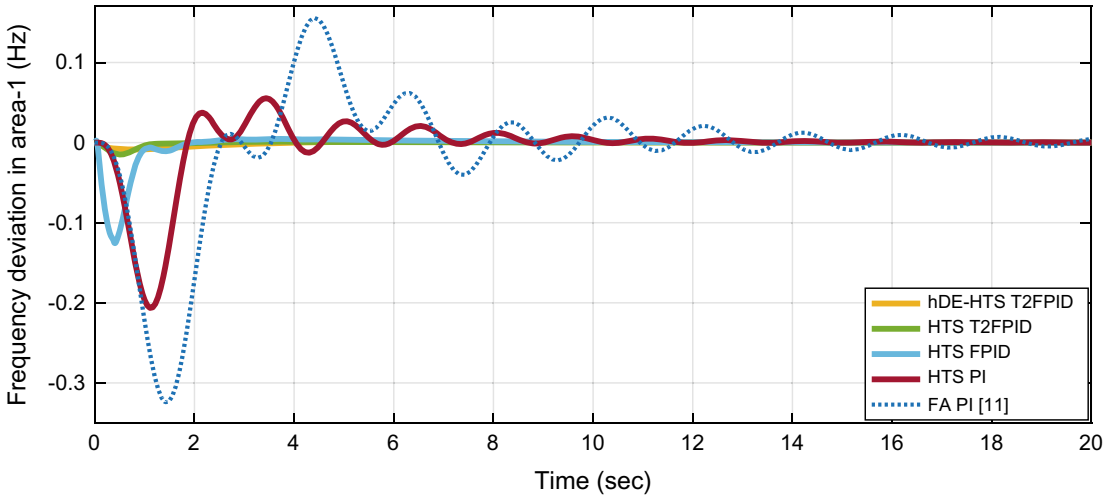


Fig. 6 Frequency deviation in area-1

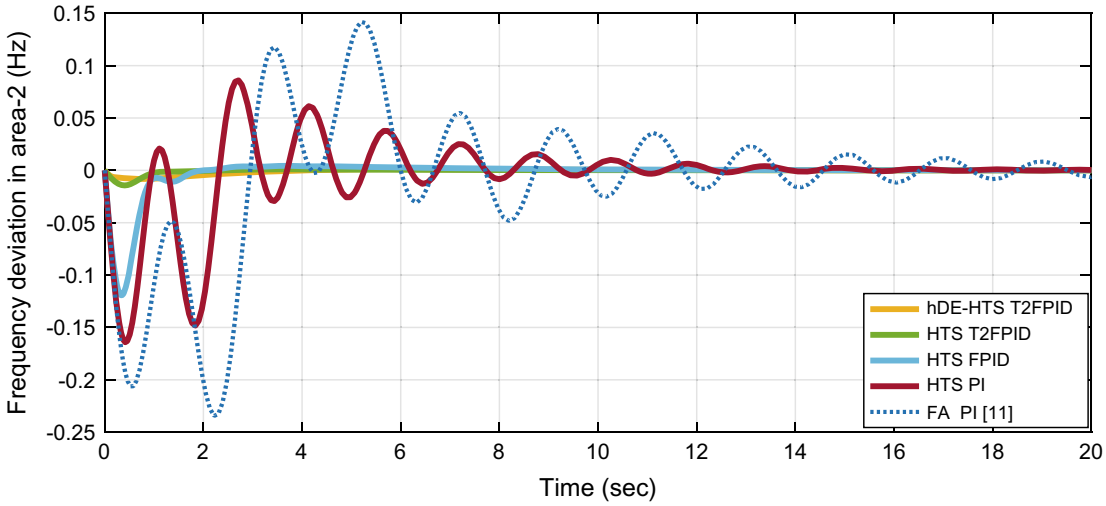


Fig. 7 Frequency deviation in area-2

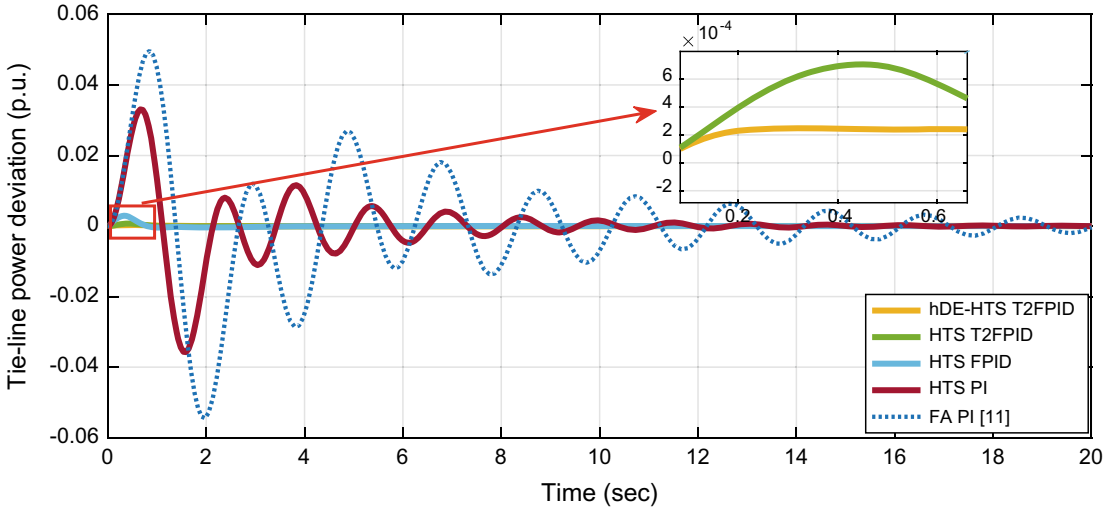


Fig. 8 Tie-line power deviation

6 Conclusion

In this work, the interconnected power system integrated with PV system with constant irradiance and temperature is simulated with various controllers. The purpose to contribute an intelligent control of frequency and power is achieved by employing PI, FPID, and T2FPID controllers. T2FPID controller yields better response over other controllers due to its capability to handle nonlinear and uncertain information. HTS and hDE-HTS algorithms are employed to extract the relevant gain parameters of controller to enhance the potential of the controller. An effort to validate the excellency of T2FPID controller and hDE-HTS algorithm is portrayed clearly through this paper.

Appendix 1: Power System Parameters

K_{ps} = 120 Hz/p.u. MW, T_{PS} = 20 s, B = 0.4249;
R = 2.4 Hz/p.u. MW; T_G = 0.08 s; T_T = 0.3 s;
K_r = 0.33; T_r = 10 s; T_{12} = 0.0707;
a/*b*/*c*/*d* = 900/-18/100/50;

Appendix 2: Assumptions of Algorithms

COF = 10; CDF = 2; RF = 10;

References

1. Kundur, P.: Power System Stability and Control by Prabha Kundur.pdf, p. 1199 (1993)
2. Ibrahim Kumar, P., Kothari, D.P.: Recent philosophies of automatic generation control strategies in power systems. IEEE Trans. Power Syst. **20**(1), 346–357 (2005)
3. Abraham, R.J., Das, D., Patra, A.: Automatic generation control of an interconnected hydrothermal power system considering superconducting magnetic energy storage. Int. J. Electr. Power Energy Syst. **29**(8), 571–579 (2007)
4. Singh, R., Sen, I.: Tuning of PID controller based AGC system using genetic algorithms, vol. 3, pp 531–534 (2004)
5. Dash, P., Saikia, L.C., Sinha, N.: Flower pollination algorithm optimized PI-PD cascade controller in automatic generation control of a multi-area power system. Int. J. Electr. Power Energy Syst. **82**, 19–28 (2016)
6. Sahu, R.K., Panda, S., Rout, U.K.: DE optimized parallel 2-DOF PID controller for load frequency control of power system with governor dead-band nonlinearity. Int. J. Electr. Power Energy Syst. **49**(1), 19–33 (2013)

7. Sinha, N., Saikia, L.C., Rahman, A.: Maiden application of hybrid pattern search-biogeography based optimisation technique in automatic generation control of a multi-area system incorporating interline power flow controller. IET Gener. Transm. Distrib. **10**(7), 1654–1662 (2016)
8. Dash, P., Saikia, L.C., Sinha, N.: Automatic generation control of multi area thermal system using Bat algorithm optimized PD-PID cascade controller. Int. J. Electr. Power Energy Syst. **68**, 364–372 (2015)
9. Alomoush, M.I.: Load frequency control and automatic generation control using fractional-order controllers. Electr. Eng. **91**(6), 357–368 (2010)
10. Arya, Y.: Automatic generation control of two-area electrical power systems via optimal fuzzy classical controller. J. Frankl. Inst. **355**(5), 2662–2688 (2018)
11. Abd-Elazim, S.M., Ali, E.S.: Firefly algorithm-based load frequency controller design of a two area system composing of PV grid and thermal generator. Electr. Eng. **100**(2), 1253–1262 (2018)
12. Nayak, J.R., Shaw, B.: Application of group hunting search optimized cascade PD-fractional order PID controller in interconnected thermal power system. Trends Renew. Energy **4**(3), 22–33 (2018)
13. Indulkar, C.S., Raj, B.: Application of fuzzy controller to automatic generation control. Electr. Mach. Power Syst. **23**(2), 209–220 (1995)
14. Çam, E., Kocaarslan, I.: Load frequency control in two area power systems using fuzzy logic controller. Energy Convers. Manag. **46**(2), 233–243 (2005)
15. Sahu, B.K., Pati, T.K., Nayak, J.R., Panda, S., Kar, S.K.: A novel hybrid LUS-TLBO optimized fuzzy-PID controller for load frequency control of multi-source power system. Int. J. Electr. Power Energy Syst. **74**, 58–69 (2016)
16. Nayak, J.R., Pati, T.K., Sahu, B.K., Kar, S.K.: Fuzzy-PID controller optimized TLBO algorithm on automatic generation control of a two-area interconnected power system. In: IEEE International Conference on Circuits, Power and Computing Technologies ICCPCT 2015, pp 4–7 (2015)
17. Pati, T.K., Nayak, J.R., Sahu, B.K.: Application of TLBO algorithm to study the performance of automatic generation control of a two-area multi-units interconnected power system. In: 2015 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems SPICES 2015 (2015)
18. Nayak, J.R., Shaw, B., Das, S., Sahu, B.K.: Design of MI fuzzy PID controller optimized by Modified Group Hunting Search algorithm for interconnected power system. Microsyst. Technol. **24**, 3615–3621 (2018)
19. Nayak, J.R., Shaw, B., Shahu, B.K.: Load frequency control of hydro-thermal power system using fuzzy PID controller optimized by hybrid DECRPSO algorithm. Int. J. Pure Appl. Math. **114**(9), 147–155 (2017)
20. Biglarbegian, M., Melek, W.W., Mendel, J.M.: Design of novel interval type-2 fuzzy controllers for modular and reconfigurable robots: theory and experiments. IEEE Trans. Ind. Electron. **58**(4), 1371–1384 (2011)
21. Nayak, J.R., Shaw, B., Sahu, B.K.: Application of adaptive-SOS (ASOS) algorithm based interval type-2 fuzzy-PID controller with derivative filter for automatic generation control of an interconnected power system. Eng. Sci. Technol. Int. J. **21**(3), 465–485 (2018)
22. Patel, K., Savsani, V.J.: Heat transfer search (HTS): a novel optimization algorithm. Inf. Sci. (NY) **324**, 217–246 (2015)
23. Storn, R., Price, K.: Differential evolution—a simple and efficient heuristic for global optimization over continuous spaces. J. Global Optim. **11**(4), 341–359 (1997)

Design of Sigma-Delta Converter Using 65 nm CMOS Technology for Nerves Organization in Brain Machine Interface

Anil Kumar Sahu, G. R. Sinha and Sapna Soni

Abstract In this paper, an overview of the present related works in the field of neuroscience is determined. The parts of the neural interface using sigma-delta converter are examined the overall ADC is driven with low voltage to improve the control utilization in the nerves organization. In this work, a basic parts of nervous system is demonstrated which is conquer to solve the problem by delivering power and transmitting data in a minimized manner. For these the above signal is first transmitted to the brain using the help of electrode into the central nervous system of the brain where the signal is diagnosed using brain computer interface by analyzing the data from analog-digital converter. Sigma-delta converter is used for visualizing low frequency signal. Major advantage of this converter is that firstly, clocking circuit need not be design and secondly, it provides good accuracy. Actually, there is no role of digital to analog converter (D/A) connected wirelessly in proposed design. A new topology based on A/D converter, which plays a wide role to minimum supply voltage, and an inactive integrator to decrease control utilization is exhibited, for empowering an in-channel advanced converter plot in a large-scale neural recording implant.

Keywords Large-scale neural recording implants · Low power sigma-delta A/D · Nerves organization · Comparator circuit

A. K. Sahu (✉)
SSGI (FET), SSTC Bhilai, Bhilai, India
e-mail: anilsahu82@gmail.com

G. R. Sinha
IIIT Bangalore, Bangalore, India
e-mail: drgrsinha@ieee.org

S. Soni
SSTC Bhilai, Bhilai, India
e-mail: mesapna93@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_28

1 Introduction

In this work, a basic parts of nervous system is demonstrated which is conquer to solve the problem by delivering power and transmitting data in a minimized manner. For these the above signal is first transmitted to the brain using the help of electrode into the central nervous system of the brain where the signal is diagnosed using brain computer interface by analyzing the data from analog-digital converter. Sigma-delta converter is used in this paper for visualizing low frequency signal; a comparator based sigma-delta converter has designed in order to change over a few input nerves path into an advanced representation with low obstruction. Sigma-delta A/D converter has been successfully designed so that a couple of info nerves way into an upgraded model with low barrier. In this type of converters clocking circuit is avoided that is why accuracy is better to maintained from low frequency signal into high resolution digital value. Right when the need is to restrict control utilization, an in-channel execution one A/D per channel can provoke huge change. The input signal is determined by utilizing an analog to digital converter by considers as an outsourcing the benefit and control of converter part on the outer collector side when joined with radio transmitter, which have a successful part for nerves organization implementation.

A new topology based on simple to advance converter that play a wide role to minimum supply voltage and an inactive integrator to decrease control utilization is exhibited, for empowering an in-channel advanced converter plot in a large-scale neural recording implant.

2 Literature Review

This section identifies and formalizes the problem faced while implementing the references. Noteworthy contributions in the proposed area have been summarized in previous section. However, some significant contributions are reported here as noteworthy contribution.

Sigma-delta converter has been used in channel medium for digitalizing small amplification of brain signal for large-scale implantation of neural implants because dynamic range of sigma-delta modulator is still less in [1]. Dynamic range will be improved depends on design parameter of operational amplifier comparator circuit and switch capacitor used in the design specification. 3-D neural chronicle micro-system has been created in [2] to drives the signify number of neural channels that can be recorded over a trans-conductions for remote connection. Some places front-end processor for multichannel neuronal narrative is depicted in [3]. It gets 12 differential-input channels of embedded account cathodes.

Low power consistent time sigma-delta modulators circuit has been used to configuration subtle elements in [4]. Lithographically portrayed microelectrode clusters currently allow high-thickness recording for recreation in the cerebrum and

are empowering new bits of information into the affiliation limit of the focal sensory system in [5]. In [6] exhibited a coordinated circuit for remote neural incitement, alongside bench top and in vivo trial comes about. The chip can drive 100 individual incitement terminals with consistent current heartbeats of shifting adequacy, span, interphase deferral, and redundancy rate. In [7] presented a multichip structure accumulated with a therapeutic survey treated steel microelectrode group expected for neural records from different channels. Transmission limit diminishing is capable through movement potential disclosure and finish catch of simulation by strategies for on chip data buffering in [8].

The SAR ADC has been utilized to controlled 32 bit channel remote incorporated in system on-a-chip (SoC). The estimated control utilization of each chronicle channel; consume less power in [9]. Ultra-low-power 32-channel neural-recording coordinated circuit chip comprises of eight neural account modules in [10] where every module contains neural recorder unit, multiplexer(MUX) unit, an A/D converter, and a serial programming interface unit. Introduced a third request incremental sigma-delta converter (ADC) for time-multiplexed signals. Incremental sigma-delta modulation has been used to convey medium to high assurance necessities of multichannel applications, while a third demand steady time execution be inspected as a possibility for low-control plans in [11]. Inside neural interface IC for neural activity dynamic figure of A/D is key factor for electronic control modules and power organization circuits in [12]. The sigma-delta is used to control the voltage controlled oscillator circuit used in neural prosthetic [13]. Different sigma-delta modulator has been proposed to diminish control utilization and size in implantable bio-interfacing frameworks in [14].

Bidirectional neural interfacing applications with filter information rates: higher rates are required for recording of uplink signals compare to downlink signals in [15]. In [16] Exhibited a mind development in vivo requires gathering bioelectrical signals from a couple of microelectrodes in the meantime in demand to get neuron associations. Brain machine interface setup for neural implants is shown in Fig. 1.

The fundamental problem are examined and reported in following:

- The basic neural-recording implants using sigma-delta converter is designed for storage purpose, in this design the area of the chip is less but the power dissipation and delay is quite high.
- In order to reduce the area, power and delay of the chip, the basic chip is designed using sigma-delta converter with different technologies and with using switch capacitor.
- The four bit sigma-delta converter has been designed and their performance is compared, in result we observed that only power is reduced, there is not much affect in area.
- We also noticed that in designing of many CMOS circuits the use of switch capacitor is very important, which plays a major impact on designing the chip by achieving a great resolution.

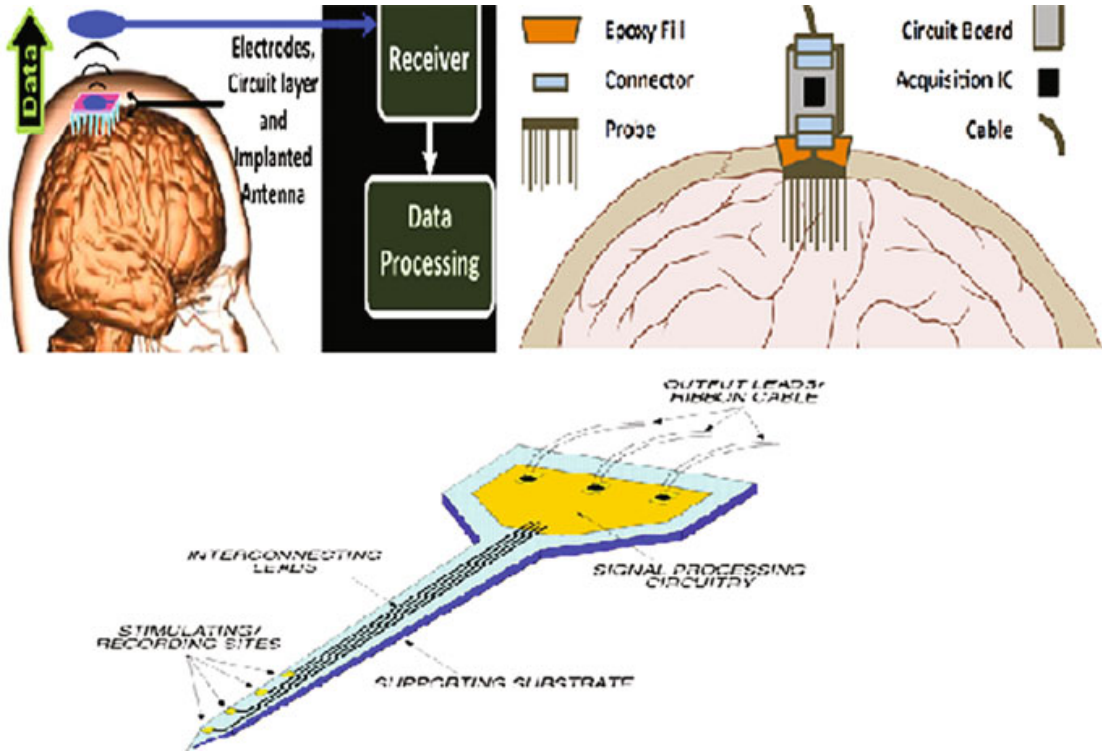


Fig. 1 Brain machine interface setup for neural implants

3 Methodology

In this part, we will examine about first neuron cell, neural action, and neural chronicle interface, second stage portray the execution of four info comparator circuit for neural account and the proposed circuit usage of the sigma-delta converter. In this postulation, proposed approach is outlined and recreated by utilizing Tanner EDA at CMOS BSIM4 innovation. To start with area clarifies the neural chronicle inserts, next segment clarifies the sigma-delta converter and last segment clarifies the proposed technique utilizing BSIM4 innovation.

The objectives of this proposed work are

(1) To design and characterize sigma-delta modulator based on comparator circuit using Quad input.
(2) To improved performance of sigma-delta modulator data converter.

Figure 2 shows the design flow chart of sigma-delta converter.

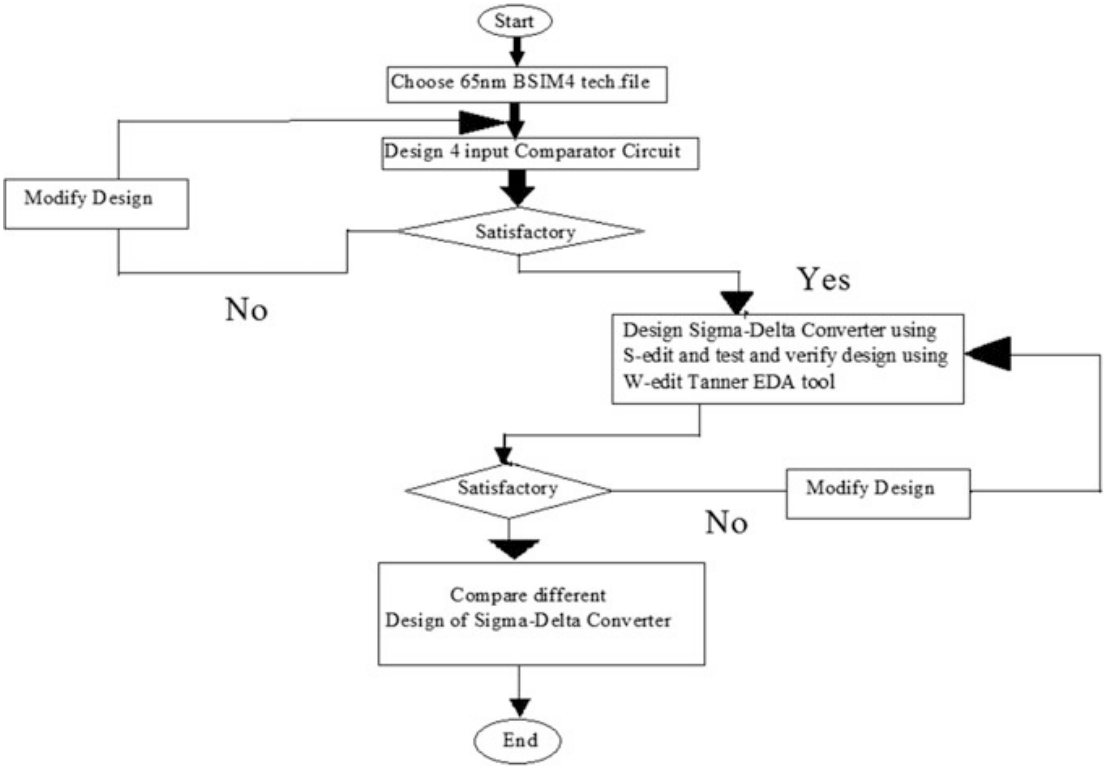


Fig. 2 Flow chart of sigma-delta converter

3.1 Design of Comparator

From Fig. 3 it can be seen that Comparator is design using 12 transistor such as M1, M2, M3, M4, M5, M6, M7, M8, M9, M10, M11, M12 and Two Inverter connected as output. It is simply differential amplifier with two input fw- and fw+. Depend on clack signal $\varphi 1$ Comparator will compare the input and produce output.

From Fig. 4, it can seen that comparator circuit using four input is implemented in tanner using 65 nm technology. Figure 5 shown top view of sigma-delta converter using four-input comparator circuit used in Nerves Organization system shown in Fig. 6. It can be seen that the nerves signal read by the terminal hold little a motivation in solitude. Warm noise from the cathode and the tissue interface and natural commotion from the neighboring neurons are accessible. These signs ought to be adjusted and arranged before the information they contain can be considered. This is where the neural interface accepts a critical part (Figs. 7 and 8).

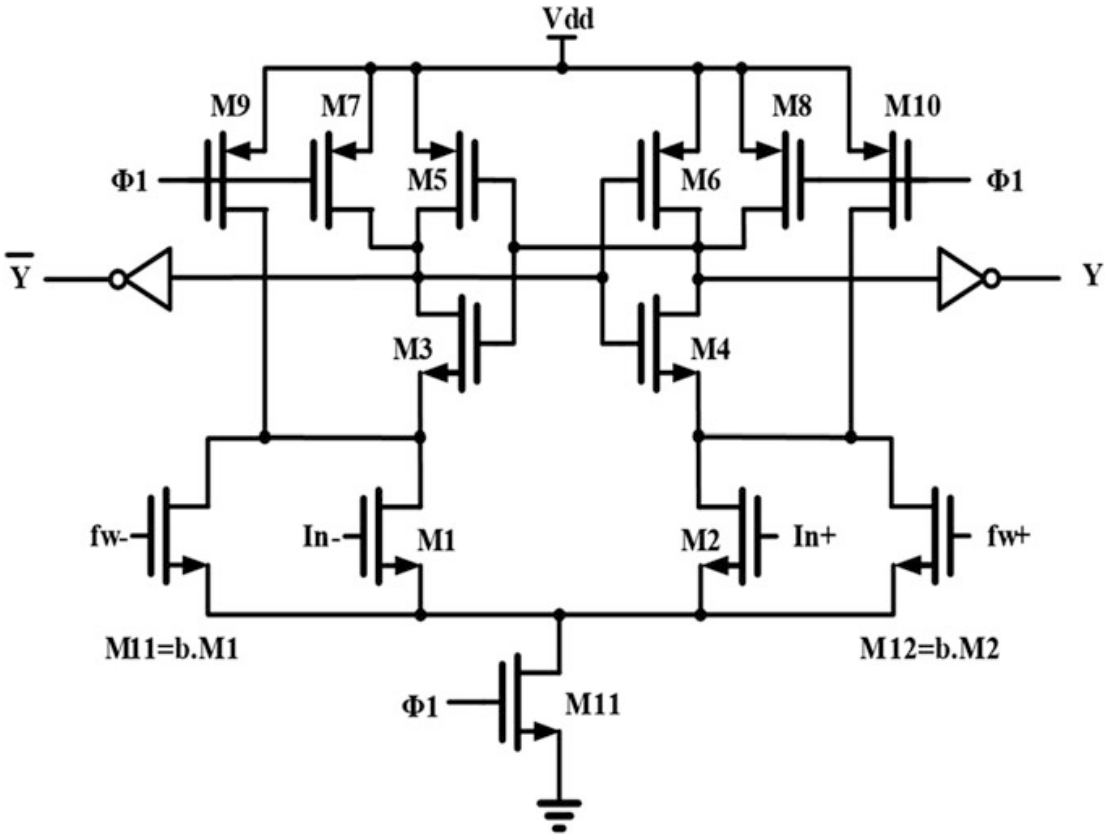


Fig. 3 Schematic of comparator circuit using four input used in sigma-delta ADC

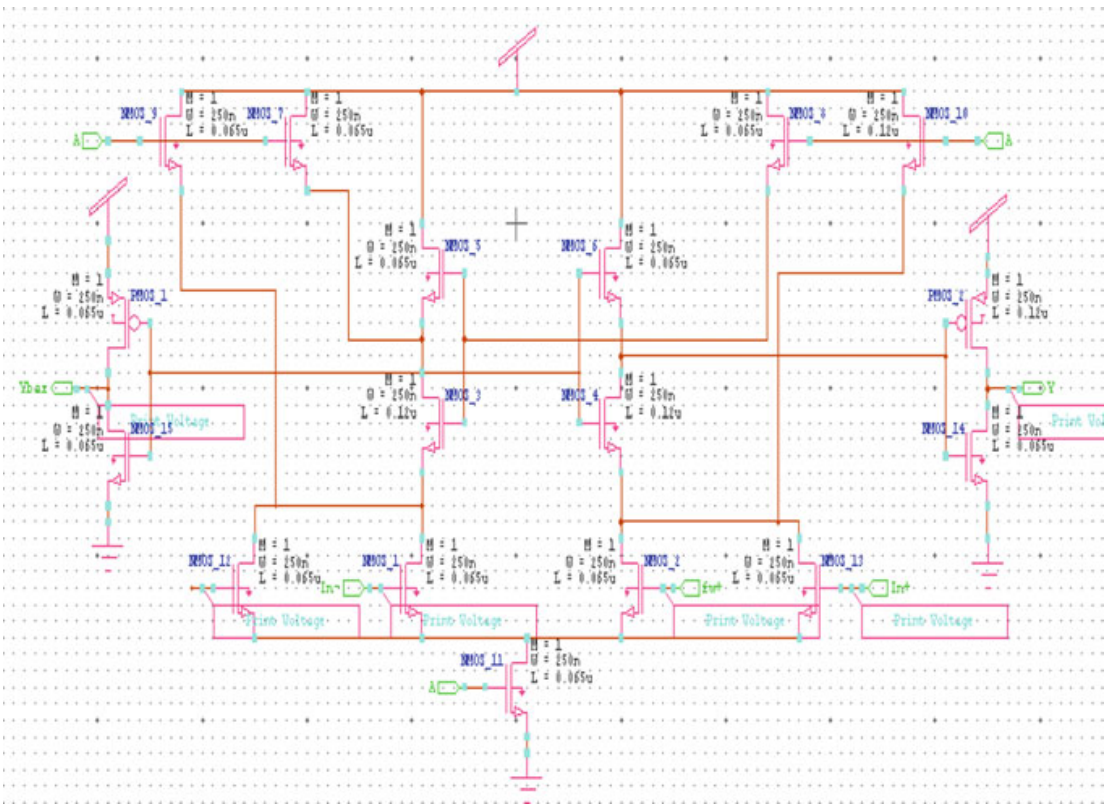


Fig. 4 Implementation of comparator circuit using quad input sigma-delta A/D

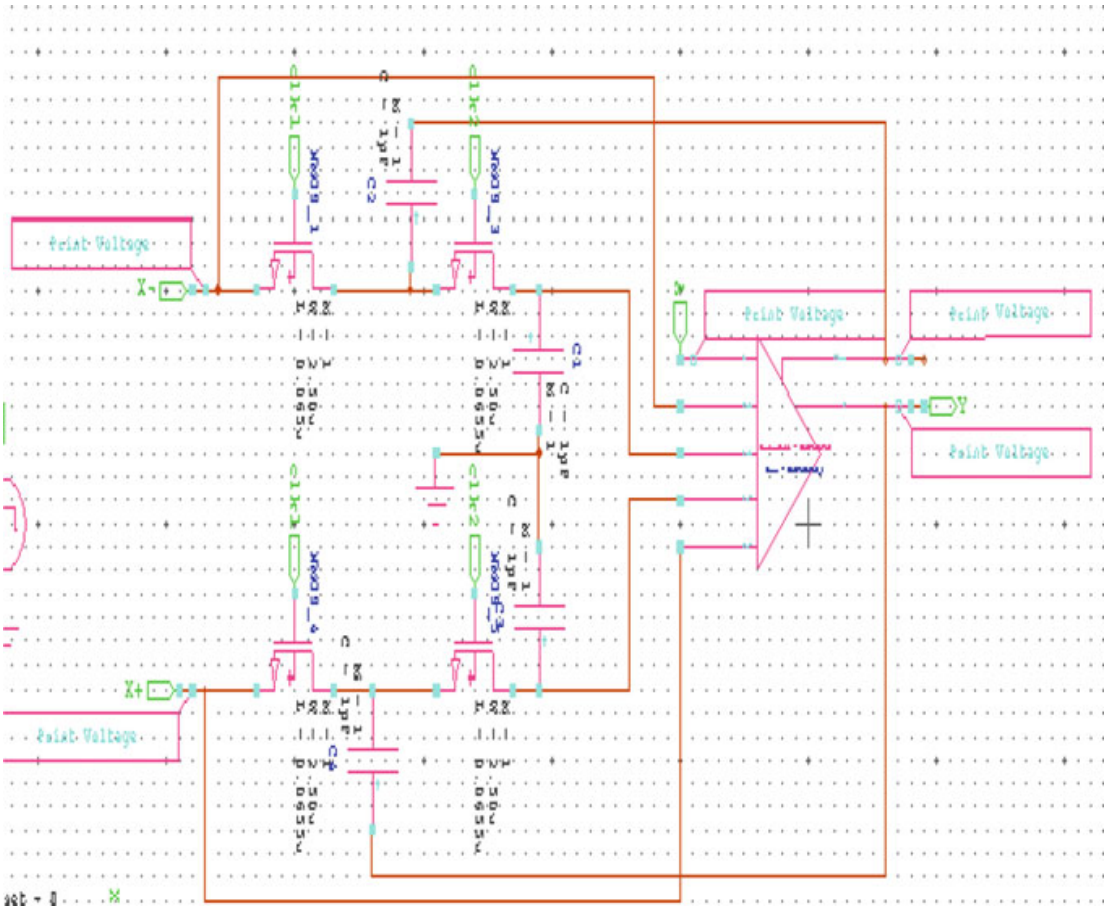


Fig. 5 Top view of sigma-delta converter using quad input comparator circuit

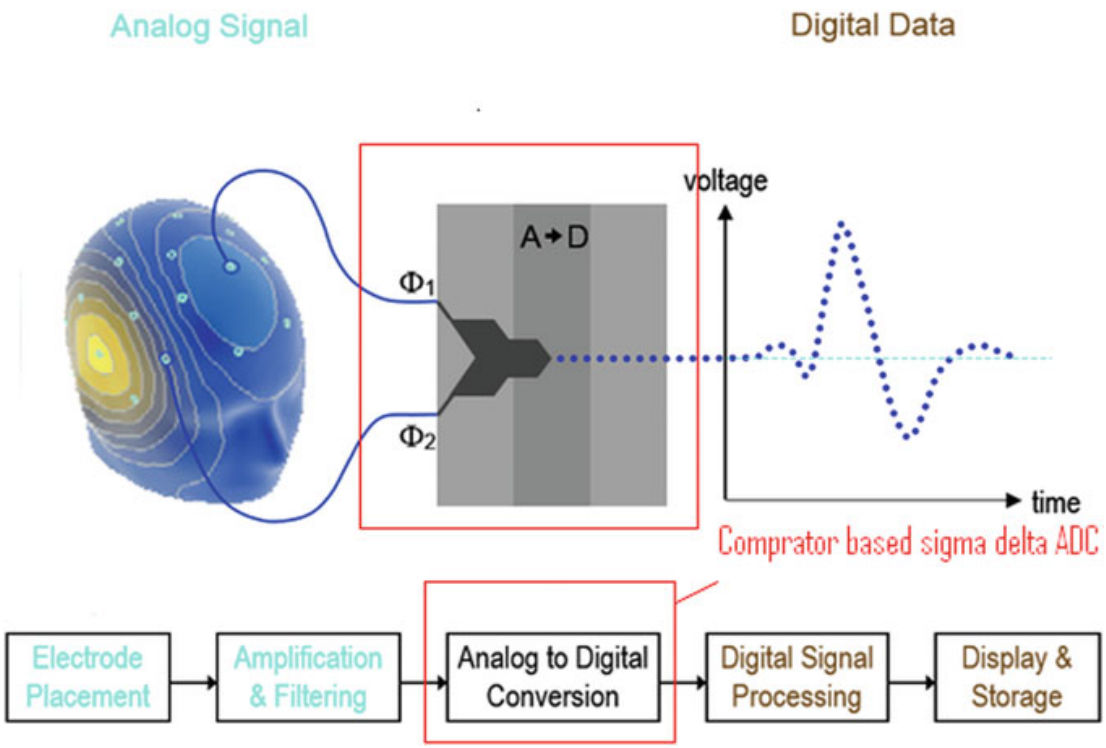


Fig. 6 Schematic view of nerves organization implementation

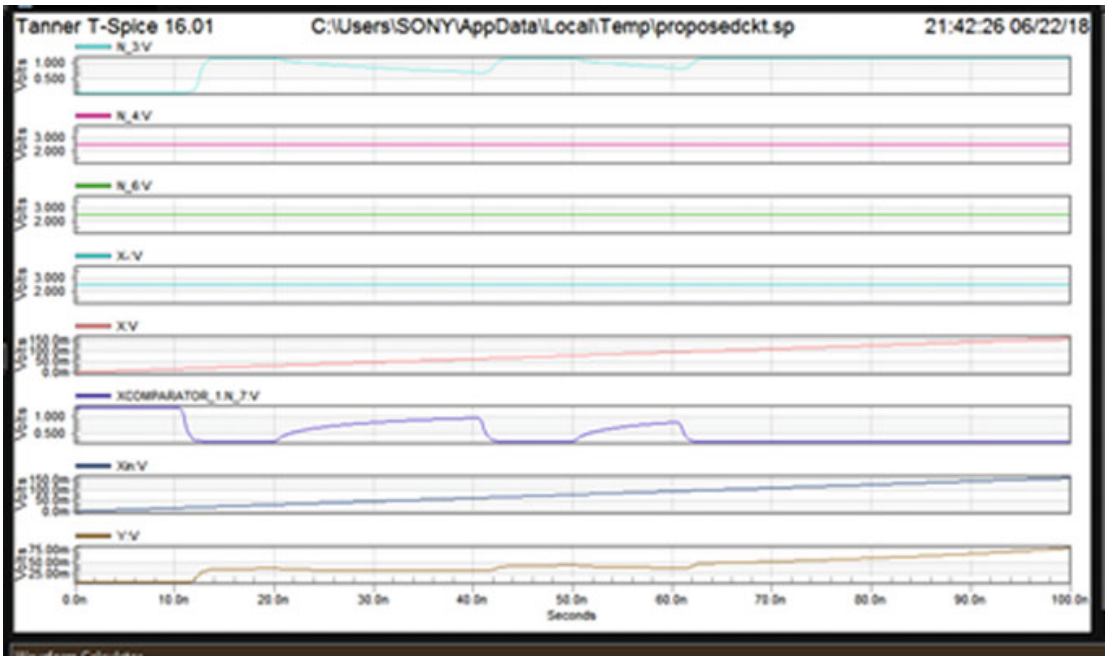


Fig. 7 Simulation result of comparator circuit using four input

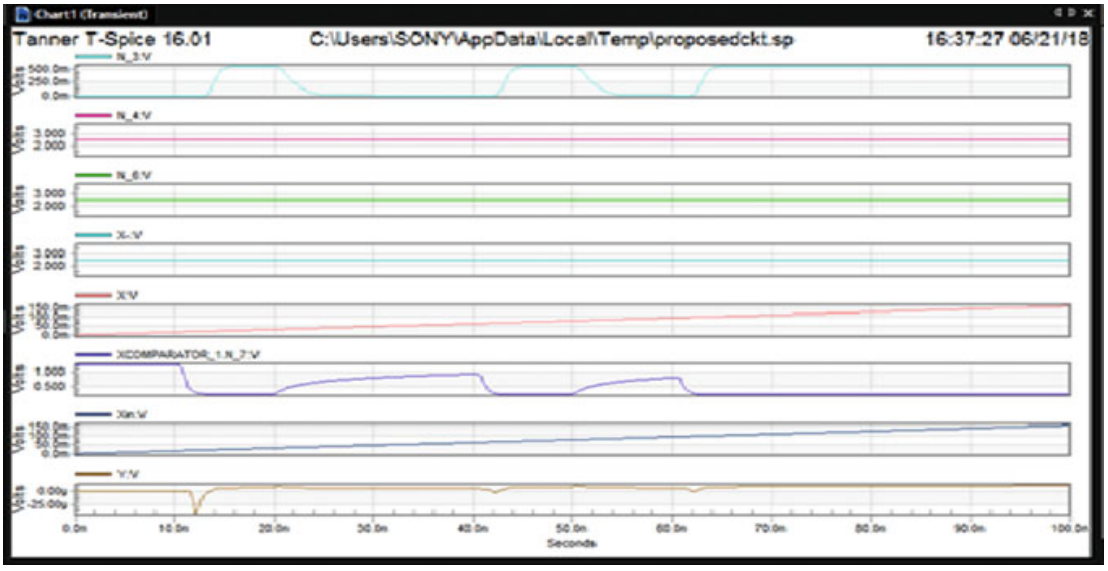


Fig. 8 Simulation results of sigma-delta converter

Table 1 Performance comparisons

Performance parameter	Ref. [3]	Ref. [1]	This work
Technology (nm)	180	180	65
Topology	DT-∑Δ	DT-∑Δ	DT-∑Δ
BW (KHz)	5	10	10
SNDR (dB)	105	55.2	50.2
ENOB (bits)	17.15	8.88	7.58
OSR	–	50	–
Power (μW)	280	0.11	0.4
FOM	190	11	45

4 Results and Discussion

The proposed circuit has been made in a TSMC 65-nm CMOS circuit. In this paper the input in the above circuit is feeded into the central nervous system where the flexible power coil is implanted using implantable flexible board in which the array of electrode is fully distributed inside the cerebellum by the help of four-input comparator circuit and the output driven signal of the above comparator circuit is used to drive the sigma-delta converter which holds a great impact of driving low frequency signal with low power consumption is determined in this paper, The power utilizes in the above circuit is above 0.4 μW that include only two non-covering clock generator. In this paper, the circuit utilizes real time user interface, which holds a great impact of determining damage nerves inside the brain to cure various chronic disease, brain tumor, various trauma inside the cerebellum (Table 1).

5 Conclusion

In this paper, we presented outline of a novel sigma-delta converter that can be used for high-thickness neural record and biotelemetry applications. A FOM has been relate and figured for a couple of determined diagrams, and the implemented circuit poses great impact with relative data converters. The over all circuit inside a multichannel recording neural testing activity is implemented successfully.

Analyzing of data on recorded. To fulfill the predetermined unique fluctuate, in any case utilization of low consumption of control power. Here the style investigations, circuit execution and stream of digital signal determination whole power utilization of the modulator is 0.4 μW that compares to a FOM of 45 fJ/change step. The focused circuits details make this style a good possibility for building high accuracy neurosensory. Highlights drawback is to control the input relate to sigma-delta converter using comparator circuit as shown above. The proposed arrangement includes speed frequency range of 1 kHz of 10 kHz, and gives 7.58

bits value of precise division by using another feed forward topology, while including minimum number of additional portions and switches. A four-channel straightforward front-end has been executed in this proposed structure and made in a CMOS 65-nm innovation.

The major limitation of our proposed methodology is to operate switch capacitor and overall circuit design is quite complex and increases the chip size. Here the overall power delay of the entire circuit may be vary when operated with different technology. In future, this sigma-delta A/D design approach using comparator circuit can be designed for n number of bits, which results in low power and lower delay. It is beneficial for larger circuits, in order to reduce the delay.

References

1. Razaei M., Maghsoudlo E., Sawan M., Gosselin B.: A 110-nW in channel sigma-delta converter for large-scale neural recording implants. in: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), April, pp. 5741–5744 (2016)
2. Roy, H.O.I.I.I., Kensall, D.W.: A three-dimensional neural recording microsystem with implantable data compression circuitry. IEEE J. Solid-State Circuits **40**(12), 2796–2804 (2005)
3. Perelman, Y., Ginosar, R.: An integrated system for multichannel neuronal recording with spike/LFP separation, integrated A/D conversion and threshold detection. IEEE Trans. Biomed. Eng. **54**(1), 130–137 (2007)
4. Pavan, S.: Excess loop delay compensation in continuous-time delta-sigma modulators. IEEE Trans. Circuits Syst. II Express Briefs **55**(11), 1119–1123 (2008)
5. Wise, K.D., Sodagar, A.M., Yao, Y., Gulari, M.N., Perlin, G.E., Najafi, K.: Microelectrodes, microelectronics, and implantable neural microsystems. IEEE J. Mag. **96**(7), 1184–1202 (2008)
6. Thurgood, B.K., Warren, D.J., Ledbetter, N.M., Clark, G.A., Harrison, R.R.: A wireless integrated circuit for 100-channel charge-balanced neural stimulation. IEEE Trans. Biomed. Circuits Syst. **3**(6), 405–414 (2009)
7. Gosselin, B., Ayoub, A.E., Roy, J.F., Sawan, M., Lepore, F., Chaudhuri, A., Guitton, D.: A mixed-signal multichip neural recording interface with bandwidth reduction. IEEE Trans. Biomed. Circuits Syst. **3**(3), 129–141 (2009)
8. Lee, S.B., Lee, H.M., Kiani, M., Jow, U.M., Ghovanloo, M.: An inductively powered scalable 32-channel wireless neural recording system-on-a-chip for neuroscience applications. IEEE Trans. Biomed. Circuits Syst. **4**(6), 360–371 (2010)
9. Shahrokhi, F., Abdelhalim, K., Serletis, D., Carlen, P.L., Genov, R.: The 128-channel fully differential digital integrated neural recording and stimulation interface. IEEE Trans. Biomed. Circuits Syst. **4**(3), 149–161 (2010)
10. Wattanapanitch, W., Sarpeshkar, R.: A low-power 32-channel digitally programmable neural recording integrated circuit. IEEE Trans. Biomed. Circuits Syst. **5**(6), 592–602 (2011)
11. Garcia, J., Rodriguez, S., Rusu, A.: A low-power CT incremental 3rd order $\sum$-Δ ADC for biosensor applications. IEEE Trans. Circuits Syst. **60**(1), 25–36 (2012)
12. Yin, M., Borton, D.A., Aceros, J., Patterson, W.R., Nurmikko, A.V.: A 100-channel hermetically sealed implantable device for chronic wireless neurosensing applications. IEEE Trans. Biomed. Circuits Syst. **7**(2), 115–128 (2013)

13. Song, H., Chen, C., Lin, M.W., Li, K., Christen, J.B.: A neural rehabilitation chip with neural recording, peak detection, spike rate counter, and biphasic neural stimulator, pp. 415–419 (2014)
14. Rezaei, M., Maghsoudloo, E., Sawan, M., Gosselin, B.: A novel multichannel analog-to-time converter based on a multiplexed sigma delta converter. In: 2015 IEEE 13th International New Circuits and Systems Conference (NEWCAS), Grenoble, pp. 1–4 (2015)
15. Mirbozorgi, S.A., Bahrami, H., Sawan, M., Rusch, L.A., Gosselin, B.: A single-chip full-duplex high speed transceiver for multi-site stimulating and recording neural implants. IEEE Trans. Biomed. Circuits Syst. **10**(3), 643–653 (2016)
16. Rezaei, M., Maghsoudloo, E., Bories, C., Koninck, Y., Gosselin, B.: A low-power current-reuse analog front-end for high-density neural recording implants. IEEE Trans. Biomed. Circuits Syst. **12**, 1–10 (2018)

Performance Comparison of Machine Learning Techniques for Epilepsy Classification and Detection in EEG Signal

Rekh Ram Janghel, Archana Verma and Yogesh Kumar Rathore

Abstract Epilepsy is a neurological affliction that in impact around 1% of humankind. Around 10% of the United States populace involvement with minimum a solitary convulsion in their activity. Epilepsy has recognized respectively tendency of the cerebrum outcomes unforeseen blasts of weird electrical action which disturbs the typical working of the mind. Since spasms by and large happen once in a while and are unforeseeable, seizure identification frameworks are proposed for seizure discovery amid long haul electroencephalography (EEG). In this exploration, we utilize DWT for highlight extraction and do correlation for all kind of Machine learning order like SVM, Nearest Neighbor Classifiers, Logistic relapse, Ensemble classifiers and so on. In this examination classification accuracy of Fine Gaussian SVM recorded as 100% and it has better as compare to other existing machine learning approaches.

Keywords Decision tree · Ensemble classifier · Electroencephalogram (EEG) · Epileptic seizure · *k*-Nearest neighbor · Support vector machine

1 Introduction

Epilepsy problem is a bunch of typical neurological conditions of the mind distinguished by recurrent unprovoked convulsions which are consequence of sudden bursts of abnormal electrical discharges in the brain [1–3]. According to a report by WHO [4] around 50 million people are in effected by epilepsy worldwide [5–8].

R. R. Janghel (✉) · A. Verma · Y. K. Rathore
Department of Information Technology, National Institute of Technology, Raipur, India
e-mail: rrjanghel.it@nitrr.ac.in

A. Verma
e-mail: archana1694@gmail.com

Y. K. Rathore
e-mail: yogeshrathore23@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_29

Nearly one in every 100 persons effected by a convulsion in their lifespan [8], generally 2.4 million new instances of epilepsy accounted for consistently internationally [6, 9]. Until now, the manifestation of an epileptic is eccentric and its course of process has been less understood [10], Epileptic patients experience varying symptoms depending on the section and the expanse of the brain affected, Epileptic seizures can induce dull physical, social consequences and psychological, which include loss of consciousness, injury and abrupt death. Epilepsy is of two samples which are depending upon the degree of involvement of the brain tissue, which is, generalized seizures and, partial involves closely to the complete brain, partial seizures create in a specific area of the mind and stay constrained to that segment [8]. Electroencephalogram (EEG) distinguished by the electrophysiological technique to appreciate the problematic activity of the human brain [11]. EEG exactly gauges and registers the electrical action of the mind. Natural EEG signals have classifieds into many rhythms based on their frequencies, which are a band (3–4 Hz); band (4–8 Hz); band (8–13 Hz); band (13–30 Hz) [12, 13]. An EEG is particularly helpful on occasion when the cerebrum is in danger by giving a delicate sign of cerebral working. Such interim is as a rule of long time ranges, thus a broadened EEG recording required early investigations have demonstrated proof of this unusual movement to be an advantageous guide in the identification of epilepsy and cerebral tumors. These days EEG signals are utilized to get data significant to the determination, visualization, and cure of these anomalous case. EEG can be register using electrodes arranged on the head and have tiny amplitudes in the range of 20 V [14, 15, 16]. The anodes are put according to the 10–20 international system which is shown in the Fig. 1.

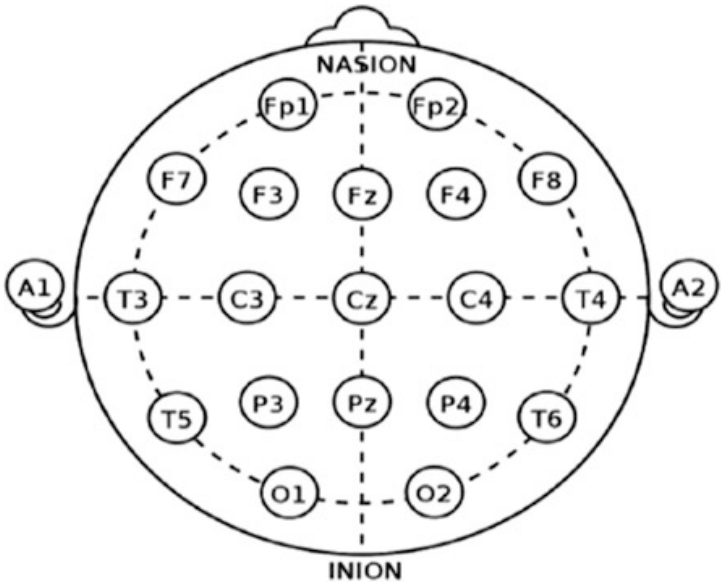


Fig. 1 Standardized electrode placement scheme

2 Proposed Work

SVM algorithm is used to a classification of healthy and seizures in 17 full term dataset of newborns baby with seizure by the author [17]. Using PCA with help of decomposing WPD getting an accuracy of 99% with a less computational cost for extracting features on Bonn University data set [18]. Using Novel wavelet-base and SVM for repeated seizure detection and getting sensitivity is 94.46% and specificity is 95.26% [19]. Using the Expectation–Maximization algorithm used for learning the Mixture of experts network structure getting an accuracy of 93.17% for classification seizure and non-seizure [20]. Fuzzy c-mean algorithm and MLP sub-network are used for ECG beat classification in this MIT/BIH dataset are used and enhance their performance [21, 22].

2.1 Dataset

The data elucidated by Andrzejak et al. [8] was employed for the current research. The total data set is comprised of five subsets, (denoted as Z, O, N, F, S) every each subset having 100 single-channel EEG fragments each being on 23:6 s time span sampled at 173:6 Hz. The fragments were chosen and hand-picked from regular continue multi-channel EEG recordings after ocular examination for artifacts for e.g., eye movements or muscle action etc. shown in Fig. 2.

2.2 Discrete Wavelet Transforms

DWT is best for DE noising and it think about signs and pictures as it helps indicates numerous regular happening signs and pictures with the assistance of fewer coefficients this empower a sparser introduction base scale in dwt is organized by you can get diverse scales by raising.

$$2^j, \quad \text{where } (j = 1, 2, 3, 4 \ldots)$$

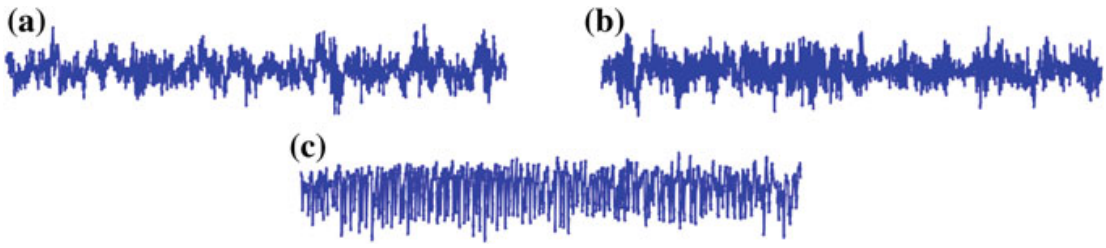


Fig. 2 Raw EEG data belonging to each of subsets **a** Z, **b** O and **c** S

Base scale to integer value shows in below

$$2^j m, \quad \text{where } (m = 1, 2, 3, 4\ldots)$$

The interpretation comes at whole numbers numerous speaks to in this condition this procedure is frequently alluded to as power scaling and moving this sort of inspecting wipes out repetition in coefficients. The output of the change yields the same number of coefficients from the length of the discrete wavelet transform process is proportional to contrasting a signal with discrete multi-rate filter banks conceptually.

The discrete wavelet change is a flexible flag handling device that searches many building and logical functions. One zone in which the DWT has been especially fruitful is the epileptic seizure identification since it catches transient highlights and restricts them in both time and recurrence content precisely. Different data sets are shown in Fig. 3.

After applying Discrete Wavelet transform we apply multiple methods have been put forward for epileptic seizure detection and compare their accuracy with help of confusion matrix. Dataset divided in the 10-k fold for training and testing. Multiple classifications we use for comparing data accuracy of various algorithms.

2.3 Different Machine Learning Techniques

A Detailed flowchart of the proposed approach is shown in Fig. 4.

2.3.1 Support Vector Machine

SVM is a supervised learning Model. It associated learning algorithms analyzing data used for classification and also regression analysis [23]. In this paper we use 2-class. Linear SVM, Quadratic SVM, Cubic SVM, Fine Gaussian SVM, Medium Gaussian SVM, and Coarse Gaussian SVM is prediction speed is fast in binary input and easy interpretability. It uses medium memory usage shown in Fig. 4.

2.3.2 KNN

The *k*-closest neighbor's calculating (*k*-NN) is a non-parametric approach utilized for characterization and relapse [24]. In the two cases, the data comprises of the *k* nearest preparing cases in the element space shown in Fig. 5. In this paper, we use Fine KNN, Cubic KNN, Coarse KNN, Weighted KNN, Medium KNN, Cosine KNN and calculate accuracy and compare each other (Fig. 6).

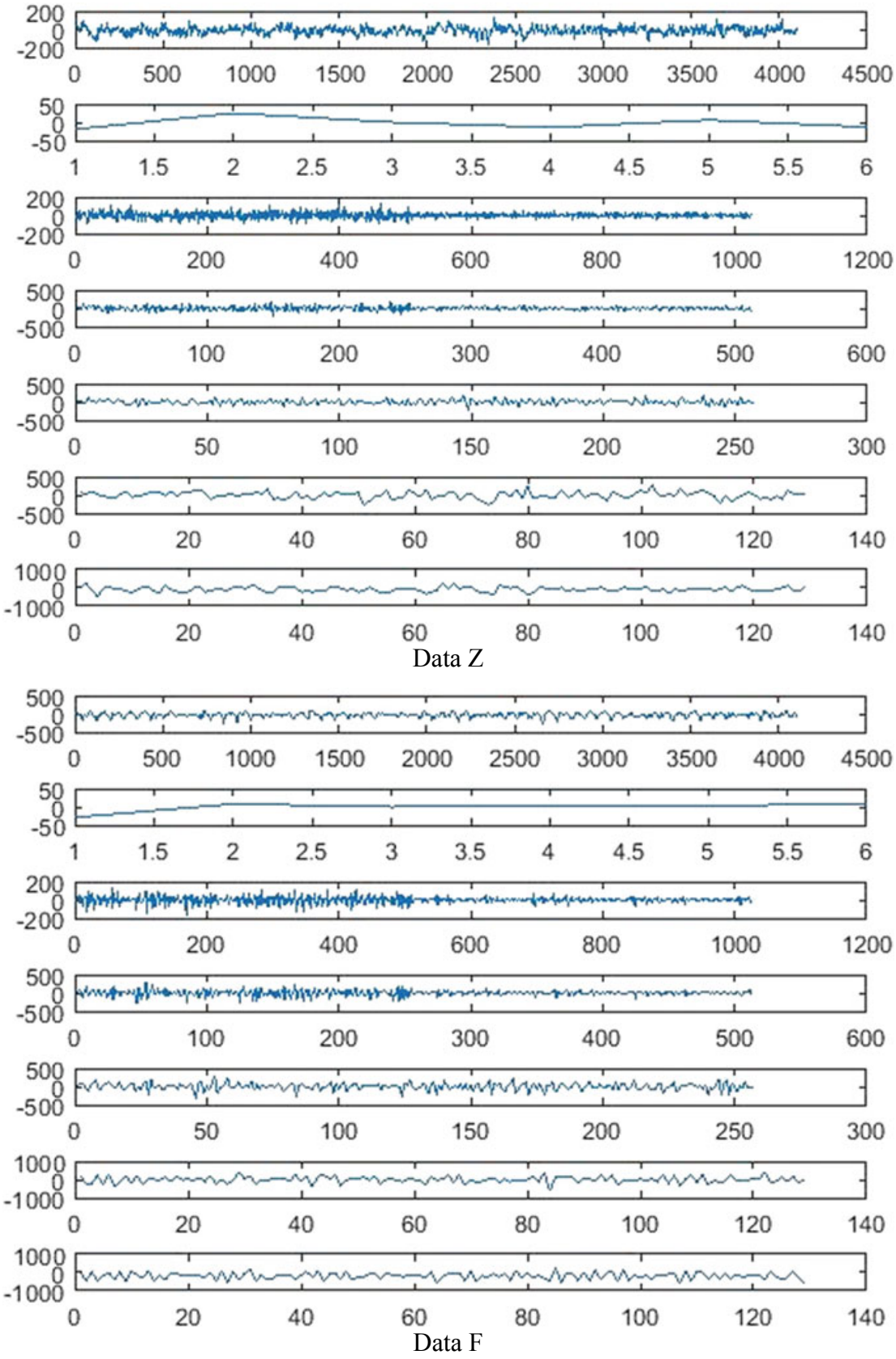


Fig. 3 Decompose dataset by using DWT

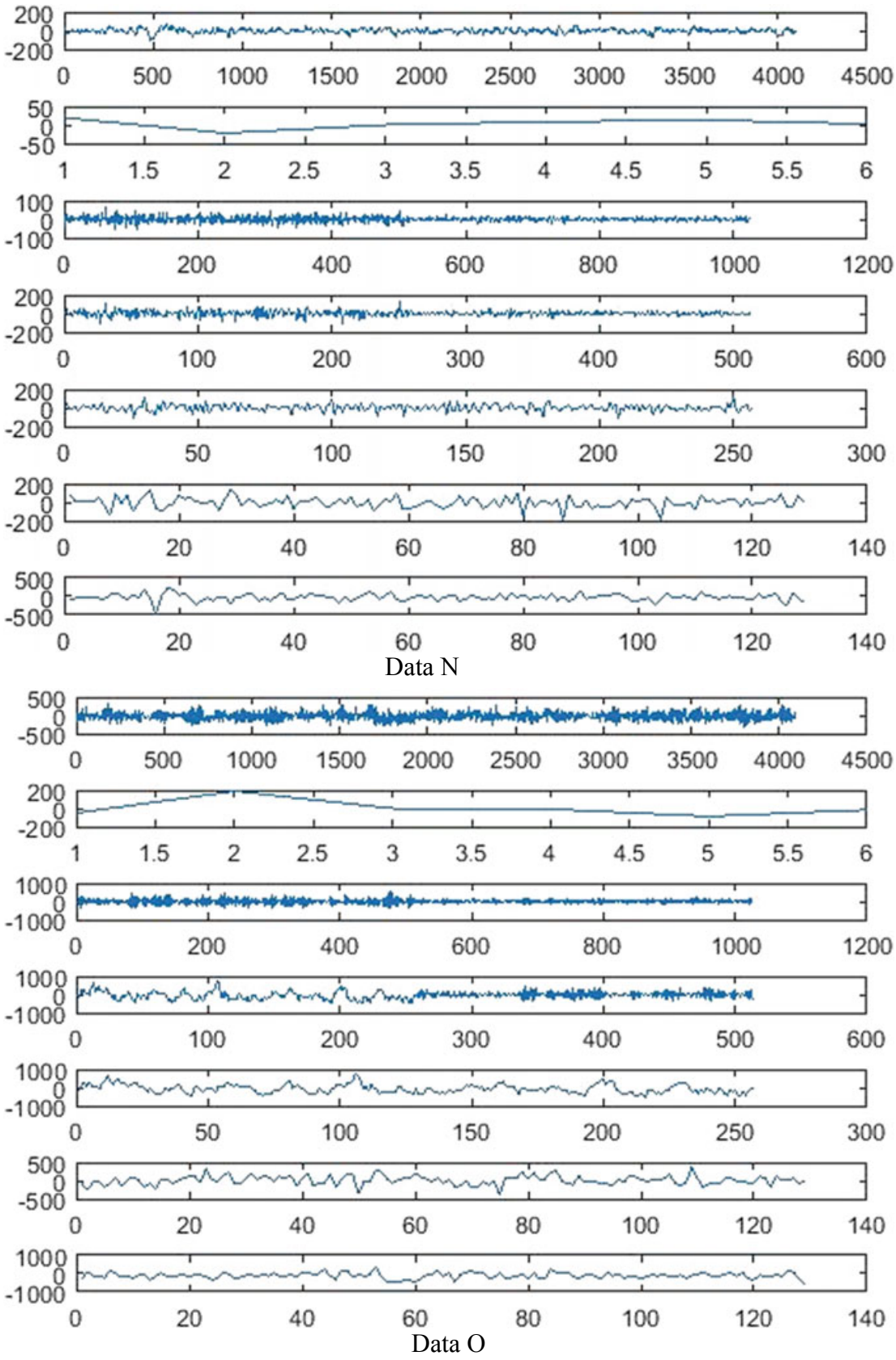


Fig. 3 (continued)

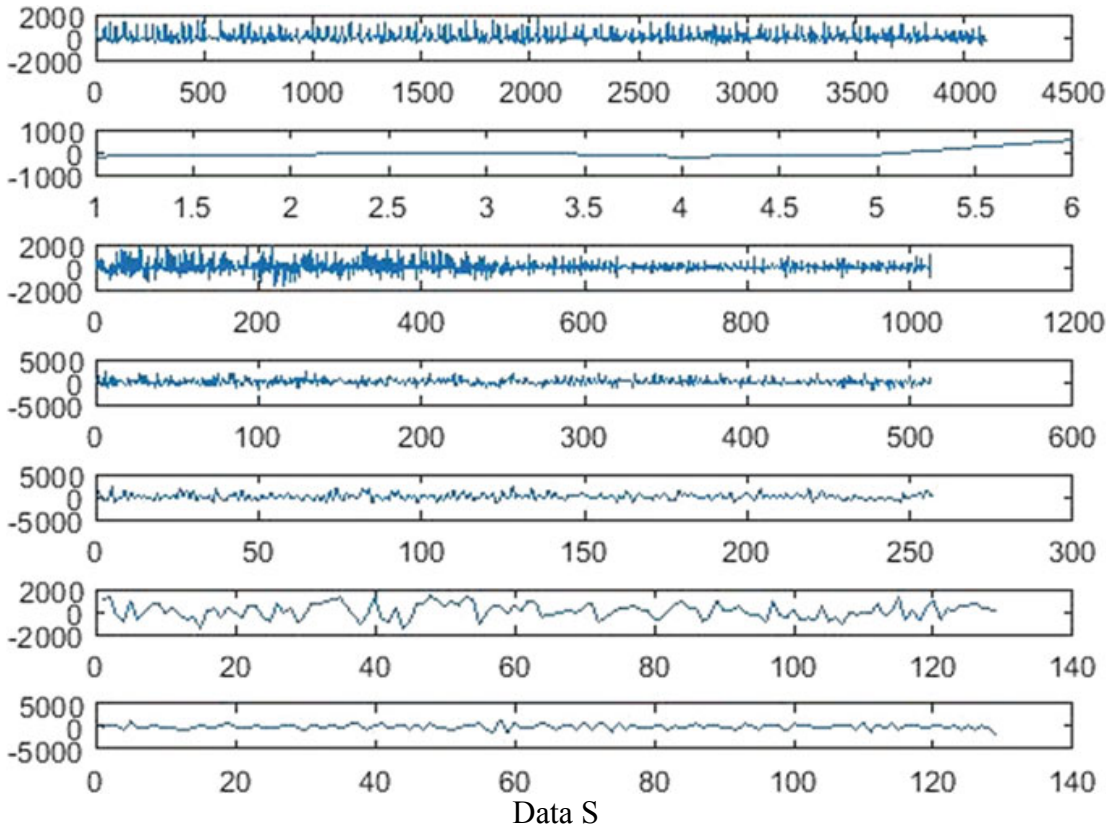


Fig. 3 (continued)

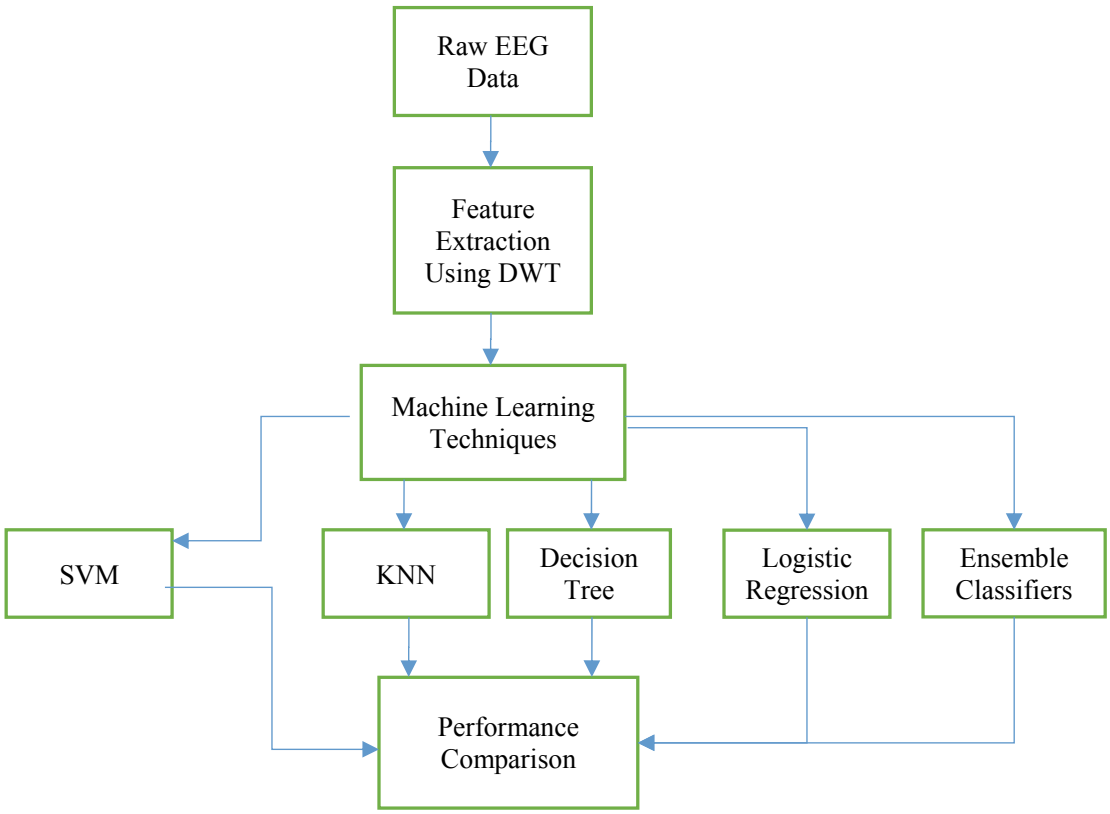


Fig. 4 Flowchart of Proposed Approach

Fig. 5 SVM Classifier [6]

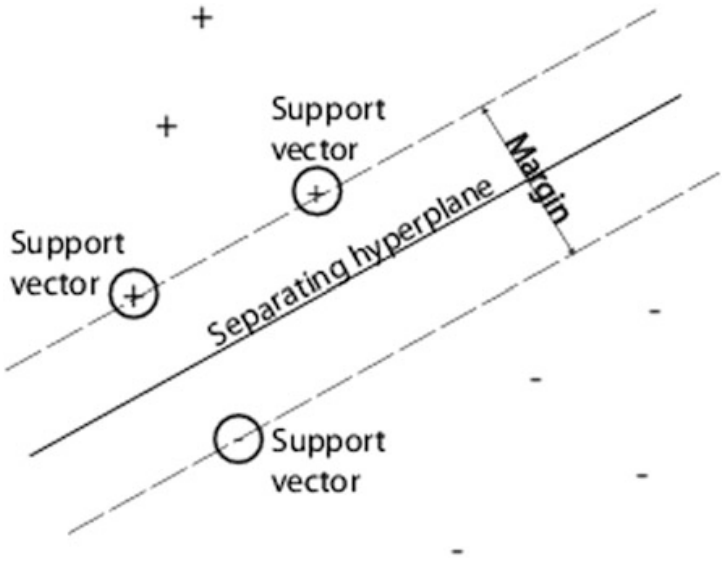
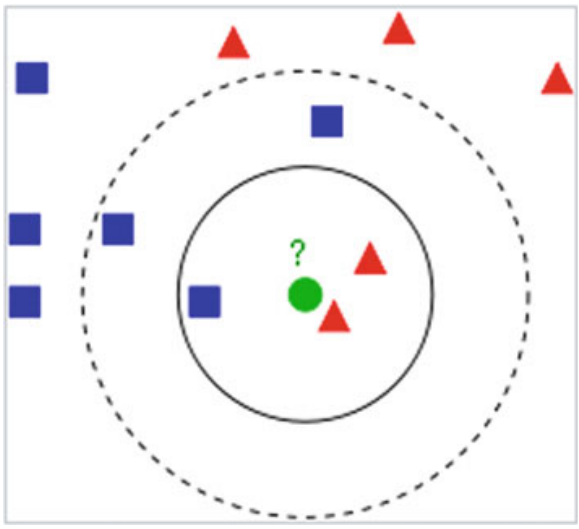


Fig. 6 KNN Classifier [10]



2.3.3 Logistic Regression

Logistic regression is a generally utilized factual displaying procedure in which the probability, $P1$, of dichotomist result occasion is identified with an arrangement of illustrative factors in the shape

$$\begin{aligned} \text{logit}(P1) &= \text{In}\left(\frac{p1}{1-p1}\right)\beta_0 + \beta_1 X_1 + \cdots + \beta_n X_n \\ &= \beta_0 + \sum_{i=0}^{n} \beta_i X_i \end{aligned} \tag{1}$$

In Eq. (1), β_0 is the intercept and $\beta_1, \beta_2, \ldots, \beta_n$ are the coefficients combine with the analytical variable $X_1, X_2, \ldots, X_n$. These info factors are the norm of the wavelet coefficients (D3—D5 and A5) of four-channel EEG signals [25]. A dichotomous variable is confined to two qualities, for example, yes/no, on/off, survive/bite the dust or 1/0, for the most part speaking to the event or non-event of some occasion (for instance, epileptic seizure/not). The illustrative (free) factors might be consistent, dichotomous, and discrete or blend. The utilization of common straight relapse (OLR) in light of slightest squares strategy with a dichotomous result.

2.3.4 Decision Tree

Decision tree learning is a standout amongst the usually broadly utilized and practical strategies for inductive inference. It is a technique for estimation of discrete-valued function that is hearty to uproarious information and fit for learning disjunctive expression [26].

Decision tree study is a route for looking like discrete-valued function, in which the prepared function is spoken to by a decision tree. Decision trees can likewise be spoken to an asset or it at that point guidelines to enhance human intelligibility. Decision tree study is an analytic, one-advance look forward (slope climbing), non-backtracking seek through the space of all conceivable Decision trees.

2.3.5 Ensemble Classifier

It is used for multiple learning algorithms to enhance better predictive performance. It can do increments with the number of training or splitting model flexibility. It's prediction better and memory usage is average. Here, we work on Boosted Tree, Bagged Tree, Subspace Discriminant, Subspace KNN, and RUSBoost tree ensemble classifiers.

3 Experimental Results and Analysis

From the Table 1, we can see that italic value shows the highest performance of a classifier whereas bold value shows the best classification result for a particular feature. On below graphs (Figs. 7, 8, 9, 10) and tables (Table 1 and 2) comparison of classifiers on the basis of accuracy is shown.

Table 1 Results of the implementation

Z versus F						Z versus O					Z versus N					Z versus S				
Feature/method	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5	D1	D2	D3	D4	D5
Cross tree	56.7	62.9	50.6	53.9	77.5	56.2	67.4	68.0	71.9	*93.3*	65.7	66.3	58.4	53.4	61.2	87.6	82.9	80.9	84.8	88.2
Fine tree	59.0	67.4	63.5	60.7	83.1	57.9	68.5	68.0	70.2	*94.4*	69.7	66.3	61.8	52.8	65.2	89.9	85.4	87.1	87.1	91.6
Medium tree	59.6	67.4	56.7	60.7	83.1	56.2	68.5	68.0	70.2	*94.4*	69.1	66.3	61.8	52.8	65.2	89.9	85.4	87.1	87.1	91.6
Linear SVM	54.5	69.7	71.9	70.2	69.1	51.1	57.3	*78.7*	56.7	69.1	55.1	68.0	78.1	77.5	*78.7*	60.7	57.3	55.6	58.4	62.4
Quadratic SVM	74.2	68.5	74.7	77.0	85.4	64.0	61.8	79.8	61.8	84.4	**75.8**	62.4	77.5	**82.6**	83.7	*87.1*	62.9	62.9	63.5	75.8
Cubic SVM	71.3	65.2	78.1	74.7	83.1	**68.0**	62.9	78.7	64.0	82.6	69.7	57.3	**78.7**	82.0	86.5	*89.9*	63.5	64.6	66.9	79.2
Fine Gaussian SVM	**74.7**	70.2	69.7	68.5	91.6	65.2	64.6	65.7	62.4	*100*	69.1	65.7	64.0	65.2	64.0	**95.5**	**99.4**	**99.4**	**99.4**	*100*
Medium Gaussian SVM	68.5	74.2	**78.7**	**81.5**	89.9	**68.0**	**87.1**	83.1	**84.3**	93.3	74.2	**77.5**	83.1	80.9	**89.3**	86.5	*93.8*	89.9	90.4	*93.8*
Coarse Gaussian SVM	55.6	55.1	56.2	56.7	68.5	57.3	57.3	66.7	56.7	68.5	56.7	59.6	62.9	61.8	66.9	56.7	63.5	64.6	68.5	*70.2*
Logistic regression	50.0	59.6	60.7	55.6	*75.3*	47.2	45.5	56.7	49.4	66.3	56.2	59.0	57.3	59.0	60.7	66.9	53.4	56.7	53.4	66.9
Fine KNN	73.0	56.2	61.2	68.0	*83.1*	59.6	62.9	59.6	68.0	80.3	66.3	55.1	57.3	62.9	82.0	65.2	71.3	74.7	75.8	82.0
Medium KNN	69.7	50.0	50.0	60.7	53.4	58.4	50.0	50.0	50.0	53.9	71.3	50.0	50.0	50.6	66.9	*80.3*	50.6	51.1	53.4	52.8
Coarse KNN	50.0	50.0	*50.6*	50.0	50.0	*50.6*	50.0	51.1	50.0	50.0	50.0	50.0	48.9	50.0	50.0	50.0	50.0	50.0	50.0	50.0
Cosine KNN	59.0	**80.9**	78.7	70.8	78.7	46.6	70.2	**84.8**	61.8	79.2	65.2	76.4	*86.0*	79.8	75.3	71.9	69.2	67.4	69.1	62.9
Cubic KNN	69.1	49.4	48.3	56.2	53.4	59.6	50.0	50.0	50.0	53.4	71.3	50.0	50.0	51.1	65.7	*79.2*	52.8	51.1	52.8	50.6
Weighted KNN	71.9	50.0	51.1	62.9	57.9	61.8	50.0	50.0	50.0	59.6	71.9	50.0	50.0	51.7	78.1	*83.1*	55.6	56.2	56.7	55.6
Boosted trees	63.5	49.4	49.4	49.4	49.4	66.3	49.4	49.4	49.4	49.4	*69.1*	49.4	49.4	49.4	49.4	49.4	49.4	49.4	49.4	49.4
Bagged trees	72.5	79.2	77.5	75.8	**96.1**	63.5	83.1	83.7	79.8	98.3	71.9	77.0	80.3	78.7	81.5	92.7	96.6	97.8	97.8	*98.9*
Subspace discriminant	50.0	58.4	67.4	70.8	72.5	45.5	57.9	70.2	51.4	71.9	56.7	60.1	69.7	78.1	*78.7*	64.6	69.1	70.2	70.8	75.8
Subspace KNN	69.1	56.2	63.5	70.2	85.4	65.7	61.8	60.7	67.4	86.0	71.3	55.1	58.4	62.9	84.4	*92.7*	70.2	86.5	75.8	86.5
RUSBoosted tree	59.6	52.8	49.4	49.4	49.4	60.1	49.4	49.4	53.9	49.4	69.1	49.4	52.2	51.7	49.4	49.4	52.2	*74.7*	52.8	49.4

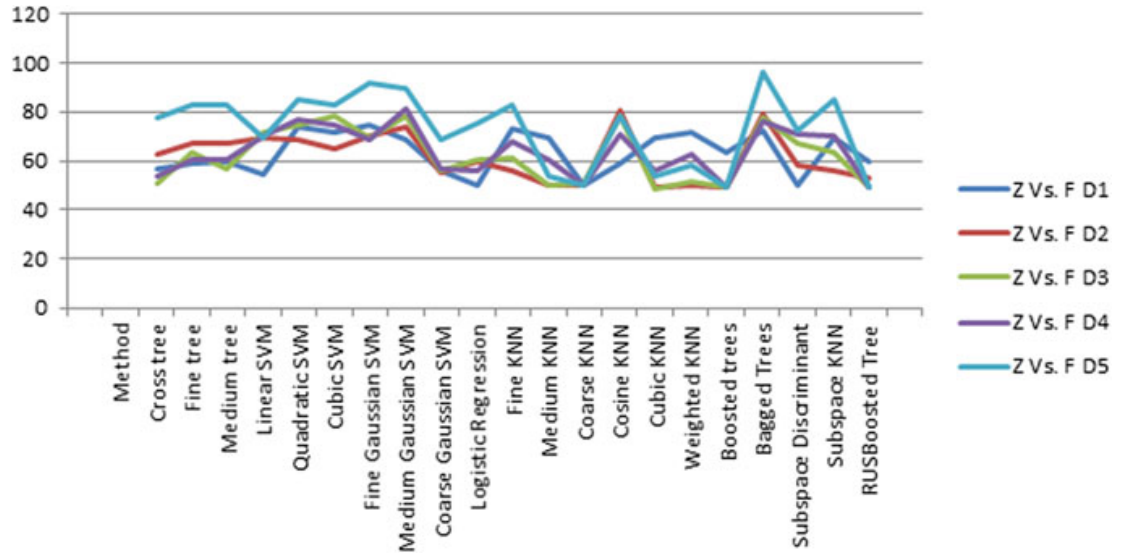


Fig. 7 The performance of different classifiers on different DWT features (Z vs. F)

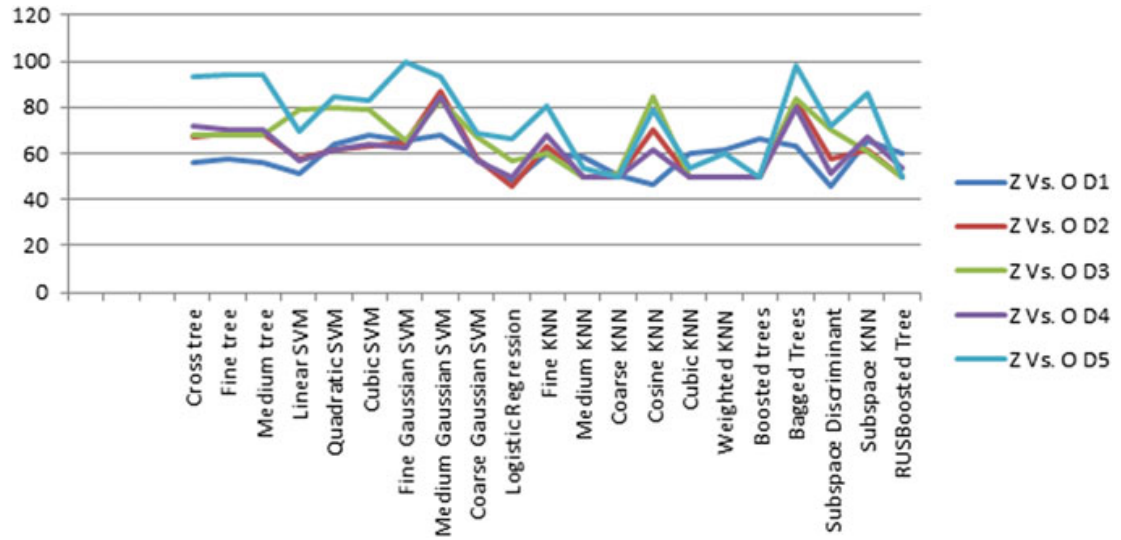


Fig. 8 The performance of different classifiers on different DWT features (Z vs. O)

4 Conclusion and Future Work

In this work Bonn University dataset which consist of five sub samples is individually subjected to all machine learning techniques i.e. SVM, KNN, Decision Tree, Logistic Regression and Ensemble classifiers. For feature extraction to increase performance we have used DWT. Our experimental result shows highest accuracy on using Fine Gaussian SVM machine learning technique on this data set.

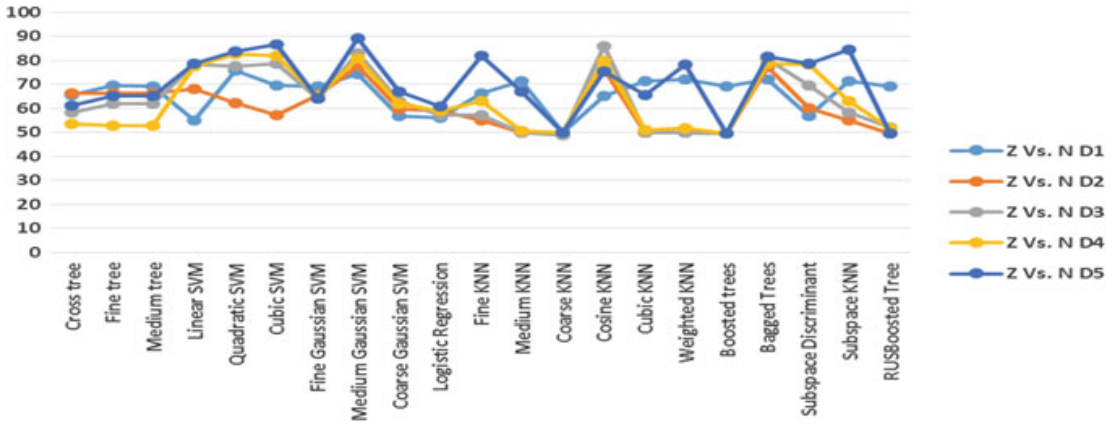


Fig. 9 The performance of different classifiers on different DWT features (Z vs. N)

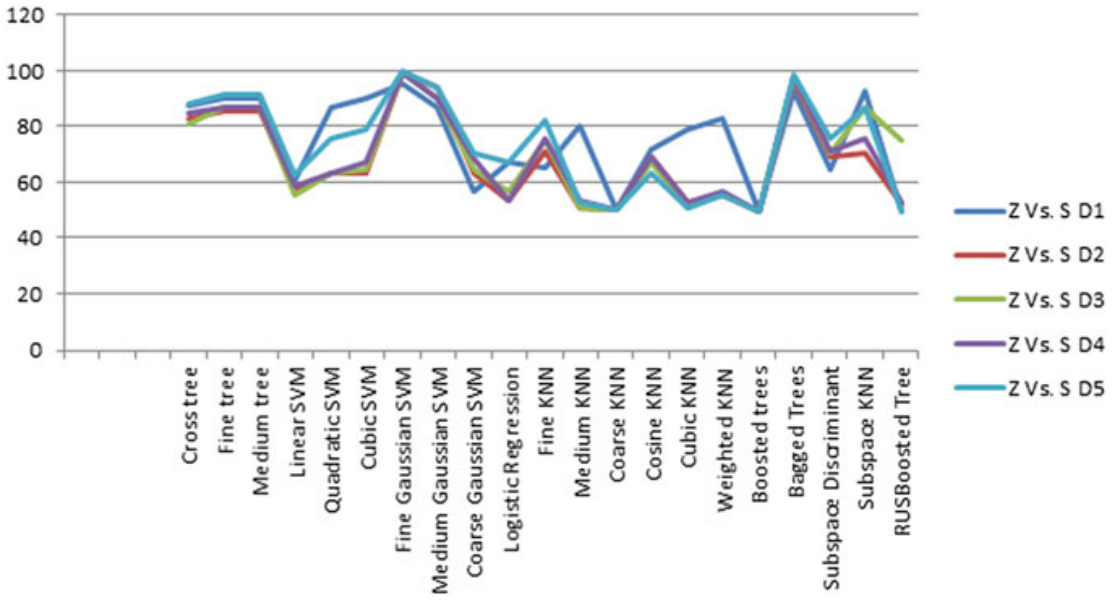


Fig. 10 The performance of different classifiers on different DWT features (Z vs. S)

Table 2 Compare research for detection of normal and epileptic classes

Author	Classifier	Accuracy (%)
Inan and Elif [27]	SVM	99.28
Varun and Ram [28]	Bandwidth features (BAM and BF M) and the LS-SVM	99.50
Abdulhamit [29]	Discrete wavelet transform (DWT), mixture of expert model	95.00
Kamal and Salih [30]	FFT-decision tree classifier	98.72
Yuan and Qi [31]	Nonlinear features + ELM	96.5
Our work	Fine Gaussian Support Vector Machine (SVM)	100.00

Here, we achieved highest accuracy with 100% for seizure classification and detection. Since the proposed framework was designed for limited number of EEG signals and machine learning modalities, it can also be applied and executed for a large data set and different machine learning or deep learning architectures.

References

1. Guo, P., Wang, J., Gao, X.Z., Tanskanen, J.M.: Epileptic EEG signal classification with marching pursuit based on harmony search method. In: 2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC), pp. 283–288 (2012)
2. Guo, L., et al.: Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks. J. Neurosci Methods **191**(1), 101–109 (2010)
3. Selvan, S., Srinivasan, R.: Removal of ocular artifacts from EEG using an efficient neural network based adaptive filtering technique. IEEE Signal Process. Lett. **6**(12), 330–332 (1999)
4. WHO Report. http://www.who.int/mental_health/neurology/epilepsy/en/. Accessed Jan 2018
5. Talathi, S.S.: Deep Recurrent Neural Networks for Seizure Detection and Early Seizure Detection Systems. arXiv preprint arXiv:1706.03283 (2017)
6. Acharya, U.R., et al.: Automated EEG analysis of epilepsy: a review. Knowl. Based Syst. **45**, 147–165 (2013)
7. Salanova, V., et al.: Long-term efficacy and safety of thalamic stimulation for drug-resistant partial epilepsy. Neurology **84**(10), 1017–1025 (2015)
8. Cook, M.J., et al.: Prediction of seizure likelihood with a long-term, implanted seizure advisory system in patients with drug-resistant epilepsy: a first-in-man study. Lancet Neurol **12**(6), 563–571 (2013)
9. Iasemidis, L.D., et al.: Adaptive epileptic seizure prediction system. IEEE Trans. Biomed. Eng. **50**(5), 616–627 (2003)
10. Orosco, L., Agustina, G.C., Eric, L.: A survey of performance and techniques for automatic epilepsy detection. J. Med. Biol. Eng. **33**(6), 526–537 (2013)
11. Guo, L., et al.: Automatic feature extraction using genetic programming: an application to epileptic EEG classification. Expert Syst. Appl. **38**(8), 10425–10436 (2011)
12. Selvan, S., Srinivasan, R.: Removal of ocular artifacts from EEG using an efficient neural network based adaptive filtering technique. IEEE Signal Process. Lett. **6**(12), 330–332 (1999)
13. Cho, K., Van Merriënboer, B., Bahdanau, D., Bengio, Y.: On the Properties of Neural Machine Translation: Encoder-Decoder Approaches. arXiv preprint arXiv:1409-1259
14. Zhou, W., Gotman, J.: Removal of EMG and ECG artifacts from EEG based on wavelet transform and ICA. In: 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEMBS'04, vol. 1, pp. 392–395 (2004)
15. Parvez, M.Z., Paul, M.: Epileptic seizure prediction by exploiting spatiotemporal relationship of EEG signals using phase correlation. IEEE Trans. Neural Syst. Rehabilit. Eng. **24**(1), 158–168 (2016)
16. Cho, K., Van Merriënboer, B., Bahdanau, D., Bengio, Y.: On the Properties of Neural Machine Translation: Encoder-Decoder Approaches. arXiv preprint arXiv:1409-1259 (2014)
17. Ammar, S., Senouci, M.: Seizure detection with single-channel EEG using Extreme Learning Machine. In: 2016 17th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA, pp. 776–779 (2016)
18. Temko, A., et al.: EEG-based neonatal seizure detection with support vector machines. Clin. Neurophysiol. **122**(3), 464–473 (2011)
19. Acharya, U.R., et al.: Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework. Expert Syst. Appl. **39**(10), 9072–9078 (2012)

20. Liu, Y., et al.: Automatic seizure detection using wavelet transform and SVM in long-term intracranial EEG. IEEE Trans. Neural Syst. Rehabilit. Eng. **20**(6), 749–755 (2012)
21. Übeyli, ElifDerya: Wavelet/mixture of experts network structure for EEG signals classification. Expert Syst. Appl. **34**(3), 1954–1962 (2008)
22. Engin, Mehmet: ECG beat classification using neuro-fuzzy network. Pattern Recognit. Lett. **25**(15), 1715–1722 (2004)
23. Garrett, D., et al.: Comparison of linear, nonlinear, and feature selection methods for EEG signal classification. IEEE Trans. Neural Syst. Rehabilit. Eng. **11**(2), 141–144 (2003)
24. Yazdani, A., Ebrahimi, T., Hoffmann U.: Classification of EEG signals using Dempster Shafer theory and a k-nearest neighbor classifier. In: 4th International IEEE/EMBS Conference on Neural Engineering, NER'09, IEEE, (2009)
25. Subasi, Abdulhamit, Ercelebi, Ergun: Classification of EEG signals using neural network and logistic regression. Comput. Methods Programs Biomed. **78**(2), 87–99 (2005)
26. Polat, Kemal, Güneş, Salih: Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. Appl. Math. Comput. **187**(2), 1017–1026 (2007)
27. Panda, R., et al.: Classification of EEG signal using wavelet transform and support vector machine for epileptic seizure diction. In: 2010 International Conference on Systems in Medicine and Biology (ICSMB), IEEE (2010)
28. Bajaj, V., Pachori, R.B.: Classification of seizure and nonseizure EEG signals using empirical mode decomposition. IEEE Trans. Inf. Technol. Biomed. **16**(6), 1135–1142 (2012)
29. Subasi, Abdulhamit: EEG signal classification using wavelet feature extraction and a mixture of expert model. Expert Syst. Appl. **32**(4), 1084–1093 (2007)
30. Polat, Kemal, Güneş, Salih: Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform. Appl. Math. Comput. **187**(2), 1017–1026 (2007)
31. Yuan, Q., et al.: Epileptic EEG classification based on extreme learning machine and nonlinear features. Epilepsy Res. **96**(1–2), 29–38 (2011)

Novel Approach for Plant Disease Detection Based on Textural Feature Analysis

Varinderjit Kaur and Ashish Oberoi

Abstract The image processing is the technique which can propose the information stored in the form of pixels. The plant disease detection is the technique which can detect the disease from the leaf. The plant disease detection algorithms have various steps like preprocessing, feature extraction, segmentation, and classification. The KNN classifier technique is applied which can classify input data into certain classes. The performance of KNN classifier is compared with the existing techniques and it is analyzed that KNN classifier has high accuracy, less fault detection as compared to other techniques. This paper presents methods that use digital image processing techniques to detect, quantify, and classify plant diseases from digital images in the visible spectrum. In plant leaf classification leaf is classified based on its different morphological features. Some of the classification techniques used are neural network, genetic algorithm, support vector machine, and principal component analysis. In this paper results are compared between KNN classifier and SVM classifier.

Keywords GLCM · KNN · K-means · WDDIP-KNN · Plant disease detection

1 Introduction

Agriculture helps to create wealth, improvise farmer's livelihood and decrease the rate of hunger and poverty especially in the remote areas. Agriculture plays an important role to improve the food and life security. But nowadays farming is suffering from various diseases, parasites, shortage of irrigation facilities and lack of best-suited fertilizers. Diseases are not the only reason for the downfall of farm marketing but unhygienic and improper sanitation at farms are also the major issues

V. Kaur · A. Oberoi (✉)
Department of CSE, RIMT University, Mandi Gobindgarh, India
e-mail: a_oberoi01@yahoo.co.in

V. Kaur
e-mail: Vari006rupi@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_30

for the diseases to human. Various methods are used for the detection of plant diseases in which a vast image processing techniques are used. The plant diseases can be analyzed by the texture, color, and structure of the leaf [1]. The information regarding the diseases and species of any plant or crop is given by the leaves, by studying the structure of a leaf researcher can generalize about the diseases of any crop. Since there are huge impacts of plant disease detection on the quality and quantity of the agricultural products, there is a need to study this field and present improvements such that the issues that are being faced can be eliminated. Agriculture is the field on which around 70% of the population of India relies. Appropriate fruit and vegetable crops to be grown can be chosen by the farmers from diverse choices available. However, the technicality level with which these crops are cultivated such that optimum yield and quality product can be produced, is very high. Technical support is highly useful in fulfilling these requirements. There is a need to monitor the fruit crops very closely so that they can be managed efficiently such that no diseases affect their productivity. Depending upon the high resolution multispectral and stereo images, the leaf diseases are to be detected and classified automatically and to do so several researchers have proposed various techniques. Providing a direct economical optimization of the agricultural products is not the only requirement here [2]. There is a need to ensure that the living beings or the environment is not harmed due to them. The crop production particularly needs to reduce the level at which the water, soil, and food resources are being contaminated due to the pesticides. Further, the proposed methods also need to have higher speed and accuracy to provide efficient results as per the technology advancements. The plant disease detection process includes some basic steps which are

a. *Input image* Capturing a sample image using a digital camera is the initial step of this process. Further, the features are extracted from the image to perform operations. The important features are stored initially within the database and the further processing will be performed in the next steps.
b. *Image database* In the next step, an image database is generated that includes all the images which will further be used to train and test the extracted data. The application is clearly the most important factor on which the image database is constructed. The efficiency of the classifier is responsible to decide the robustness of an algorithm and it completely relies on the image database.
c. *Image preprocessing* The operations that are performed on the images at the lowest level of abstraction are collectively known as image preprocessing. The image data is improved and the distortions that are unnecessarily present are suppressed within this process [3]. The processing and task analysis are performed with the help of image features which can be improved through this process. The information content on the image is not increased due to this process. Sufficient amount of redundancy is provided in the images with the help of this method. The brightness value, of neighboring pixels that belong to

the same real object, is also same. In the form of an average value of the neighboring pixels, it is possible to restore a pixel that has been distorted within an image. The image that has been stored within the image database is captured with the help of image preprocessing techniques as well.

d. *Feature extraction* The features that help in determining the meaning of a given sample are used for the identification and extraction of features from it. Color, shape and texture features are mainly included as image features in case of image processing.
e. *Recognition* and *classification* Training and classification are the two broader phases in which the complete recognition process is divided. The numerical characteristics of several image features are analyzed using image classification. Further, the data is categorized in an organized way through this process [4]. Training and testing are the two phases that are used by classification algorithms. The isolation of properties of basic image features is done within the initial training process. Further, a training class is introduced with the help of these isolated features, which has a unique description for each classification category. The image features are classified within the further testing phase with the help of these feature-space divisions created here.

A nonlinear classifier that is used to solve several pattern recognition issues is known as Support Vector Machine (SVM). For achieving better performance of classification, mapping of nonlinear input data is done to the linearly separated data within certain high dimensional space in SVM. The marginal distance present among various classes is increased through SVM [5]. Various kernels are used to divide the classes. Only two classes are used by SVM and to partition them, a hyper plane is defined by it. To do so, the margin that exists between the hyper plane and the two classes is maximized. Support vectors are known as the samples that are nearest to the margin and were chosen such that the hyper place could be determined.

2 Literature Review

Mattihalli et al. [9] proposed a novel approach for detecting the diseases from leaves. Here, the leaf images are taken such that few important features can be extracted which can further be used in this proposed work. A device named as Beagle bone black is used in this work which also consists of a web camera that helps in identifying diseases from the leaf samples. The device includes a database in which there are few pre-stored leaves images. This paper compares the pre-stored images with the images captures from the Beagle bone black device. GSM is used to transmit the information related to detect the diseases and operate the valves. As per the simulation results, it is seen that the proposed approach provides very low cost, user friendly and highly efficient results.

Tichkule and Gawali [10] reviewed various techniques through which plants diseases can be detected. Image processing technique is used for the detection and identification of diseases caused by bacteria, fungi, virus, and excessive use of insecticides. Therefore, this method is used to classify the diseases caused in agricultural field. The authors accurately detected and classified diseases on various plants using all above techniques. K-Means Clustering method is used to detect infected plants and Neural Networks provides the accuracy in detection and classification of diseases. So, these methods have ability to be used in Agrobot system.

Tlhobogang and Wannous [11] proposed the study that investigated the problems related to unavailability, irrelevant, and less accurate farming information. The main objective was to deal with infected plants and to diagnose them. Image analysis, convolutional neural networks and the knowledge of machine learning offer a stable and movable solution. A Science Research Methodology was used in framing the prototype. Proper and systematic studies on farmers' agricultural techniques and effect of recently on-going projects are used to understand the benefits of mobile based agricultural services. The contribution of ICT in agricultural fields, collecting, storing, and disseminating the development are required by the farmers to increase their market value.

Gandhi et al. [12] presents an image-related classification which identifies the plant diseases. As, there are many datasets used in other countries but none of them are connected to India. So, there is a need to develop a local dataset for the benefits of Indian farmers. Generative Adversarial Networks (GANs) are used to limit the number of local available images. A Convolutional Neural Networks (CNN) is employed on the smart mobile applications. This application can be easily installed in any smart phone and the farmer needs to move through the field and capture the images. This could capture the several images which can be further send to the server through which the model runs and gives the classification.

Khan et al. [13] presented multilevel segmentation method in which initial segmentation is performed using expectation maximization algorithm. In this method, there is a very few chances of information loss. Finally, a region is taken out from the quantized image by empowering the binary partitioned tree which is further used in principle Eigen vector. Post processing methods are used to eradicate the useless fragments and to resolve the image labeling problem. When the salient region of the image belongs to single class this technique is mostly used. The new cascaded design in primary color segmentation with the confirmation of infected regions extractions gives the experimental analysis.

3 Research Methodology

This research work is based on the detection of plant disease. To detect plant disease with the proposed technique following steps are followed

1. *Preprocessing* The image is taken as input on which disease need to detect. The image is converted to the grayscale for the feature extraction.
2. *Feature extraction* In the second phase the algorithm of GLCM is applied which can extract features of the input image and store in the database. A tabular description which shows the number of times various combinations of pixel brightness values occur within an image is known as Gray-Level Co-occurrence Matrix. The locations of pixel that has alike gray level values are presented by GLCM [7]. The pairs of gray level values are provided as input from GLCM calculation units. When the original and predictive images are compared, the deviation present in them is presented here. The relationship among two neighboring pixels is considered within GLCM. The initial pixel is called the reference pixel whereas the neighbor pixel is known to be the second one. A two dimensional array in which a set of possible image values are presented in rows and columns is known as a co-occurrence matrix.
3. *Segmentation* The k-mean segmentation algorithm is applied which can segment input image into certain parts. K-means is the data clustering algorithm in which the numbers of clusters within the data are pre-specified is known as k-means algorithm. A trial-and-error method in which appropriate numbers of clusters for particular data set are identified has made it difficult to define correct clustering method [8]. Through the selection of K value, the performance of a clustering algorithm is affected. Thus, there is a need to adopt a set of values instead of using a single predefined K. For reflecting the special properties of data sets, the consideration of large number of values is important. Further, in comparison to the number of objects present in the data sets, there need to be less number of selected values.
4. *Classification* The classification detects the name of the disease with which plant is affected. The system is training with the number of images and test set is given as input after extracting image features. The Proposed Model (WDDIP-KNN) classifier is applied to detect the disease name and flow chart is blow. A classifier, that performs classification through the recognition of neighbors that are closest to the query examples, is known as k-Nearest Neighbor. The class of query is determined with the help of these identified neighbors. In this algorithm, on the basis of the least distance present between the given point and other points is calculated which is then used to perform classification that helps in determining the class within which a particular point belongs [6]. There is no training process involved within this classifier. Since, the robustness to noisy data for this classifier is zero it is not applied within very large number of training examples. The calculation of Euclidean distance between the test and training samples is done which is further used within plant leaf classification. Thus, similar measures are identified in this manner which further helps in identifying the class for test samples as well (Fig. 1).

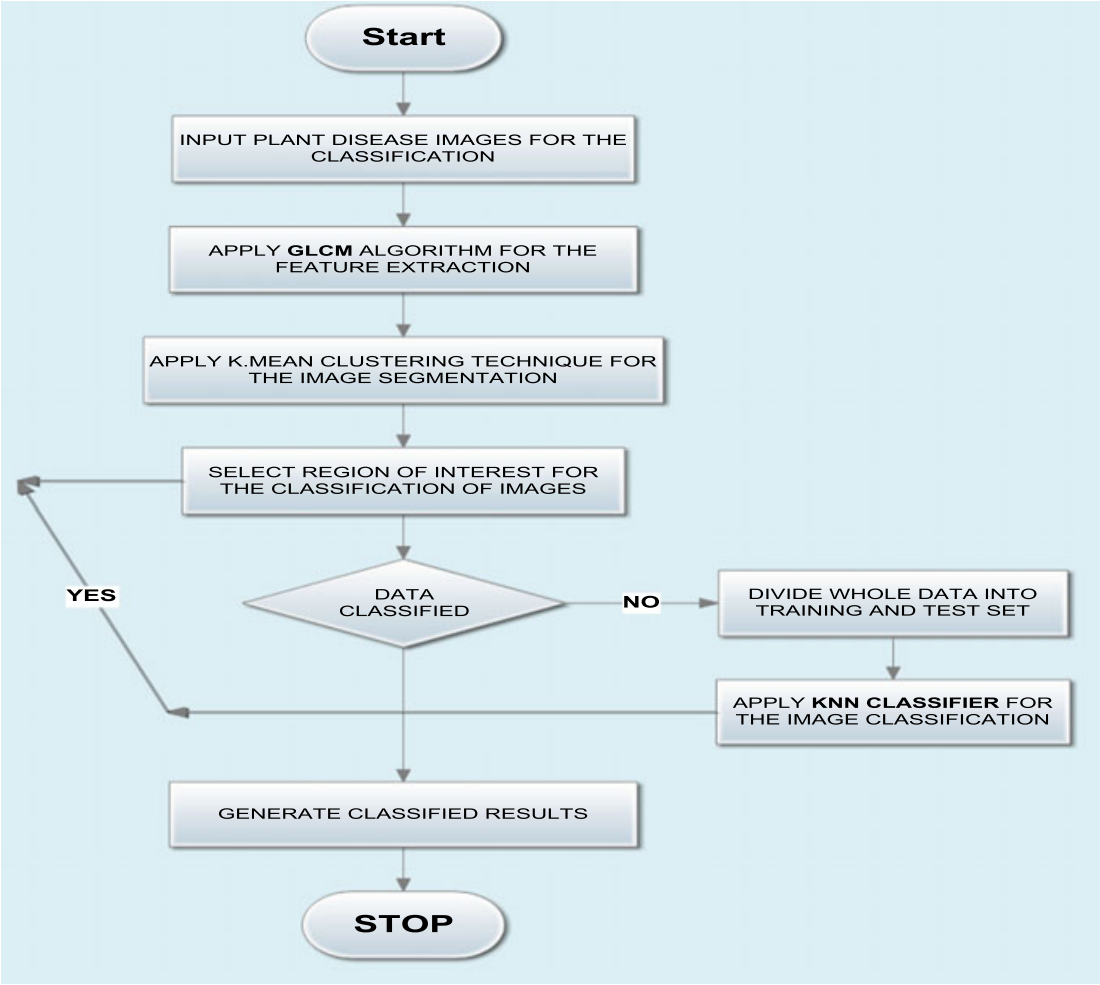


Fig. 1 Proposed model (WDDIP-KNN) control flow

4 Results and Discussion

The technique is proposed for the plant disease detection based on feature detection and classification. The GLCM algorithm is applied for the feature extraction. The k-mean clustering is applied for the image segmentation and KNN algorithm is applied for the disease classification. The Data set of about 25 images is taken as input to prepare the training set. The results are analyzed in terms of certain parameters which are described below.

As shown in Fig. 2, the fault of the proposed and existing algorithm is compared for the performance analysis. It is analyzed that proposed KNN technique has less faults as compared to existing technique SVM (Table 1).

$$\text{Fault Detection Rate} = 100 - \text{accuracy}$$

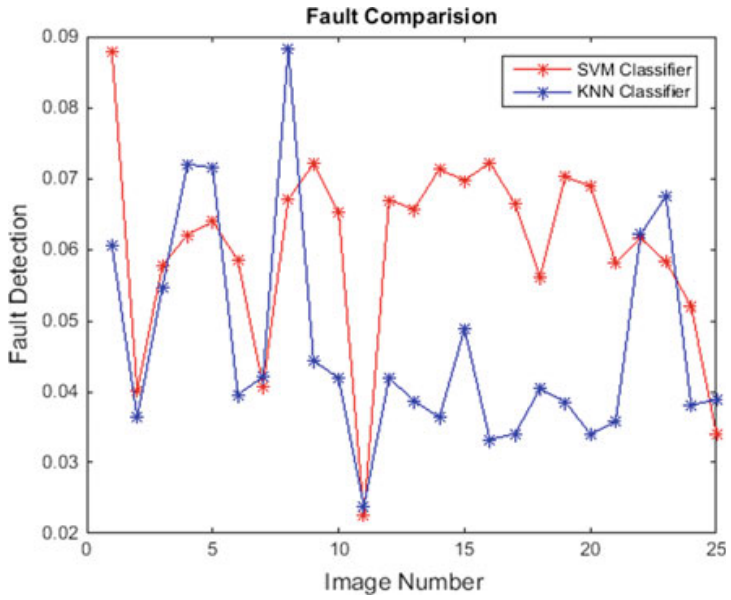


Fig. 2 Fault detection comparison

Table 1 Fault detection rate comparison

Image no	Existing technique (SVM classifier)	Proposed technique (WDDIP-KNN)
1	0.087951	0.060472
2	0.040103	0.036336
3	0.057707	0.054626
4	0.062027	0.07203
5	0.063903	0.071472
6	0.058424	0.039559
7	0.040812	0.042147
8	0.067034	0.088288
9	0.072044	0.044439
10	0.065199	0.041873
11	0.022471	0.023817
12	0.066916	0.041897
13	0.065681	0.038756
14	0.071313	0.036336
15	0.069679	0.048892
16	0.072199	0.033249
17	0.066446	0.034057
18	0.056053	0.040433
19	0.070209	0.038563
20	0.068922	0.033993
21	0.058104	0.035843
22	0.061631	0.062198
23	0.058188	0.067419
24	0.051977	0.038162
25	0.033993	0.038839

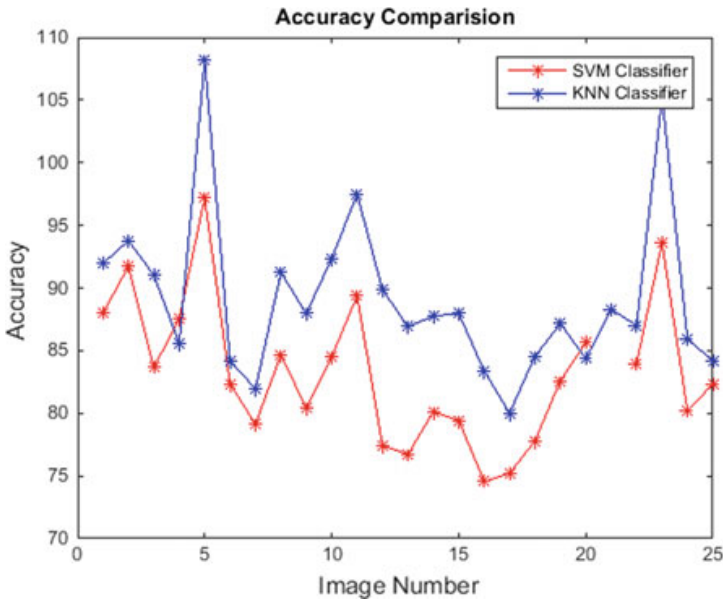


Fig. 3 Accuracy comparison

As shown in Fig. 3, the accuracy of the proposed KNN technique is compared with the existing SVM technique. It is analyzed that accuracy of the proposed algorithm is high as compared to existing algorithm.

As shown in Fig. 4, the accuracy of the SVM technique is 80.02954 and fault rate is 0.84811 (Fig. 5).

As shown in Fig. 4, the accuracy of the KNN Classifier is 92.25974 and fault rate is 0.94918 (Table 2).

$$\text{Accuracy} = \frac{\text{Total no of points classified}}{\text{Total no of points}} * 100$$

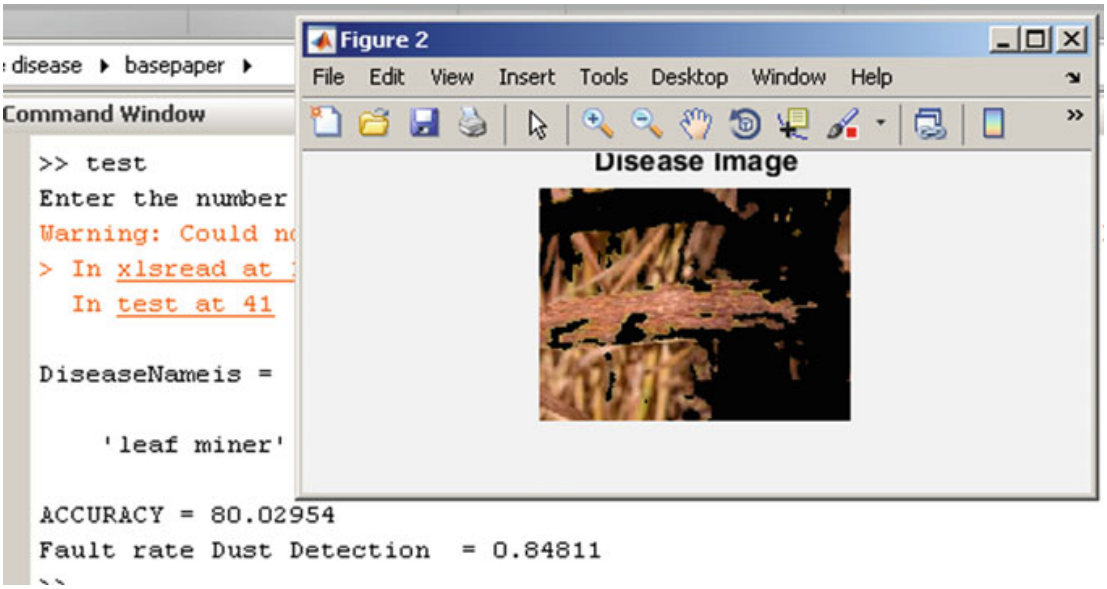


Fig. 4 Disease Name, accuracy and fault rate for one image by SVM classifier

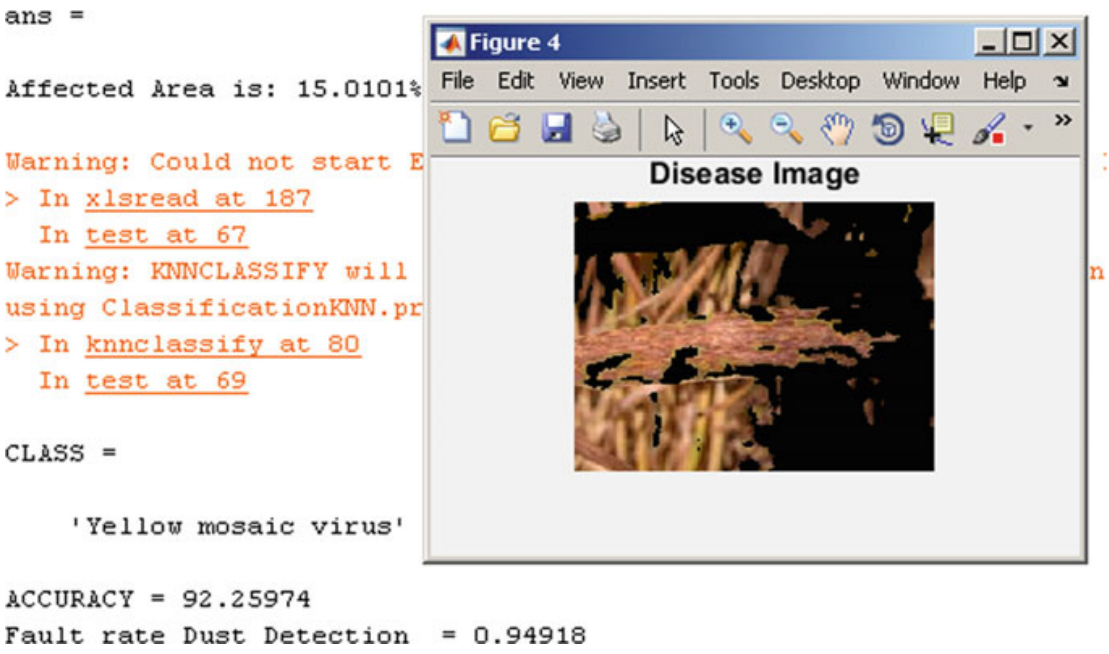


Fig. 5 Disease name, accuracy and fault rate for one image by KNN classifier

Table 2 Accuracy comparison

Image no	Existing technique (SVM classifier)	Proposed technique (WDDIP-KNN)
1	91.97587	88.0339
2	93.75906	91.69698
3	91.02909	83.65844
4	85.55519	87.54323
5	108.1948	97.18758
6	84.1623	82.33363
7	81.95833	79.13637
8	91.23152	84.67711
9	87.96284	80.39575
10	92.37004	84.52256
11	97.47806	89.3719
12	89.84999	77.41904
13	86.92375	76.71102
14	87.75028	80.12171
15	87.96963	79.43393
16	83.36096	74.55252
17	80.01366	75.19481
18	84.559	77.79646
19	87.15097	82.59238
20	84.35514	85.65128
21	88.27116	85.46842

(continued)

Table 2 (continued)

Image no	Existing technique (SVM classifier)	Proposed technique (WDDIP-KNN)
22	87.02477	83.97361
23	105.2089	93.60827
24	93.37007	95.35002
25	90.85422	92.46523

5 Conclusion

In this work, it is concluded that plant disease detection is the technique (WDDIP-KNN) which is applied to detect disease from the plants. The technique of plant disease detection has various steps like feature extraction, segmentation, and classification. In this paper, novel technique(WDDIP-KNN) is proposed based on the GLCM algorithm for the feature extraction, K-mean segmentation is applied for the image segmentation and KNN classification technique is applied for the disease classification. The performance of proposed algorithm is compared with the existing algorithm in terms of accuracy and fault rate. It is analyzed that KNN classifier has high accuracy and less fault rate as compared to existing technique.

References

1. Camargo, A., Smith, J.S.: An image-processing based algorithm to automatically identify plant disease visual symptoms. Bio Syst. Eng. **102**, 9–21 (2008)
2. Camargo, A., Smith, J.S.: Image processing for pattern classification for the identification of dis-ease causing agents in plants. Comput. Electron. Agric. **66**, 121–125 (2009)
3. Guru, D.S., Mallikarjuna, P.B., Manjunath, S.: Segmentation and classification of tobacco seedling diseases. In: Proceedings of the Fourth Annual ACM Bangalore Conference (2011)
4. Zhao, Y.X., Wang, K.R., Bai, Z.Y., Li, S.K., Xie, R.Z., Gao, S.J.: Research of maize leaf disease identifying models based image recognition. In: Crop Modeling and Decision Support, pp. 317–324. Tsinghua University Press, Beijing (2009)
5. Fury, T.S., Cristianini, N., Duffy, N.: Support vector machine (SVM) classification and validation of cancer tissue samples using microarray expression data. Proc. BioInform. **16** (10), 906–914 (2000)
6. Al-Hiaryy, H., Bani Yas Ahmad, S., Reyalat, M., Ahmed Braik, M., AL Rahamnehiahh, Z.: Fast and accurate detection and classification of plant diseases. Int. J. Comput. Appl. **17**(1), 31–38 (2011)
7. Mohanaiah, P., Sathyanarayana, P., GuruKumar, L.: Image texture feature extraction using GLCM approach. Int. J. Sci. Res. Publ. **3**(5), 1 (2013)
8. Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R., Wu, A.Y.: An efficient k-means clustering algorithm: analysis and implementation. IEEE Trans. Pattern Anal. Mach. Intell. **24**(7), 881–892 (2002)
9. Mattihalli, C., Gedefaye, E., Endalamaw, F., Necho, A.: Real time automation of agriculture land, by automatically detecting plant leaf diseases and auto medicine. In: 32nd International Conference on Advanced Information Networking and Applications Workshops (2018)

10. Tichkule, S.K., Gawali, D.H.: Plant diseases detection using image processing techniques. In: Online International Conference on Green Engineering and Technologies (IC-GET) (2016)
11. Tlhobogang, B., Wannous, M.: Design of plant disease detection system: a transfer learning approach work in progress. IEEE (2018)
12. Gandhi, R., Nimbalkar, S., Yelamanchili, N., Ponkshe, S.: Plant disease detection using CNNs and GANs as an augmentative approach. IEEE (2018)
13. Khan, Z.U., Akra, T., Naqvi, S.R., Haider, S.A., Kamran, M., Muhammad, N.: Automatic detection of plant diseases; utilizing an unsupervised cascaded design. IEEE (2018)

Novel Approach for Brain Tumor Detection Based on Naïve Bayes Classification

Gurkarandesh Kaur and Ashish Oberoi

Abstract The brain tumor detection is the approach which can detect the tumor portion from the MRI image. To detect tumor from the image various techniques has been proposed in the previous times. The technique which is adapted in research work is based upon morphological scanning, clustering, and Naïve Bayes classification. The morphological scanning will scan the input image and clustering will cluster similar and dissimilar patches from image then Naïve Bayes classifier spot the tumor portion from magnetic resonance imaging. The advance algorithm is implemented in MATLAB and results are analyzed in terms of PSNR, MSE accuracy, and fault detection and also calculate overlapping area with dice coef. The proposed method has been tested on data set with more than 25 slide scanned images. This proposed method achieved accuracy with 86% best cell detection.

Keywords Brain tumor · Clustering · MRI · Morphological scanning · Naïve Bayes NBC-BTD model

1 Introduction

Several lives have been affected because of a common brain disease known as brain tumor. The patients suffering from this disease have not survived in most of the cases. For fighting this disease, several techniques have been proposed such that the knowledge related to medicine can be expanded and one can understand calculations in a better manner such that the tumor can be detected. Due to the high complexity of brain images and the fact that only expert physicians can analyze the tumors, brain tumor detection is a challenging task within medical image processing [1]. To detect brain tumor from various images, the two most common tests that are

G. Kaur · A. Oberoi (✉)
Department of CSE, R.I.M.T University, Mandi Gobindgarh, India
e-mail: a_oberoi01@yahoo.co.in

G. Kaur
e-mail: deesh.mangat@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_31

applied are Magnetic Resonance Image (MRI) and Computer Tomography (CT) scan of brain. Further, to perform various treatments, the location of tumor is also identified through this technique. To heal brain tumor, several treatment techniques are proposed today such as radiation therapy, chemotherapy as well as surgery. On the basis of size, type of tumor as well as its grade, the treatment type is chosen. To check whether other parts are being affected by this tumor or not, it is also important to perform certain analysis. When the appropriate treatment method has to be chosen by the doctor, there are certain factors that are to be considered. The possible side effects of a treatment, consideration of complete health and checking whether the central nervous system is affected due to the tumor or not, are few of these factors. Radio imaging is the most commonly applied technique within MRI due to its dynamicity and flexibility [2]. Various pulse sequences and modification in imaging parameters that are based on Longitudinal Relaxation Time (LRT/T1) and Transverse Relaxation Time (TRT/T2) are used to perform acquisition of variable image contrast. Particular tissue properties are provided in relevance to signal intensities provided on the weighted images T1 and T2. On the basis of pulse sequence parameters, the contrast on MR images is provided. For knowing the details of structures of various organs of the body such as liver, chest, and brain, MRI imaging of the body is done. The treatment can be monitored in the patient efficiently with the help of this approach. There are certain steps performed in order to identify the tumor in the patient's body [3]. Preprocessing, segmentation, feature extraction as well as classification are the commonly applied steps. The MRI samples are gathered at the initial stage.

a. Preprocessing and Enhancement: The chances that a suspicious region can be detected can be improved through this initial step being performed in image processing. From the image, the noise is eliminated and finer details are extracted. The accuracy of an image is minimized when noise is present within the MRI image. The noise is removed by applying different filters on the image. The filters are also applied to sharpen the image. Since the detection of boundary of tumor can be done more effectively and easily, it is important to sharpen the image with the help of various low pass filters once the noise has been completely eliminated from the images.
b. Segmentation methods: The procedure where the image is broken down to smaller parts and segments is known as image segmentation. The analysis can be performed in easy manner through this step. Several image segmentation methods have been developed over the time. The approach in which the object boundaries are assumed to be defined by the detected edges and which further helps in recognizing these objects is known as edge-based segmentation approach [4]. There is a need to achieve very distinct and closed boundaries to perform direct segmentation which can be done through this approach. False edge detection can occur many times and the partial edges can be joined within an object boundary through edge linking process. The approach in which the bordering pixels present in one area assume to have similar values is known as region based approach. Instead of identifying the edges, the identification of

object region is more important in this case. The pixels are compared with the neighboring pixels. The pixels is said to belong to the cluster in the form of one or more of its neighbors in case when the congruence criteria is satisfied.

c. Feature Extraction: To detect brain tumor from images, the extraction of exact tumor image is very important since the structure of brain is very complex [5]. In order to extract certain features, it is important to consider few parameters. The tumor can easily be classified with the help of results achieved from feature extraction process.

2 Literature Review

Kaur et al. [11] analyze technique of Magnetic Resonance Image (MRI) for brain tumor detection. It shows difficult structure of brain cells with thin network. It also considers solid growth. If we want to study the growth we need to study the fragmentation process, which is a huge disadvantage. This disadvantage can be solved by clustering technique. For this extraction of segmented brain tumor from its area a sobel edge detection is used. In clustering technique, the no of clusters is counted by computing them on the peak of histogram. The size and location can be analyzed by the segmented part of the binary image. The final fragmented part is then use to analyze size and perimeter of the tumor. It concludes that. The brain tumor can be detected using MRI and clustering techniques. So, it is used on the nature of image and the number of peaks, the clusters can be computed.

Hazra et al. [12] reviewed detection and localization of tumor region present in the brain by using patient's MRI. It contain three levels namely, preprocessing, segmentation, and edge detection. Preprocessing converts the original image into grayscale image and eradicates noise if any which further followed by the edge detection using Sober and Canny algorithms with technique of image enhancement. The segmentation is applied to display the tumor affected region. Lastly, the clustering algorithm is used for the image clusters. It results identification of the brain tumor is done efficiently using MRI and K-means algorithms. In order to detect the tumor more accurately the algorithms can be improvised.

Chauhan et al. [13] proposed preprocessed median filtering MRI brain images. In order to separate the area from image- and color-based segmentation and edge detection is done. Histogram of oriented gradients and gray level co-occurrence matrix is used to represent the images. The respected extracted features are stored in the transactional database to classify the tumor into normal benign. The classified accuracy is being calculated 86/6%. This summarized that the proposed system help to know about the type of brain tumor and its further treatment. This system has been successfully tested on the large-sized brain scanned images of brain tumor.

Reema Mathew et al. analysed that the Magnetic Resonance Image (MRI) is effective technique of the brain tumor detection and classification. This classification is done in various steps like preprocessing, filtration of sound, feature

extraction, and segmentation. These methods preprocessed the MRI brain image using anisotropic diffusion filters. The discrete wavelet transforms are extracted in the feature extracted step. These features are further given as the input to the segmentation step. A support Vector Machine was used segmentation and tumor classification. Hence, it concludes that the accuracy of proposed system is 86%. The validation of this method with the recent results can be used in the future proposals.

3 Research Methodology

This research work is relies on brain tumor detection.

The technique to detect tumor are based on following steps:

Step 1: Morphological Operations:

The process through which the structure or shape of an object can be deformed or reconstructed is known as morphology. For the representation of shape of an object, the operations that are applied on binary images are known as morphological operations. While performing pre or post processing, these operations are applied such that the shape of objects or areas can be known in more appropriate way. Following are few of the most commonly used morphological operations:

a. Erosion: The operation with the help of which the boundaries of areas of front-end pixels are eroded from the binary images is known as erosion. In terms of size and holes present within it, the regions of foreground pixels are shrunk. There are two inputs given here [6]. The image is eroded within the initial input and the structuring element is given as the second input. The structuring element place upward of given image such that the erosion of binary image can be calculated. Thus, the origin of structuring element and input pixel coincide with each other.
b. Dilation: The approach through which the holes are filled by adding the pixels to the boundaries of objects present within the image is known as dilation. Two pieces of data are taken as input in this operator. Image is dilated in the initial one and elements are structured in the second one. On the input image, the structuring element is placed for every background pixel such that the given image pixel position and structuring element coincide [7]. Increase the area of foreground pixels is the basic effect of dilation on the binary image. There is a complete closing up of the operation however, in this operation which is its only demerit. There are several classifiers used in the process of detection brain tumor from images. A data structure in the form of a tree is created within a decision tree classifier. On the basis of one particular feature, each interior node that includes decision criteria is based. The entropy reduction that presents the purity of samples is used to calculate the features that are in relevance to classification [8]. The classifier through which two classes are separated using a hyper plane is known as Support Vector Machine (SVM). From the empirical data, an optical

function can be calculated in case when the classes are separated by hyper plane. A basic feed forward based artificial neural network classifier was introduced known as multi-layer perceptron classifier. For performing simple functions, a single hidden layer is used here at first. Further, to improve the classification performance, two hidden layers were included here. For every data set, different hidden units were selected. Across a number of trails, the numbers of hidden neurons were identified. Back propagation algorithm was used to train the neural network.

c. Clustering: In image processing is basically defined as the technique in which groups of identical image primitive are identified. Clustering is a method in which objects are unified into groups based on their characteristics. A cluster is basically an assembly of objects which are similar between them and are not similar to the objects fitting to additional clusters. C-mean clustering is mainly assigning points to cluster or class. In this clusters are mainly indentified by similarity measure, in terms of distance, connectivity and intensity. Moreover, in this technique, each data point belongs to more than one cluster.

Step 2: Naïve Bayes Classifier: Naive Bayes algorithm is effective method of text classification. It works on large training sample set and gives an accurate result. It is a probabilistic classifier based on Bayes theorem with independent assumption which assumes the presence of particular features of a class is unrelated to presence of other features.

Naive Bayes classifier is a simple probabilistic classifier based on applying Bayes' theorem (from Bayesian statistics) with strong (naive) independence assumptions. A more descriptive term for the underlying probability model would be "independent feature model". In simple terms, a naive Bayes classifier assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature. The morphological scanning technique will scan the image and technique of Naïve Bayes is applied which mark the tumor in the image. The classifier that includes all independent attributes when the value of class variable is given is known as Naïve Bayes classifier. Conditional independence is another name for this classifier and it is known to be the easiest form of Bayesian network [9]. Here, the Bayes' theorem is applied along with the naïve assumption that shows the independence among every pair of features within the set of supervised learning algorithms. Following relationship is stated by the Bayes' theorem:

$$P(y|x_1,\ldots,x_n) = \frac{P(y)P(x_1,\ldots,x_n|y)}{P(x_1,\ldots,x_n)} \tag{1}$$

Here, y is a class of variable and from x_1 to x_n a dependent feature vector is included.

PSNR and MSE: The PSNR signal is used to measure the quality of loss and lossless compression (e.g., for image compression). The peek signal used original data. The noise is the error introduced by compression. When comparing

compression codecs, PSNR is an approximation to human perception of reconstruction quality. Although a higher PSNR generally indicates that the reconstruction is of higher quality. When PSNR signal is maximum, the MSE signal is minimum (Fig. 1).

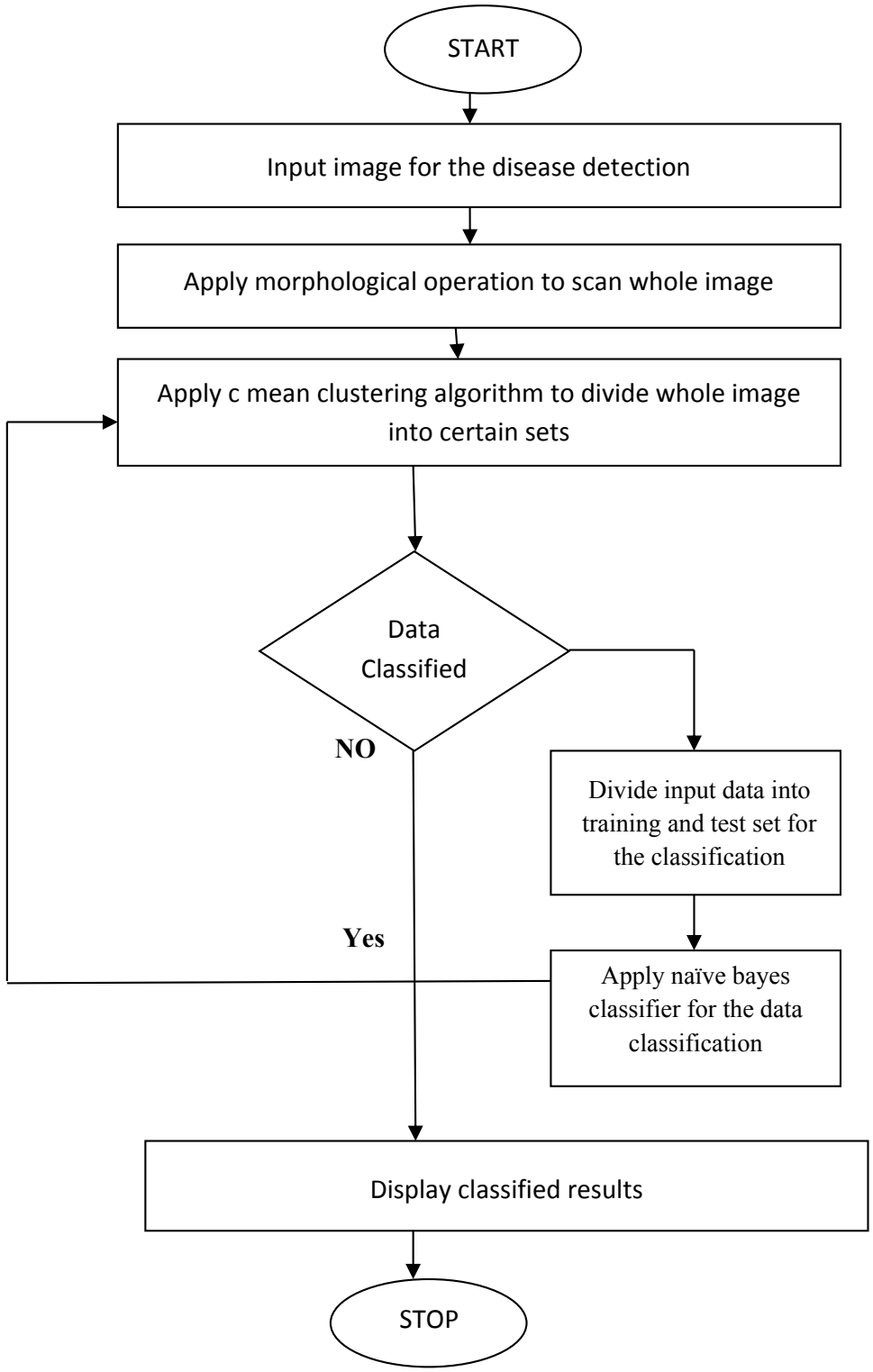


Fig. 1 Proposed flowchart NBC-BTD (brain tumor detection)

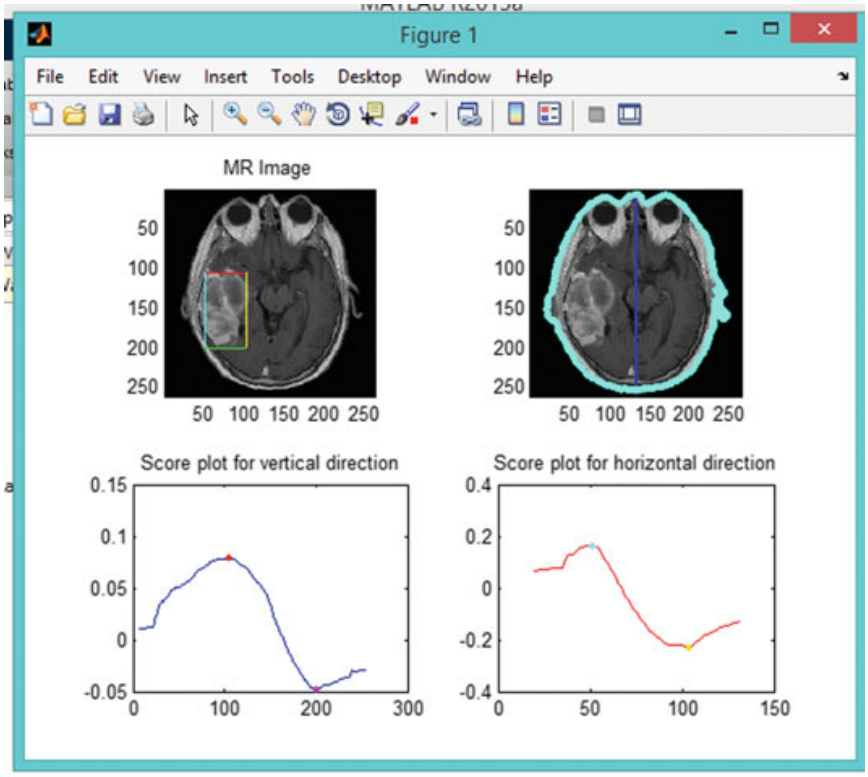


Fig. 2 Naïve Bayes classifier

Result and Discussion: This research work was based on the brain tumor detection; the 20 images taken from data respiratory set. To detect tumor from the MRI images technique of classification was applied. Input the test data for tumor detection, applied in the morphological operations; divided input class into training set; then classified the data. The technique of Naïve Bayes classifier NBC-BTD model marks the tumor portion in the image with horizontal and vertical plot with segmented tumor area. The tumor region grew which segmented tumor would portion from non-tumor region. That gave false positive and negative rate.

As shown in Fig. 2, the technique of Naïve Bayes classifier was applied which marked the tumor portion on the image. The vertical and horizontal position was also calculated from the input MRI Image (Fig. 3; Table 1).

The PSNR value of the Advance and previous research algorithm has compared for the performance analysis. It has analyzed that PSNR value of advance algorithm was high as that to previous research algorithm (Fig. 4; Table 2).

The MSE value of advance and previous research algorithm is compared for the performance analysis. It is analyzed that MSE value of advance algorithm is less as compared to previous research algorithm (Fig. 5; Table 3).

The accuracy value of the advance and previous research algorithm was compared for the performance analysis. It was analyzed that proposed algorithm has high accuracy as compared to previous research algorithm (Fig. 6; Table 4).

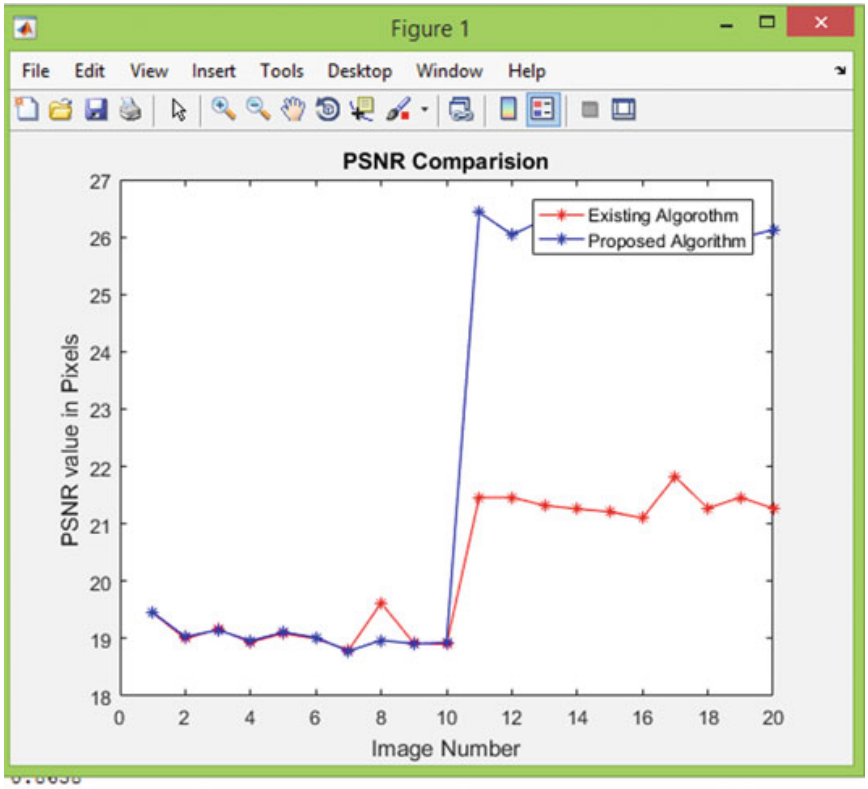


Fig. 3 PSNR comparison

Table 1 PSNR comparison

Image no.	Existing technique (S.V.M classifier)	Proposed technique NBC-BTD
1.	19.45	19.46
2.	19	19.03
3.	19.16	19.15
4.	18.94	18.96
5.	19.09	19.11
6.	19	19.02
7.	18.8	18.78
8.	19.62	18.97
9.	18.92	18.91
10.	18.9	18.93
11.	21.46	26.43
12.	21.46	26.04
13.	21.32	26.3
14.	21.26	26
15.	21.21	26.31
16.	21.1	26.14
17.	21.82	25.83
18.	21.27	26.05
19.	21.46	25.99
20.	21.27	26.12

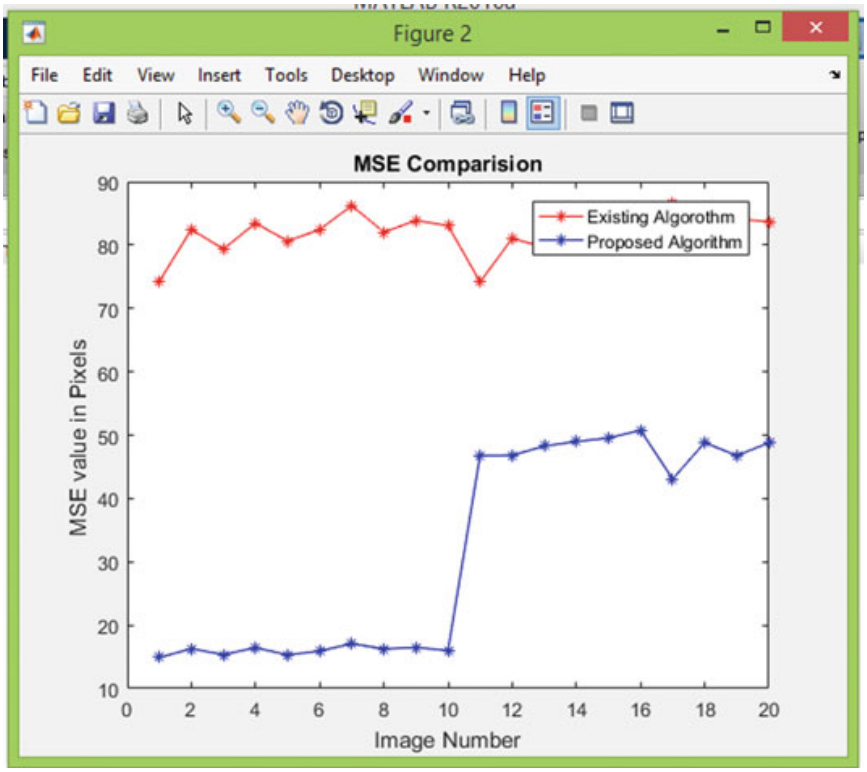


Fig. 4 MSE comparison

Table 2 MSE comparison

Image no.	Existing technique (S.V.M classifier)	Proposed technique NBC-BTD
1.	74.25	14.9
2.	82.45	16.29
3.	79.37	15.36
4.	83.46	16.45
5.	80.63	15.31
6.	82.44	15.92
7.	86.23	17.09
8.	82.03	16.25
9.	83.89	16.48
10.	83.05	16.01
11.	74.18	46.8
12.	81.03	46.81
13.	79.66	48.27
14.	83.19	48.99
15.	80.37	49.55
16.	81.98	50.76
17.	86.69	43.08
18.	83.05	48.83
19.	84.14	46.75
20.	83.68	48.81

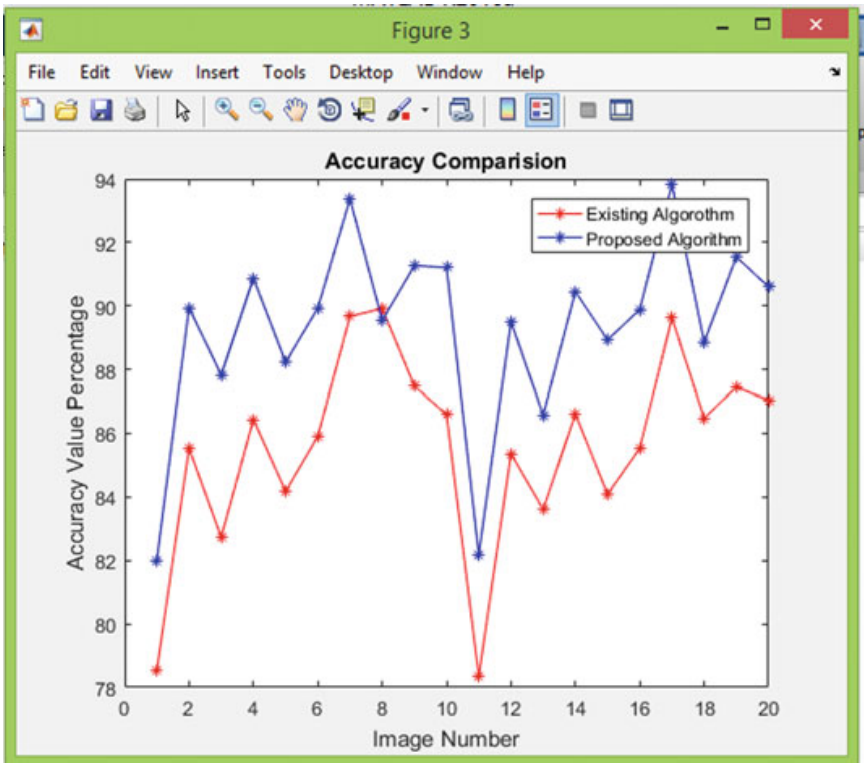


Fig. 5 Accuracy comparison

Table 3 Accuracy comparison

Image no.	Existing technique (S.V.M classifier)	Proposed technique NBC-BTD
1.	78.53	82
2.	85.53	89.94
3.	82.74	87.83
4.	86.44	90.87
5.	84.18	88.23
6.	85.91	89.93
7.	89.67	93.37
8.	89.92	89.54
9.	87.48	91.27
10.	86.59	91.21
11.	78.34	82.18
12.	85.36	89.5
13.	83.61	86.58
14.	86.61	90.45
15.	84.09	88.93
16.	85.53	89.89
17.	89.62	93.83
18.	86.48	88.86
19.	87.45	91.54
20.	87.04	90.6

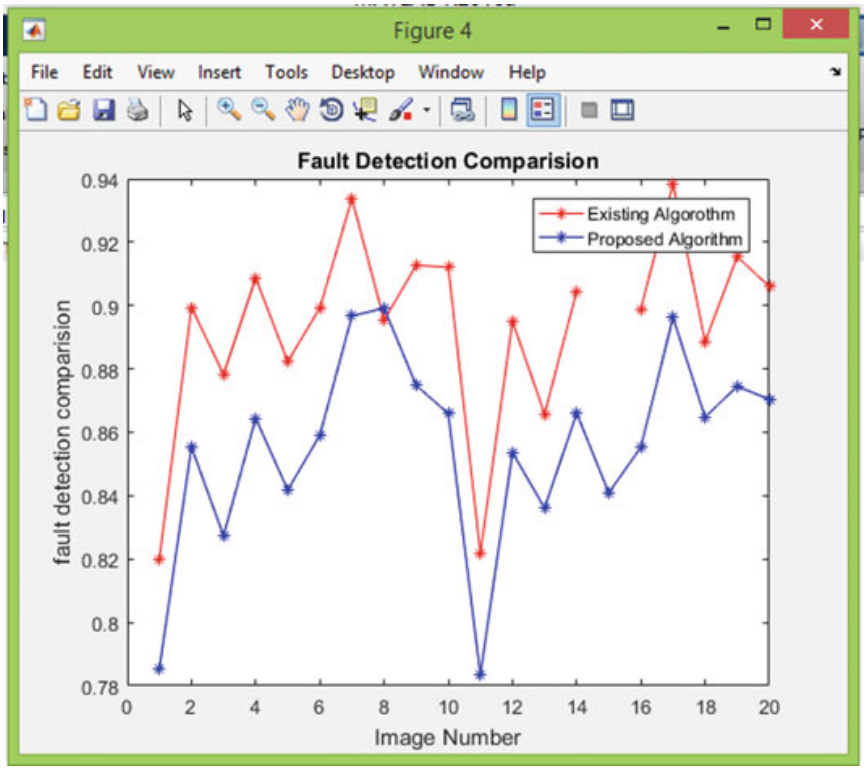


Fig. 6 Fault detection comparison

Table 4 Fault detection

Image no.	Existing technique (S.V.M classifier)	Proposed technique NBC-BTD
1.	0.7853	0.82
2.	0.8553	0.8994
3.	0.8274	0.8783
4.	0.8644	0.9087
5.	0.8418	0.8823
6.	0.8591	0.8993
7.	0.8967	0.9337
8.	0.8992	0.8954
9.	0.8748	0.9127
10.	0.8659	0.9121
11.	0.7834	0.8218
12.	0.8536	0.895
13.	0.8361	0.8658
14.	0.866	0.9045
15.	0.8409	0.8893
16.	0.8553	0.8989
17.	0.8962	0.9383
18.	0.8648	0.8886
19.	0.8745	0.9154
20.	0.8704	0.906

The fault detection rate value of the present and previous research algorithm was compared for the performance analysis. It was analyzed that present algorithm has high fault detection as compared to previous research algorithm.

4 Conclusion

In this work, it is concluded that image processing is the technique which can process information stored in the form of pixels. The brain tumor detection is the technology which can detect tumor portion from the MRI image of brain. In this research work, novel technique is proposed which is based on the morphological operation and Naïve Bayes classifier and clustering techniques. The performance of present algorithm is collate with past and it is analyzed that present algorithm performs well in terms of PSNR, MSE, and accuracy and fault detection with 86% ratio.

References

1. Dhanwani, D.C., Bartere, M.M.: Survey on various techniques of brain tumour detection from MRI images. IJCER **4**(1), 24–26 (2014)
2. Joshi, M.A., Shah, D.H.: Survey of brain tumor detection techniques through MRI images. AIJRFANS **9** (2015)
3. Poonam, J.P.: Review of image processing techniques for automatic detection of tumor in human brain. IJCSMC **2**(11), 117–122 (2013)
4. Kowear, M.K., Yadev, S.: Brain tumor detection and segmentation using histogram thresholding. Int. J. Eng. Adv. Technol. **1**, 16–20 (2012)
5. Patil, R.C., Bhalchandra, A.S.: Brain tumor extraction from MRI images Using MAT Lab. IJECSCSE **2**(1), 1 (2012)
6. Parmeshwarappa, V., Nandish, S.: A segmented morphological approach to detect tumor in brain images. IJARCSSE **4**(1), 408–412 (2014)
7. Karuna, M., Joshi, A.: Automatic detection and severity analysis of brain tumors using GUI in matlab. IJRET Int. J. Res. Eng. Technol. **2**(10), 586–594 (2013)
8. Kamil, M.Y.: Brain tumor area calculation in CT-scan image using morphological operations. IOSR J. Comput. Eng. (IOSR-JCE) **17**(2), 125–128 (2015)
9. Wan, E.W.: Neural network classification: a Bayesian interpretation. IEEE Trans. Neural Netw. **1**(4), 303–305 (1990)
10. Chew, K.M., Yong, C.Y., Sudirman, R., Wei, S.T.C.: Bio-signal processing and 2D representation for brain tumor detection using microwave signal analysis. In: IEEE (2018)
11. Kaur, N., Sharma, M.: Brain tumor detection using self-adaptive K-means clustering. In: IEEE (2018)
12. Hazra, A., Dey, A., Gupta, S.K., Ansari, M.A.: Brain tumor detection based on segmentation using MATLAB. In: IEEE (2017)
13. Chauhan, S., More, A., Uikey, R., Malviya, P., Moghe, A.: Brain tumor detection and classification in MRI images using image and data mining. In: IEEE (2017)

Automatic Classification of Carnatic Music Instruments Using MFCC and LPC

Surendra Shetty and Sarika Hegde

Abstract With a large collection of digital music in recent days, the challenge is to organize and access the music efficiently. Research in the field of Music Information Retrieval (MIR) focuses on these challenges. In this paper, we develop a system which automatically identifies the instrument in a given Carnatic music on ten different types of instruments. We extract the well-known features namely, MFCC and LPC, and analyze the capability of these features in distinguishing different instruments. Then, we apply, the classification techniques like, Artificial Neural Network, Support Vector Machine, and Bayesian classifiers on those features. We compare the performances of those algorithms along with different features for Carnatic music instruments identification.

Keywords Audio data mining · Carnatic music · Instrument classification · MIR

1 Introduction

Music Information Retrieval (MIR) research explores the different applications involving the retrieval of contents or information from a given music which can be useful for categorizing and indexing music. Due to rapid increase in digital media collections, organization and accessing of the music from large collection has become one of the challenges. Usually, metadata is used to label and segment the songs and is used to search and access a particular music. But this process needs lot of manual intervention for preparing metadata and labeling process which is not so

S. Shetty (✉)
Department of Computer Applications, NMAM Institute of Technology Nitte, Udupi District, Karnataka 574110, India
e-mail: hsshetty@nitte.edu.in

S. Hegde
Department of CSE, NMAM Institute of Technology Nitte, Udupi District, Karnataka 574110, India
e-mail: sarika.hegde@yahoo.in

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_32

efficient. Hence, this process can be automated using Machine Learning techniques including both labeling and searching. The digital music may be labeled based on instrument type, singer, genre of music, etc.

Lot of research has been done in this area on western music for content-based information retrieval. However, research on Indian music started very recently with mainly Raga identification as focused application. Instruments like Harmonium, Tabla, etc., are integral part of Indian music. The instrumental music is very melodic and pleasant, played in important Indian occasions like, marriage, festivals, funeral. The Indian instruments can be divided into three types, rhythmic instruments, wind instruments and string instruments. In our paper, we have selected the wind and string types of instruments and applied machine learning technique for automatic identification of instrument. The basic steps include preprocessing and feature extraction from audio, visualizing the features and applying classification algorithms. The set of standard audio features namely, Mel-frequency cepstral coefficient (MFCC) and Linear Predictive Coefficient (LPC) are used. We compare the capability of these features in representing an instrument. The research work in this paper explores the following objectives,

1. Extraction of MFCC and LPC features
2. Analysis of suitability of features through visualization
3. Classification of audio features using Artificial Neural Network (ANN), Bayesian classifiers, and Support Vector Machine (SVM)
4. Comparison of classifiers performance based on classification accuracy.

The content discussed in this paper is organized in five sections. We briefly discuss the literature works on MIR in second section. The methodology and the techniques used in our work are described in third section. The method of experiments and the corresponding results is covered in fourth section which is followed by conclusions in fifth section.

2 Background

Martin and Kim [13] have applied statistical pattern recognition technique for instrument identification. They have demonstrated the usefulness of organizing the musical instruments hierarchically for identification and illustrated the usefulness of timbre feature for this application. An automated system for signer identification from western music is proposed by Kim and Brian [12]. The music is first analyzed to extract the vocal segments and then the singer is identified using GMM and SVM algorithms. Similar work is proposed in Zhang [25] using a statistical model with 80% accuracy. MFCC feature with 32 coefficients is used to characterize the voice of singer and further classified using Neural network in Mesaros and Astola [15]. SVM algorithm along with LPC feature is used by Chetry and Sandier [4] for classifying six instruments successfully. Benetos et al. [3] have performed

instrument identification experiments on large dataset with 20 instruments using variety of classifiers and have reported accuracy ranging between 88.7 and 95.3%. Usefulness of MFCC coefficients for identification of instruments in monophonic signals is investigated in Sturn et al. [21]. A new nonlinear method involving the fractal theory is proposed by Zlatintsi and Maragos [26] for instrument identification. Polyphonic signals are considered in Giannoulis and Klapuri [7] using local features and missing features for instrument identification. Gaussian mixture model based semi supervised learning algorithm is applied on instrument identification by Diment et al. [5].

A continuous polyphonic music is analyzed to perform two tasks namely, automatic music transcription and instrument assignment by Giannoulis et al. [8]. Factorial Gaussian Mixture-HMM (F-GM-HMM) algorithm is used by Hg and Sreenivas [11] for solving the problem instrument detection in a given polyphonic music. Uhlich et al. [24] have used deep neural network for extracting instrument from music. Isolated Instruments identification from noisy clips has been proposed by Mukherjee et al. [16] using the statistical feature LPCC-S which is computed well-known method, Linear Predictive Cepstral Coefficient. Banerjee et al. [2] have proposed a novel approach for string instrument identification using frequency and wavelet domain analysis and the classification algorithms namely, k-NN and random forest method. A very detailed review on Music Information retrieval, its applications and the related work is described in Murthy and Koolagudi [17].

3 Methodology and Techniques

The instrument classification problem can be solved using Machine Learning algorithms. The input is an instrumental audio and the output is the label corresponding to the identified instrument. Before giving the input to Machine Learning algorithm, the input needs to be processed using Digital signal processing technique for extracting the features. In this section, we discuss briefly about the considered dataset, the feature extraction techniques and classification algorithms used in our experiment.

3.1 Dataset

There is no standard benchmark dataset for Indian instrumental music recordings. We collected few CD's consisting of music. From these songs, we extracted the first portion with duration in the range of 2 s–3 min based on the availability. The collection of recordings prepared in this manner consists of instruments namely, *Flute, Harmonium, Mandolin, Nagaswara, Santoor, Saxophone, Sitar, Shehnai, Veena and Violin.* We collected up to 15 audio clips of each instrument category with total number of audios as 150.

3.2 Audio Data Preprocessing and Feature Extraction

To extract useful information from audio which is a non-stationary data, we apply "framing". It is a mechanism of dividing the audio into smaller chunks of duration of 10–30 ms along with overlapping and windowing. Overlapping is done to ensure continuity in the data and windowing is done for smoothening the data at the edges. Each frame is then processed to extract the audio features namely, Mel-Frequency Cepstral Coefficient (MFCC), Linear Predictive Coefficients (LPC). We get a features vector of size "*d*" after applying feature extraction technique. The feature for each frame may be denoted as $\mathrm{FV} = \{\mathrm{fv}_1, \mathrm{fv}_2, \ldots, \mathrm{fv}_d\}$. The set of features extracted from all the recording is further used as input for Machine Learning algorithms.

a. *Mel Frequency Cepstral Coefficients (MFCC)*

It is a well-known feature used in the areas of speech recognition, MIR applications [9, 20]. This feature simulates the behavior of human ear where the time domain signal is first converted into frequency domain signal using discrete Fourier algorithm. The computed power spectrum from DFT coefficients are filtered using a set of triangular filters with mel-scale spacing. In the last step, inverse Fourier transform is applied by computing DCT from logarithm of Mel-spectrum given as

$$C_i = \sum_{k=1}^{K} (\log S_k) \cdot \cos\left(\frac{i\pi}{K}\left(k - \frac{1}{2}\right)\right) \quad i = 1, 2, \ldots, K \tag{1}$$

Here, the value C_i is the ith MFCC coefficient computed, S_k is the output from kth filterbank channel and K is the number of filterbanks used in the set of triangular filters [19].

Using this technique, 12 MFCC coefficients and its first/second derivative can be computed. The number of coefficients to be used is not fixed. Depending on the requirements of application, number of MFCC can be changed. Most of the time 12 MFCC's are used. Less number of MFCC makes the model simple. Aucouturier and Pachet [1] in their work have used *eight* MFCC coefficients. In our work, we have experimented with less than eight. The MFCC feature is computed using Auditory Toolbox [20]. Out of the 12 coefficients computed, we consider first *six* coefficients.

b. *Linear Predictive Coefficients (LPC)*

The time domain signal of a given audio can be analyzed using Linear predictive analysis tool to extract features from it. This tool is found to be very useful in signal processing applications. In this technique, each sample is defined with the weighted sum of previous "*d*" samples. The weights assigned to each previous sample are termed as LPC coefficient which is unknown term. The predicted value for a sample is given as [18]

$$\widehat{s(k)} = \sum_{i=1}^{d} a_i s(k-i) \tag{2}$$

Here, the $\widehat{s(k)}$ is a predicted value for kth sample, $s(k-i)$ is the sample value at index $(k-i)$. The coefficient a_i is an unknown term. We get such equations with linear combinations for all the samples in the frame where a_i's are unknown and to be computed. The constraint for computing a_i's is that the difference between the actual sample $s(k)$ and the predicted sample value $\widehat{s(k)}$ termed as Mean Squared Error (MSE) must be as minimum as possible. This is a least square problem which can be solved using normal equations.

Using this mechanism, we get "d" LPC coefficients which are equal to 13 by default. To keep the number of MFCC's and LPC's same, we consider only first *six* LPC coefficients. For a given frame f_i, the set of six LPC coefficients is represented as, $\mathrm{LP} = (\mathrm{lp}_1, \mathrm{lp}_2, \mathrm{lp}_3, \mathrm{lp}_4, \mathrm{lp}_5, \mathrm{lp}_6)$. This LP is used as the feature vector (FV) for discriminating different classes of audio signal in further experiments.

3.3 Machine Learning Algorithms

Classification is a task under Machine Learning algorithm with an objective to assign a given feature vector to one of the known class labels [6]. In our system, the given input is a feature extracted from audio and the label assigned is one to ten corresponding to the ten instruments. The dataset is first divided into training and testing set. The training set is used to construct ML model and the testing set is used to evaluate the performance of ML model. We use three algorithms, namely, Artificial Neural Network, Support Vector Machine, and Bayesian classifier which is discussed in his section.

a. *Artificial Neural Network (ANN) Classifier*

Artificial Neural Network is one of the popular Machine Learning algorithms mimics the behavior of human brain. It is a network of computing elements known as Neurons. The working of single neuron has two components, summation and activation function given as,

$$y = \frac{1}{1 + e^{-(\sum_{j=1}^{p} w_j x_j)}} \tag{3}$$

where x is an of "p" dimension and w is the weight corresponding to each input. The neurons can be arranged in layers. The three types of layers in ANN are, input layer, hidden layer, and output layer [6]. The number of hidden layers may be zero or more. The number of neurons in input layer is same as the number of feature (p) and in output layer is same number of class labels (10). The number of hidden

layers and number of neurons in hidden layer need to be decided through trial and error. In our work, we have one hidden layer with same "p" neurons. The following algorithm shows the usage of ANN in our work.

```
ALGORITHM 3.1: Training and Testing steps using ANN
// The algorithm takes dataset of feature vectors as input and generates the accuracy of classification using ANN
algorithm
//Input: D is a dataset (features vector along with target label)
//Output: m (Confusion matrix)
dataset=D;
[training_set   testing_set]=hold_out(dataset);
//Training the ANN
x=Store all the set of feature vectors of training set
y=Sequence of class labels in training set
c=no_of_class
//Each row in T represents the binary pattern of the corresponding value of y
For each row in y
    Initialize the T(i) to 'c' number of zeros
    Set the value of T(y(i))=1
End
network=trainANN(x,T);
//Testing the ANN
For each row from testing set
        FV=feature vector
        output_array=testANN(network, FV); / / list of values from Output neuron
        pos=index of maximum value in output_array
        predicted_target_label=pos;
        m(actual_target_label, predicted_target_label)= m(actual_target_label, predicted_target_label)+1
end
```

b. *Support Vector Machine (SVM)*

This is a supervised machine learning algorithm which constructs a hyperplane that can separate the feature vector of one class with feature vectors of all other classes. Out of the different possibilities of hyperplane, the one with widest margin is selected as classifier called Maximal Margin Classifier [23]. Let us consider a set of training observations, given as $(\boldsymbol{x_i}, y_i)$ $(i = 1, 2, \ldots, k)$ where $\boldsymbol{x_i} = \left(x_{i1}, x_{i2}, \ldots, x_{ip}\right)^{\mathrm{T}}$ corresponds to the feature vector (MFCC or LPC) for the ith input observation and $y_i \in \{-1, 1\}$ denote its class label. SVM is a two-class classifier which constructs a hyperplane separating the observations from two class. In our paper, *ten* hyperplanes will be constructed. For each instrument, a constructed hyperplane will separate the observations of that instrument from all other instruments [10]. The decision boundary (hyperplane) can be defined as follows,

$$\boldsymbol{w} \cdot \boldsymbol{x} + b = 0$$

The parameters **w** and b need to be estimated in training phase for a given dataset. Since the hyperplane need to maximal, the constraint and objective function to be minimized for computing the parameters is given as,

$$\begin{aligned} \boldsymbol{w} \cdot \boldsymbol{x} + b \geq 1 \quad \text{if } y_i = 1 \\ \boldsymbol{w} \cdot \boldsymbol{x} + b \leq -1 \quad \text{if } y_i = -1 \\ f(w) = \frac{\|\boldsymbol{w}\|^2}{2} \end{aligned} \tag{4}$$

The following algorithm illustrates the steps followed for using SVM algorithm in our work.

ALGORITHM 3.2: Pseudo code for SVM algorithm

```
// The algorithm takes dataset of feature vectors as input and computes the accuracy of classification
using SVM algorithm
//Input: D is dataset (feature vector along with target label)
//Output: m (Confusion matrix)
dataset=D;
[training_set   testing_set]=hold_out(dataset);
//Training the SVM
x=Store all the set of feature vectors of training set
y=Sequence of class labels in training set
c=no_of_class
Apply Gaussian kernel function to x
[xsupport, w, b]=trainSVM(x,y);
//Testing the SVM
For each row from testing set
        FV=feature vector
        predicted_target_label =testSVM(FV, xsupport,w,b);
        m(actual_target_label, predicted_target_label)= m(actual_target_label,
predicted_target_label)+1
end
```

c. *Bayesian Classifier*

Bayesian classifier works based on conditional probability. Posterior probability using Bayes theorem is computed for an feature vector x given as,

$$p(c_j|x) = \frac{p(c_j)p(x|c_j)}{p(x)} \tag{5}$$

where

$$p(x) = \sum_{j=1}^{c} p(w_j)p(x|w_j) \tag{6}$$

The posterior probability $p(c_j|x)$ indicates the probability that feature vector x belongs to jth class where $j = 1, 2, \ldots, 10$ in our work. To compute the class label for a given x, all the 10 posterior probabilities are evaluated and the class corresponding to the highest probability is chosen [22]. The prior probability for class j, i.e., $p(c_j)$, represents the initial degree of belief that a given set of feature vector is a

data from jth class. The class-conditional probability $p(x|c_j)$ is also called as class-conditional probability representing the probability distribution of the features for each class [14].

4 Experimental Results and Discussion

Each audio from the collected dataset is retrieved and processed to extract MFCC and LPC features. The feature vector corresponding to each frame is considered as an observation (an input) for the next stage. Total number of observations for each class is equivalent to number of audios of that class multiplied by number of frames in the audio. We get up to 600–1000 observations for each class. We have considered four different frame size, 512, 1024, 2048, and 4096 for the experiment. For the final classification, we selected the best frame size among this.

In this section, we discuss three types of experiments and the corresponding results. The first one is visualization of features, second is to select the best frame size, and the third one is the classification of instruments. For visualization on 3-D plot, we select the first three coefficients $(fv_{i1}, fv_{i2}, fv_{i3})$ of MFCC and LPC. Randomly selected 25 observations are plotted on 3-D graph with different color/symbols for different instruments. Figures 1 and 2 show the plot for MFCC and LPC features respectively. It can be clearly observed that the instruments of same type are grouped together in both the plots. We can also observe that similar types of instruments like *Saxophone(inverted triangle)* and *Nagaswara(circle)* groups are nearer to each other.

This result of this experiments just gives a brief idea about the potentiality of a feature to distinguish the various instruments and the similarity among some instruments. As per the results, both MFCC and LPC features capable to that. Such experiments are useful in Machine learning applications as a preprocessing step to

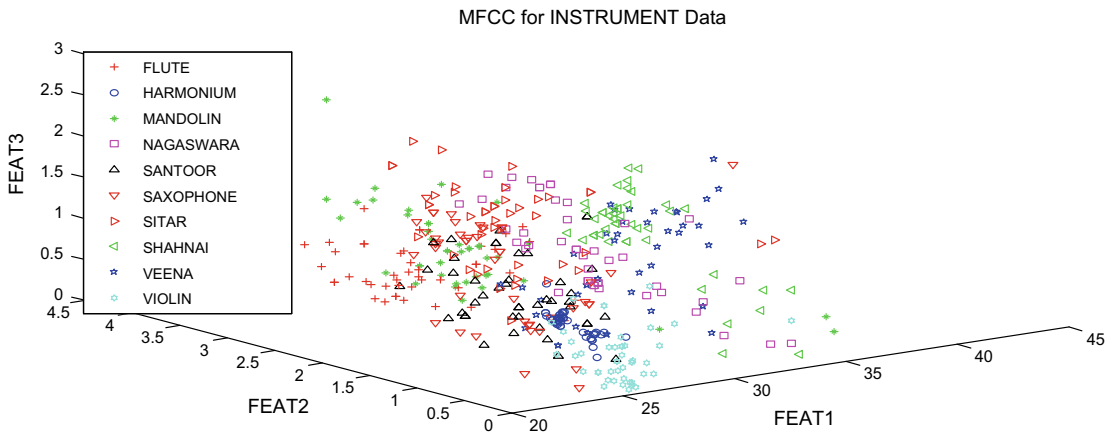


Fig. 1 Visualization of MFCC feature for Instrument data

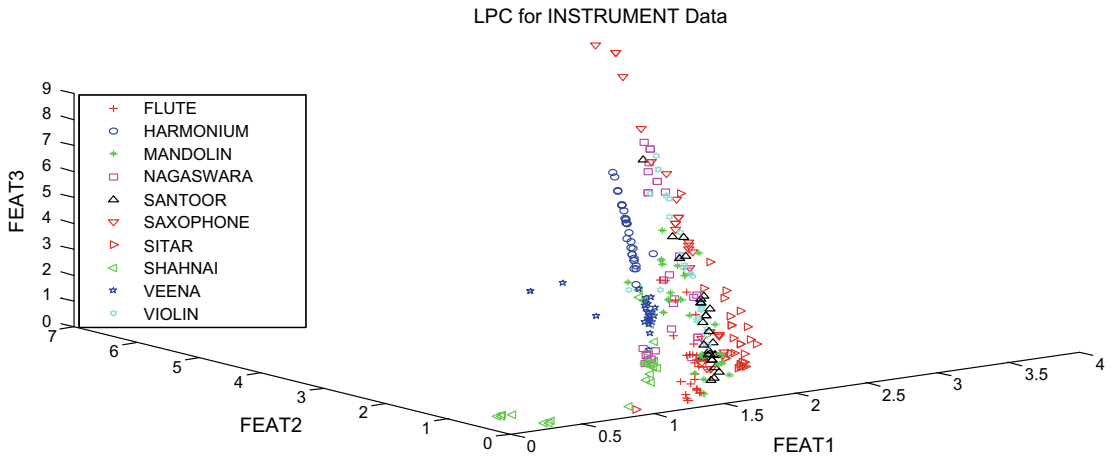


Fig. 2 Visualization of LPC feature for Instrument data

understand the features at preliminary level. The discrimination of instruments through MFCC is slightly clearer than LPC features.

The next experiment done is to select the best frame size suitable for instruments classification. For this experiment, we randomly retrieve *50* observations from each class and apply all the classification techniques using twofold cross-validation evaluation method. Various frame sizes considered are *512, 1024, 2048 and 4096* sample per frame. The features extracted with various frame size are stored as separate datasets. The frame size which gives the highest classification accuracy is selected and considered for further evaluation of classification experiments. Table 1 shows the results of varying frame size across *two* features and the different classification algorithms for the instrument dataset.

We can observe from the table that the performance of classification is better for the dataset constructed with the frame size of *2048* sample observations. For the further analysis, we consider the dataset which is constructed with the frame size 2048.

The last experiment is classification of ten instruments using, ANN, SVM, and Bayesian classifier. For this, we retrieve *500* observations from each class of instruments, totally constituting *5000* observations from *ten* classes. The dataset is divided into two sets, set1 and set2 in the ratio of *50:50* randomly. In the first iteration, set1 is used as training set and set2 is used as the testing set and the

Table 1 Comparison of class accuracy for different frame sizes for Instrument data

Classifier		512	1024	2048	4096
ANN	MFCC	56%	31%	72%	55%
	LPC	25%	61%	61%	66%
SVM	MFCC	88%	82%	86%	85%
	LPC	75%	77%	76%	70%

Table 2 Comparison of classification accuracy for Instrument data

Classifier/feature	ANN (%)	SVM (%)	Bayesian (%)
1. MFCC	**83.07**	69.13	67.12
2. LPC	81.81	67.91	51.25

accuracy of classification is calculated for the test set. Same thing is repeated in the second iteration where set2 is used as the training set and set1 is used as the testing set. Finally, the average accuracy of classification of the two iterations is calculated. Classifiers are evaluated by repeating each of the experiments for *fifteen* times and the average accuracy of classification is taken as the final accuracy of classification. Table 2, shows the accuracy of classification for the dataset of *5000* observations with MFCC and LPC features and the ANN, SVM and Bayesian classifier algorithms. We can observe from the results that MFCC feature is better representative for instrument data and the ANN algorithm classification is better compared to other algorithms. But LPC feature is also useful in representing the instrument data.

The highest accuracy rate for 500 observations is 83.07% with MFCC feature and ANN classifier and Fig. 3 shows the graphical view of the result. The *x* axis gives the name of the instrument (actual target class) and the *y* axis gives the number of observations considered. In visualization, we claimed that MFCC discriminates instrument data better than LPC. This is found to be same in classification also. Even though there is not much difference in the accuracy, MFCC has shown slightly better performance than LPC in all the cases.

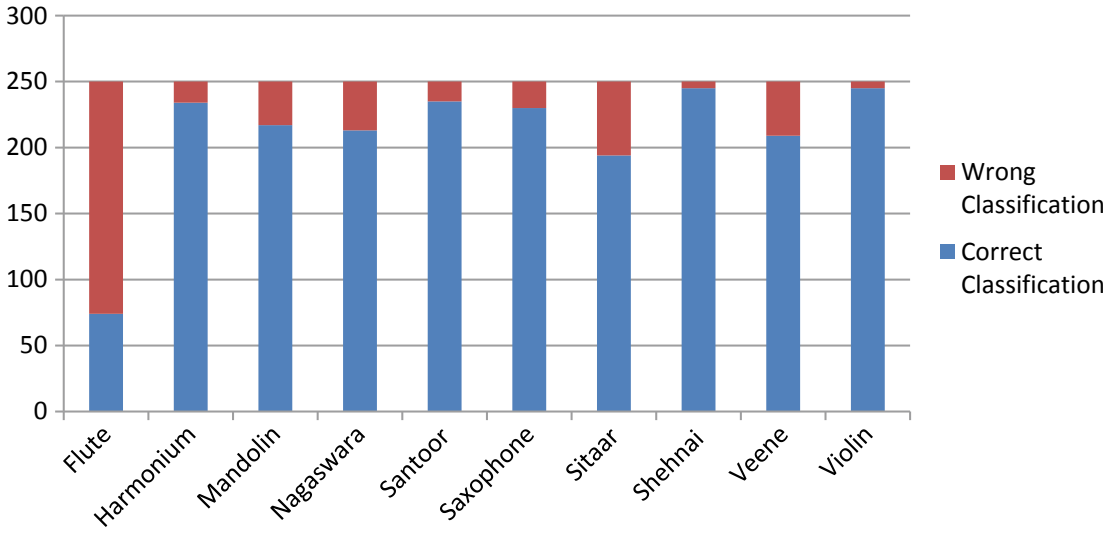


Fig. 3 Graphical representation of confusion matrix for Instrument data classification with ANN and MFCC feature

5 Conclusions

Even though, lot of research is carried for western music, the research in Indian music has lot of scope due to versatile nature of Indian music. Instruments are integral part of Indian music. The automatic identification of instruments from Carnatic music can ease the process of indexing and retrieval. We have developed a system for instrument identification using MCC and LPC features. The visualization helped in understanding the capability of MFCC and LPC features in distinguishing the various instruments. The training dataset and test dataset are selected randomly in each of the iteration and it is found that MFCC feature is shown to be more discriminative than other features. The highest rate of classification accuracy for *10* classes of instrument data is 83% with ANN classifier. More variety of instruments may be considered in future to develop an instrument identification system using deep neural network.

References

1. Aucouturier, J.J., Pachet, F.: Improving timbre similarity: how high's the sky? J. Negat. Results Speech Audio Sci. **1**(1), 1–13 (2004)
2. Banerjee, A., Ghosh, A., Palit, S., Ballester, M.A.F.: A novel approach to string instrument recognition. In: International Conference on Image and Signal Processing, pp. 165–175. Springer, Cham (2018)
3. Benetos, E., Kotti, M., Kotropoulos, C.: Large scale musical instrument identification. In: Proceedings SMC'07, 4th Sound and Music Computing Conference, Greece (2007)
4. Chetry, N., Sandier, M.: Linear predictive model for musical instrument detection. In: IEEE Conference on Acoustics, Speech and Signal Processing, vol. 5 (2006)
5. Diment, A., Heittola, T., Virtanen, T.: Semi-supervised learning for musical instrument recognition. In: 2013 Proceedings of the 21st European Signal Processing Conference (EUSIPCO), pp. 1–5. IEEE (2013)
6. Duda, R.O., Hart, P.E., Stork, D.G.: Pattern Classification. Wiley Publications, Hoboken (2006)
7. Giannoulis, D., Klapuri, A.: Musical instrument recognition in polyphonic audio using missing feature approach. IEEE Trans. Audio Speech Lang. Process. **21**(9), 1805–1817 (2013)
8. Giannoulis, D., Benetos, E., Klapuri, A., Plumbley, M.D.: Improving instrument recognition in polyphonic music through system integration. In: 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 5222–5226. IEEE (2014)
9. Gold, B., Morgan, N.: Speech and Audio Signal Processing—Processing and Perception of Speech and Music. Wiley India Pvt. Ltd., ISBN: 81–265-0822-1 (2006)
10. He, X., Zhou, X.: Audio classification by hybrid support vector machine/hidden Markov model. UK World J. Model. Simul. **1**(1), 56–59 (2005)
11. Hg, R., Sreenivas, T.V.: Multi-instrument detection in polyphonic music using Gaussian Mixture based factorial HMM. In: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 191–195. IEEE (2015)
12. Kim, Y.K., Brian, Y.: Singer identification in popular music recordings using voice coding features. In: Proceedings of ISMIR (2002)

13. Martin, K.D., Kim, Y.E.: 2pMU9: musical instrument identification: a pattern-recognition approach. In: 136th Meeting of the ASA (1998)
14. Martinez, W.L., Martinez, A.R.: Computational Statistics Handbook with Matlab. Chapman & Hall/CRC Publications, ISBN: 1-58488-566-1 (2008)
15. Mesaros, A., Astola, J.: The mel-frequency cepstral coefficients in the context of singer identification. In: Proceedings of the International Conference on Music Information Retrieval (2005)
16. Mukherjee, H., Obaidullah, S.M., Phadikar, S., Roy, K.: MISNA—a musical instrument segregation system from noisy audio with LPCC-S features and extreme learning. Multimed. Tools Appl. **77**, 27997–28022 (2018)
17. Murthy, Y.V., Koolagudi, S.G.: Content-based music information retrieval (CB-MIR) and its applications toward the music industry: a review. ACM Comput. Surv. (CSUR) **51**(3), 45 (2018)
18. Peltonen, V., Tuomi, J., Klapuri, A, Huopaniemi, J., Sorsa, T.: Computational Auditory Scene Recognition. In: IEEE International Conference on Acoustics Speech and Signal Processing, vol. 2 (2002)
19. Rabiner, L., Jaung, B.: Fundamentals of Speech Recognition. Prentice Hall, ISBN:10:013051572 (1993)
20. Slaney, M.: Auditory toolbox: a MATLAB Toolbox for auditory modeling work. Technical Report 1998-010, Interval Research Corporation, Palo Alto, CA, USA, 1998, Version 2 (1998)
21. Sturm, B.L., Morvidone, M., Daudet, L.: Musical instrument identification using multiscale mel-frequency cepstral coefficients. In: 18th European Signal Processing Conference (EUSIPCO-2010) (2010)
22. Tan, P., Steinbach, M., Kumar, V.: Introduction to Data Mining. Pearson Addison Wesley, ISBN: 978-81-317-1472-0 (2006)
23. Theodoridis, S., Koutroumbas, K.: Pattern Recognition, 4th edn. Academic Press, London (2009)
24. Uhlich, S., Giron, F., Mitsufuji, Y.: Deep neural network based instrument extraction from music. In: ICASSP, pp. 2135–2139 (2015)
25. Zhang, T.: System and method for automatic singer identification. In: Proceedings of the IEEE Conference on Multimedia and Expo, vol. 1, pp. 33–36 (2003)
26. Zlatintsi, A., Maragos, P.: Multiscale fractal analysis of musical instrument signals with application to recognition. IEEE Trans. Audio Speech Lang. Process. **21**(4), 737–748 (2013)

Semiautomated Ontology Learning to Provide Domain-Specific Knowledge Search in Marathi Language

Neelam Chandolikar, Pushkar Joglekar, Shivjeet Bhosale, Dipali Peddawad, Rajesh Jalnekar and Swati Shilaskar

Abstract In this research work, our goal is to build a self-sustainable, reproducible, and extensive domain-specific ontology for the purposes of creating a knowledge search engine. We have used online data as the primary information store using which we construct ontology by identifying concepts (nodes) and relationships between concepts. The project encompasses preestablished ideas gathered from successful NLP trials and presents a new variation to the task of ontology creation. The system, for which the ontology is being created, is a knowledge search engine in Marathi. This aims at building semiautomated ontology whose target demographic is primary school children and the selected domain is science domain. This project proposes a method to build semiautomated ontology. We use a combination of natural language processing method and machine learning method to automate the ontology learning task. Automatically learned ontology is further modified by language and domain experts to enrich the contents of ontology. Unlike, standard search engines, our knowledge search engine attempts to provide learned resources directly to the user rather than website links. This approach enables the user to directly get information without having to spend time on browsing indexed links.

N. Chandolikar (✉)
Department of IT and MCA, Vishwakarma Institute of Technology, Pune, India
e-mail: neelam.chandolikar@vit.edu

P. Joglekar · S. Bhosale · D. Peddawad
Department of Computer Science, Vishwakarma Institute of Technology, Pune, India
e-mail: pushkar.joglekar@vit.edu

S. Bhosale
e-mail: shivjeet.bhosale14@vit.edu

D. Peddawad
e-mail: dipali.peddawad16@vit.edu

R. Jalnekar · S. Shilaskar
Department of E&TC, Vishwakarma Institute of Technology, Pune, India
e-mail: rajesh.jalnekar@vit.edu

S. Shilaskar
e-mail: swati.shilaskar@vit.edu

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_33

Keywords Knowledge search engine · Natural language processing (NLP) · Ontology · Resource description framework (RDF) · Semantic search

1 Introduction

Knowledge sources are rarely available in Marathi language on World Wide Web. This paper proposes to provide domain-specific search facility in Marathi. It also aims at converting the extracted data into knowledge and provides output in more semantically coherent form. The proposed knowledge search in Marathi would be based on semiautomated ontology-based framework.

1.1 Ontology

Ontology has the ability to capture the structure of a domain. Ontology has a knowledge base which contains structures that are defined by the allowed constructs of ontology [1, 2]. Ontologies are building blocks of semantic web-based systems [3]. Creating ontologies is an intricate process [4, 5]. Ontology defines a common vocabulary for researchers who need to share information in a specific domain. Ontology of any domain is a specification of a conceptualization of that specific domain. Ontologies are concerned with the meaning of terms. Ontology typically consists of concepts and relationship between those concepts. It includes definitions of basic concepts in the specific domain and relations among them. Ontologies are used to represent domain knowledge in a digital form. Ontologies can be automated or semiautomated. Semiautomated ontology is created automatically and edited manually.

1.2 Knowledge Search Engine

Knowledge search engine [6–8] provides semantically correlated information to the input keyword. Knowledge search engine is somewhat similar to semantic search engine. A semantic search engine interprets the meaning of a given keyword and provides information either in the form of relevant links or in the form of summarized information. Ontologies allow smart linking of knowledge resources. This gives a computer the potential to use intelligent agents to access knowledge and deliver good quality information to the user. Agents also have the ability to use the ontology semantics [9, 10] to communicate with other agents and thus making computer communications smarter.

1.3 Resource Description Framework (RDF) and Web Ontologies

Ontologies are stored in RDF format [11]. RDF is the widely accepted standard framework for representing information about a certain system. RFD is built around the concept of uniform resource identifier (URI) which is used to describe website metadata. A typical RDF graph consists of RDF triples (subject, predicate, and object) which is displayed as a graph. Web ontology complements the RDF and RDF Schema by giving it an ability to express reasoning and convey logical structure.

2 Related Work

In 2007, author Horacio Saggion et al. proposed the system on ontology-based information extraction for business intelligence in which they used information from business graphics and tabular data and they applied OCR analysis of images and then output of the OCR analysis was corrected by the exploitation of collateral information found around the graphics [12]. In 1992, author Marti A. Hearst described a method for the automatic acquisition of the hyponymy lexical relation from large text corpora first they identified a set of lexico-syntactic patterns and then described a method for discovering these patterns [13]. In 2006, author Patrick Pantel et al. proposed a weakly-supervised, general-purpose algorithm, called Espresso, for harvesting binary semantic relations from raw text. In which they have extracted several standard and specific semantic relations like is a, part of, succession, reaction, and production [14].

In 2004, author Patrick Pantel et al. used top-down approach for labeling semantic classes. They used co-occurrence statistics of semantic classes discovered by algorithms like CBC to label concepts. In this paper, they proposed an algorithm for automatically inducing names for semantic classes and for finding instance/concept (is-a) relationships [15]. In 2013, author Tomas Mikolov proposed two novel model architectures for computing continuous vector representations of words from very large datasets. The author worked on distributed representations of words learned by neural networks in which they proposed two models Continuous Bag-of-Words Model and Continuous Skip-gram Model [16]. Natural language processing methods [17– 20] of pos tagging and chunking is widely used by many researchers for identification of concepts for domain-specific ontology. Machine learning techniques like classification [21–24] is used widely for relationship classification.

3 Proposed Methodology

In the proposed methodology, we have focused on creating ontology for a science domain which includes science subject terms and concepts understandable to primary school-going children of Marathi medium. However, proposed methodology can be applied to any domain. In the proposed methodology we at first define ontology structure followed by this ontology learning method is proposed and finally we provide knowledge search through knowledge search engine.

3.1 Defining Structure of Ontology

For building ontology and its nodes, we have initially defined ontology structure as per the requirement of project. Ontology ideally has terms and relationships. In this research work for every term, one ontology node is created. Structure of ontology node which is used is as follows:

1. Keyword in English.
2. Definition in English.
3. Automated Translated Marathi keyword.
4. Marathi definition by domain and language expert.
5. Image link.

Once this ontology node structure is defined, all the terms (concepts) are stored using the same node structure.

3.2 Methodology to Create Semiautomated Ontology

We have presented a combined approach to natural language processing and machine learning classification for ontology learning. Ontology learning has six individual modules, namely, seed word resolution, data scraping, pos tagging and chunking, predicate classification for identification of relationship, ontology writer and graph intelligence, and text simplification. Figure 1 shows the block diagram of ontology creation. The detailed explanation of the ontology creation are as follows.

3.2.1 Seed Word Resolution

The semantic search engine takes an input seed word for learning ontology. The search engine operates on one seed word at a time so as to limit the divergence of

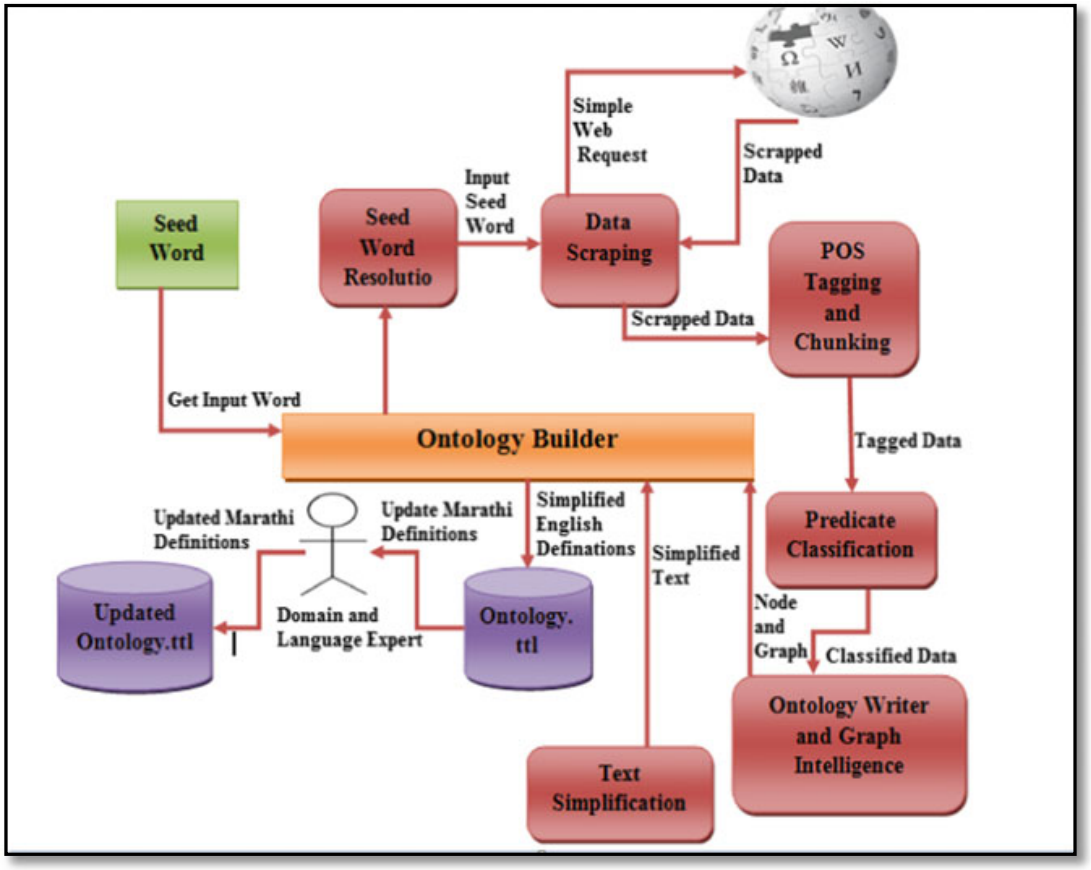


Fig. 1 Block diagram of ontology builder

the data that is gathered. Traditional encyclopedias such as Wikipedia or Britannica are replete with example of entries with topics that can be interpreted in a much broader sense. Our algorithm, in order to avoid this problem, gets additional five words that are semantically similar to the seed word. This is done using Word2vec. Word2Vec is a vector embedding space that encodes semantically similar words in an N-dimensional vector space such that their cosine similarity is minimum. We used Google's dataset which is trained on roughly 100 billion words with a vocabulary of 3 million words. The top five words similar to the seed word make sure that the algorithm will always have relevant information when displaying to the user. However, to be able to know the meaning of millions of words there is no practical classifier that will give us good description of any word. A dictionary can be used instead but it does not capture the linking between two or more terms. To be able to do this, our algorithm treats Word2vec model as the classifier for the word. This is the prime reason behind the extraction of top five similar words. They act as an index as to what words will point wherein the ontology.

3.2.2 Web Data Scraping

Web Data Scraping is a computer science application, where a programmer automates the process of retrieval of information from websites by the use of intelligent scripts. Every website typically has an HTML format which tells the browser how the raw data must look to the user. Now, say a user wants to analyze product prices on a web store, manual extraction is impractical. In such cases, a simple script can be written to talk to the website's server and retrieve the required data directly into user-defined data structures. Our search engine relies heavily on web scraping to eliminate the need for storing data offline. The top five similar words to the seed word will be the basis of getting raw text data. The algorithm will first scrape data for each word from https://simple.wikipedia.org/, which is a simplified version of Wikipedia. At this stage, the algorithm has all the relevant data for each wiki pages for each of the five words and the seed word. The scraper will extract definitions, images, (links) and other paragraphs from each web page. The scraped data objects will be provided to a module that performs a series of operations to further refine the data.

3.2.3 Part of Speech (POS) Tagging

Every word used in a sentence is assigned a specific part of speech. There are several parts of speech and the most important and commonly recognized ones are nouns, verbs, and adjectives. We have to dive a little deeper into the grammar of English language to understand the importance of this module.

A sentence in English falls into three categories, namely, simple, compound, and complex. A machine, however, does not understand these categories. To allow the computer to understand this, we exploit the patterns that the POS tags that appear in all three kinds of sentences. We use NLTK's inbuilt POS tagger for this purpose.

NLTK's POS tagger works on a saved per-trained POS tagging machine learning models. The NLTK POS tagger recognizes the various types of POS tags.

3.2.4 Chunker

The next phase that comes is chunking. Chunking is the process of converting a POS tagged sentence into a list of tree structures that recognizes a certain user-defined grammar. The root of all trees are defined by the letter 'S' and in the absence of grammar, all the words are leaves of this tree. The chunking process, allows us to recognize the subject, predicate, and the object of the sentence. NTLK's Chunker is a regular expression-based function that parses the sentences into the tree from the user-defined grammar. Our algorithm attempts to recognize

the subject, predicate, and object of every sentence. To do this, we defined our own grammar rules to recognize the subject, predicate, and object in a sentence. The grammar used is as follows.

```
“”””
        NP :    {<PRP|PRP$>?<DT>?<JJ>*<NN.*>+}
                {<JJ>?}
                {<PRP|PRP$>}
        REL : {<V.*><TO>?}
                {<DT>?<JJ>*<NN.*>+}
                }<VBZ|VBP>{
        REL_BOOL : {<RB|RBR|RBS>?<VBZ|VBP*>?<RB.>?<TO>?<IN>?}
                Commas : {<,><NP>}
                NPlist : {<NP><Commas>*<CC>?<NP>}
        SVO: {<NP|NPlist><MD>?<REL_BOOL>+<REL>*<IN>*<CD>*<NP|NPlist>}
        Prep : {<IN>+<N.|VBG|PRP.>}
        Simple: {<SVO><Prep>*}
        tag :   {<CC><Simple>}
        Compound: {<Simple><tag>+}
“”””
```

In the above-defined grammar, “NP” is the grammar used to chunk noun phrases, “REL” and “REL_BOOL” are used to chunk verb phrases from sentences. When the chunker spots the grammar defined by “SVO” tag, it will indicate that the algorithm has now safely found an SPO triple. This triple will be further analyzed to generate the ontology.

3.2.5 Concept Identification

After POS tagging and chunking phase, subjects and objects are identified as concept (term) of ontology. After identification of terms, for every term following details are stored in the node of ontology: keyword in English, definition in English, automated translated Marathi keyword, and Marathi definition by domain and language expert and Image link for term.

3.2.6 Predicate Classification

As humans, understanding the meaning of any sentence is very easy. When trying to replicate the same task for computers, the biggest challenge is where to start in order to solve this problem. Ideally one would think that there is no problem in storing the relationship directly as it would make any difference when it is presented to the user. However, the opposite is true for machines. Machines have not been taught to recognize the relationship between two or more entities. When dealing with computers, ultimately, it comes down to etching strict borders to these abstract tasks that humans perform. Ideally, a human will be able to understand if two entities are related, but to know how they are related is quite a huge and complicated task. For instance, a human can easily declare that the word "dog" and the word "golden retriever" are related. As humans, we can relate "dog" and "golden retriever" as class and instance of the class, respectively. However, it is important to note that we can do this without the presence of a predicate. Humans can also express more than one relation between the two entities. This is unique only to humans.

In order to replicate the same result in machine, we first need to define how the relationships are to be specified. This is where the ontology comes into the view. As we discussed in the initial section of this paper, ontology is a specification for conceptualization. The exact syntax and rules that are specified in RDF will form the basis for recording the knowledge in the ontology. For our purposes, we have tested four classifiers, namely, Decision Tree, SVM, Naïve Bayesian, and Maximum Entropy.

In this paper, we propose the use of maximum entropy techniques for ontology building for knowledge search. Experimental results show that max entropy gives more accurate result. Maximum entropy is a method used for a natural language processing task like text segmentation. The first step for constructing model using max entropy is to collect a large number of training data which consists of samples represented in the following format: (x_i, y_i) where the x_i includes the contextual information of the document (the sparse array) and y_i its class. The second step is to summarize the training sample in terms of its empirical probability distribution:

$$\breve{p}(x, y) = \frac{1}{N} \times T$$

where T is the number of times that (x, y) occurs in the sample and N is the size of the training dataset.

3.2.7 Graph Intelligence

Simply storing the data in the ontology is not very useful if the user has to link the data using their own intelligence. Our graph intelligence module generates a graph that records all the knowledge in terms of subject, predicate, and object represented

as the nodes. The graph intelligence automatically integrates definitions and images for the nodes from the ontology.

After the machine has recognized the word, the next step of the algorithm is to describe the word. To describe a word, we as humans have a brain wiring that allows us to record the features of the word. For example, we can say that to describe a dog, we humans will recall the image of the dog and then describe the features of the dog that we imagine. A particular node in the ontology is the feature of that word that the algorithm has learned and stored by analyzing all the raw text data. Over time all these nodes automatically form connections and which the algorithm can describe. The algorithm allows the user to traverse along the edges of the ontology nodes thus simulating a primitive thought process. This graph is well connected to the other related entities but with a few erroneous relationships. The graph currently traverses two levels deep. The graph can be visualized as follows (Fig. 2).

3.2.8 Text Simplification

Text simplification [25] is a technique in natural language processing which is used to modify text in such a way, that the grammar and structure of the given text is

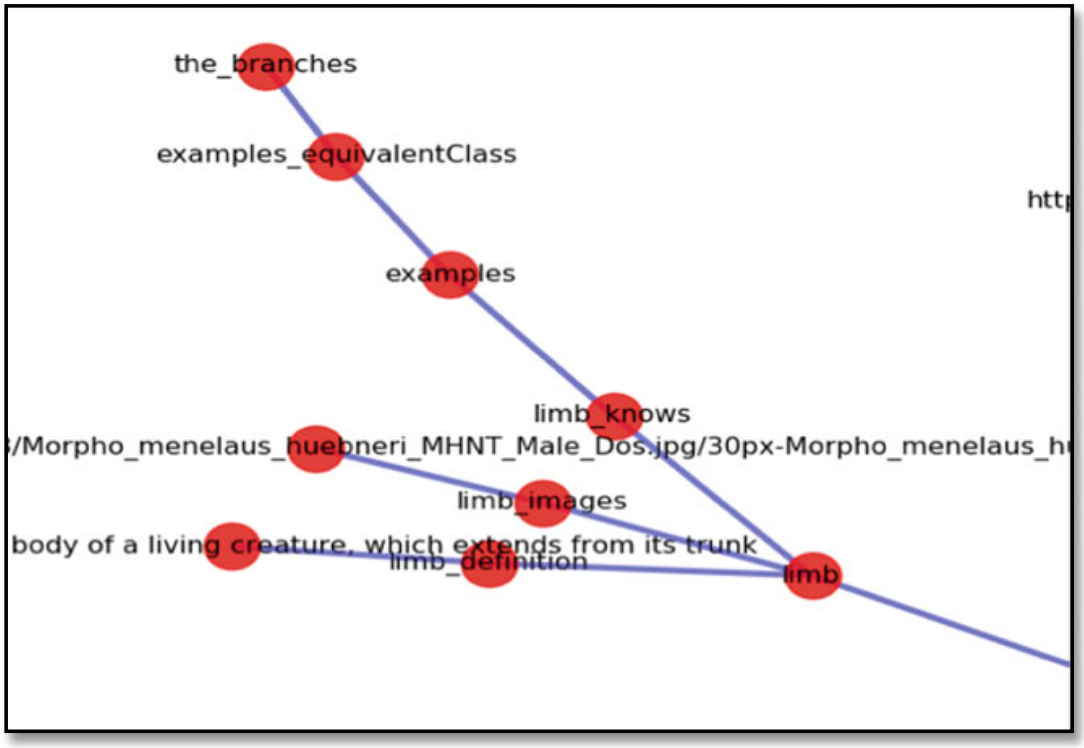


Fig. 2 A part of the subgraph, partly describing the node "limb"

greatly simplified, while the meaning and information remain the same. The target demographic of the knowledge search engine is school-going children so; it is necessary to create definitions which will be understandable by children. In-text simplification we converted complex words to easily understandable words and complex sentences to simple sentences. In this work English definition which is retrieved from World Wide Web and stored in ontology node are further simplified, so that it becomes understandable to target demographics. Further, Marathi definitions are added by domain and Marathi language expert, to complete ontology node as per defined structure.

3.3 Knowledge Search Engine

In the second part, we created a web application using Django framework. This application will help student to search the science terms. This application will give direct results to the students instead of giving a link as a result. So, it will help student to get the result without wasting time on indexed link. Figure 3 shows a block diagram for ontology search.

We will enter a science keyword in Marathi as input. Then that Marathi word is converted to English word then semantic search engine will search that keyword in updated ontology file and will return output definitions and images.

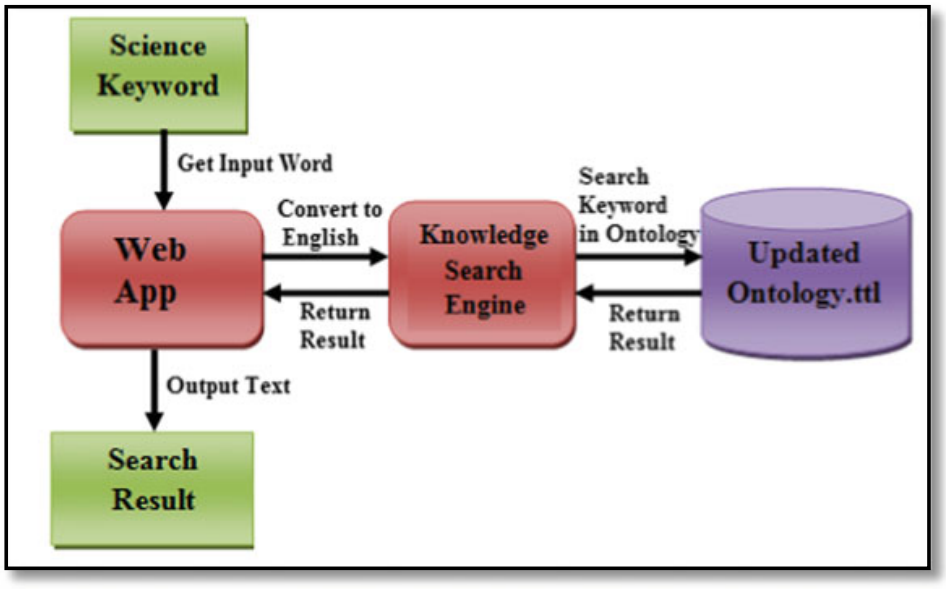


Fig. 3 Block diagram of ontology search

4 Experimental Result

During the course of the project, it has been found out that to generate good quality ontology; the most significant task identifies relevant triples. A relevant triple is a triple in which the subject, predicate, and object talk about the seed word. Current chunking methods perform fairly well for simple and complex sentences.

4.1 Comparative Analysis of Classifiers

The initial approach was to use the Naive Bayes classifier to classify the predicate. In later comparative analysis of the dataset, the following were the accuracies for each classifier (Table 1).

4.2 Classification of Predicate

The predicate classification task requires a huge amount of data along with good labels. The predicate classification works fairly well on maximum entropy classifier with an average accuracy of 76%. The feature of the predicate can be a dictionary of POS tags as the keys and the corresponding word in the predicate as the value of the dictionary.

Maximum entropy classifier gives the highest accuracy. Therefore, this method is used for building ontology nodes and relationships. Relationships are mainly classified into subclass, equivalent, object property, and irrelevant. The relationships that are classified as irrelevant are of varying nature. They do not directly fall into subclass, equivalent, or object property relations but they usually are not completely useless. They are stored as a friendly relation.

Table 1 Comparative analysis of classifiers on our dataset

Classifier	Accuracy (avg of 20 shuffled iterations) (%)
Naive Bayes classifier	72.42
Maximum entropy classifier	76.39
Support vector machine (SVM)	75.15
Decision tree	73.18

4.3 Created Ontology

The proposed method provides a good mechanism to generate domain-specific ontology. Given a seed word, relevant terms are added into ontology. Created ontology has terms and relationships. For every term one ontology node is created. Following examples illustrates ontology node and ontology structure.

(a) Example of ontology nodes

1. Keyword in English:- "Sunlight".
2. Definition in English (simplified definition):- "It is the light of the sun. Sunlight, also called sunshine. What we experience as sunlight is actually solar radiation. It is the radiation and heat from the Sun in the form of waves. The atmosphere affects the amount of solar radiation received. When solar radiation travels through the atmosphere, some of it is soaked up by the atmosphere (16%)".
3. Marathi definition (provided by language and domain expert):- "सूर्यप्रकाश म्हणजे सूर्याचा प्रकाश. आपण जो सूर्यप्रकाश अनुभवतो, ती खरे म्हणजे सौर प्रारणे आहेत. सूर्यापासून नघिालेली सौर प्रारणे आर्णि उष्णता तरंगांच्या रूपात असतात. सौर प्रारणांमुळे वातावरणावर परणिाम होतो. सौर प्रारणे जेव्हा वातावरणातून जातात तेव्हा त्यातील काही प्रारणे (सुमारे 16%) वातावरणात शोषली जातात".

3. Image link: https://upload.wikimedia.org/wikipedia/commons/thumb/7/71/Antelopecanyonjh1.jpg/220px-Antelope_canyon_jh1.jpg.

5 Conclusion

In the context of the current project, we are planning to use the ontology for knowledge search engine where users are school-going children. The semantic search engine is mainly focused on science keywords which will provide a Marathi definition of a keyword as output. Creating ontologies is not an easy task and there is no unique correct ontology for any subject. So, in this project, we used a combined approach to natural language processing and machine learning method for creating ontology. The predicate classification task requires a huge amount of data along with good labels. The predicate classification works fairly well on maximum entropy classifier with an average accuracy of 76%. The maximum entropy classifier is the most suitable for our ontology creation. This paper proposes the building of ontology which provides knowledge search in Marathi, which takes a keyword as an input and uses ontology to provide semantically coherent information and present summarized information in the structured form.

Acknowledgements This work was supported by the Maharashtra Government Project funded by the Rajiv Gandhi Science & Technology Commission Mumbai. We also thank to Marathi Vidnyan Parishad, Pune Vibhag for their immense help in writing Marathi definitions.

References

1. Maedche, A.: Ontology Learning for the Semantic Web. Kluwer Academic Publishers, Amsterdam (2002)
2. Maynard, D., Funk, A., Peters, W.: Using Lexico-syntactic ontology design patterns for ontology creation and population. In: Proceedings of the 2009 International Conference on Ontology Patterns, vol. 516 (2009)
3. McDonald, R.: Extracting relations from unstructured text. UPenn CIS Technical Report (2004)
4. Maynard, D., Tablan, V., Ursu, C., Cunningham, H., Wilks, Y.: Named entity recognition from diverse text types. In: Recent Advances in Natural Language Processing Conference (2001)
5. Navigli, R., Velardi, P., Cucchiarelli, A., Neri, F.: Automatic ontology learning: supporting a per concept evaluation by domain experts
6. Maynard, D., Funk, A., Peters, W: SPRAT: a tool for automatic semantic pattern-based ontology population
7. Carlson, A., Betteridge, J., Kisiel, B., Settles, B., Hruschka, Jr., E.R., Mitchell, T.M.: Toward an architecture for never-ending language learning. In: Association for the Advancement of Artificial Intelligence (2010)
8. Davulcu, H., Vadrevu, S., Nagarajan, S.: OntoMiner: bootstrapping and populating ontologies from domain specific web sites
9. El-gaya, M.M., Mekky, N., Atwan, A.: Efficient proposed framework for semantic search engine using new semantic ranking algorithm. Int. J. Adv. Comput. Sci. Appl. **6**(8), P136–P143 (2015)
10. Cai, B., Li, Y.: Design and development of semantic-based search engine model. In: 7th International Conference on Intelligent Computation Technology and Automation, pp. 145–148 (2014)
11. Dilek, S., Karacan, H., Jahangiri, N., Afzali, S.: Ontology Creation for an Educational Center 978-1-4673-1740-5/12/©2012 IEEE
12. Saggion, H., Funk, A., Maynard, D., Bontcheva, K.: Ontology-based information extraction for business applications. In: Proceedings of the 6th International Semantic Web Conference (ISWC 2007), Busan, Korea (2007)
13. Hearst, M.A.: Automatic acquisition of hyponyms from large text corpora. In: Conference on Computational Linguistics (COLING'92), Nantes, France. Association for Computational Linguistics (1992)
14. Pantel, P., Pennacchioni, M.: Espresso: leveraging generic patterns for automatically harvesting semantic relations. In: Proceedings of Conference on Computational Linguistics/ Association for Computational Linguistics (COLING/ACL-06), Sydney, Australia, pp. 113–120 (2006)
15. Pantel, P., Ravichandran, D.: Automatically labeling semantic classes. In: Proceedings of HLT/NAACL-04, Boston, MA, pp. 321–328 (2004)
16. Mikolov, T., Yih, W.T., Zweig, G.: Linguistic regularities in continuous space word representations. In: NAACL HLT (2013)
17. Bach, N., Badaskar, S.: A review of relation extraction. Literature review for language and statistics II (2007)
18. Saranya, K., Jayanthy, S.: Onto-based sentiment classification using machine learning techniques. In: International Conference on Innovations in information Embedded and Communication Systems (2017)
19. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781 (2013)
20. Collobert, R., Weston, J., Bottou, L., Karlen, M., Kavukcuoglu, K., Kuksa, P.: Natural language processing (almost) from scratch. J. Mach. Learn. Res. **12**, 24932537 (2011)

21. Friedman, N., Geiger, D., Goldszmidt, M.: Bayesian network classifiers. Mach. Learn. **29**(2–3), 131–163 (1997)
22. Liu, J., Qin, L., Wang, H.: An ontology mapping method based on support vector machine
23. Johnson, I., Ab, J., Charnomordic, B., Destercke, S., Thomopoulos, R.: Making ontology-based knowledge and decision trees interact: an approach to enrich knowledge and increase expert confidence in data-driven models
24. Chieu, H.L., Ng, H.T.: Named entity recognition: a maximum entropy approach using global information. In: Proceedings of the 19th International Conference on Computational Linguistics, vol. 1 (2002)
25. Siddharthan, A.: An architecture for a text simplification system. In: Proceedings of the Language Engineering Conference (LEC'02) 0-7695-1885-0/02© 2002 IEEE

Identifying Influential Users on Social Network: An Insight

Ragini Krishna and C. M. Prashanth

Abstract The advancement in the speed of the internet connection on handheld devices has led to an increase in the usage of social media. This drew the attention of advertisers to use social media as a platform to promote their products thus leading to an increase in the sales of their product, increasing the brand awareness. To increase the rate of information dissemination within a short period of time, influential users on social media were targeted, who would act as the word-of-mouth advertisers of the product. However, there are various parameters on which the influence of a user has to be determined. The parameters can be (1) the connectivity of the user in the network (2) knowledge/interest of the user on a particular topic/product/content (3) activity of the user on the social media. This survey focuses on the various methods and models for identifying influential nodes and also the effect of compliance, where a user falsely agrees to the content of another influential user by retweeting, just to gain status or reputation and thus increasing his influential score. Thus, the list of influential nodes of a social network can be faked upon, due to this issue.

1 Introduction

It is an era of automation and people's digital presence is increasing and finding the influence of the user based on their digital presence is the talk of the time. The digital presence of the user on various social media like facebook, twitter, linkedln, google+, etc., are used to find the influence of the user on the digital social networks.

R. Krishna (✉) · C. M. Prashanth
Department of Computer Science and Engineering, Acharya Institute of Technology, Bangalore 107, India
e-mail: raginis.17.pfcs@acharya.ac.in

C. M. Prashanth
e-mail: prashanthcm@acharya.ac.in

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_34

With the vast usage of the social network [1] to connect to friends and pass information across the network of friends, the social networks are being targeted for the use of business. Thus, social media is used for promoting a product or for passing on information or news. To speed up the process of the spread of information, people target the influential users.

It is already proven and now it is a fact that social media marketing affects the brand awareness and the purchase intentions [2] but the author "Romy Sassine" [1], also discusses that the effect of influencers for a brand marketing is decreasing due to the overuse of the social media.

The reasons for such a trend can be the overuse or wrong placement of the hashtags in a particular message or it may be due to the competition among the influencers to get more brands contract instead getting connected to the audience.

According to [3], the advertizing done using the word-of-mouth strategy increased the sales of the product by twice and the customers gained from this method had a higher retention rate of 37%. The author also states that just having a large number of followers does not indicate an influence, moreover to determine the influence of a person, his credibility towards the topic and strength of the relationship with followers along with the total numbers of followers should be considered.

To determine the influential users in the social network is one of the most trending research areas. There have been many research work carried to find the influential users in a social network. In this paper, Sect. 2 introduces the methods used to find the centrality of the nodes in a network. In Sect. 3, the various methods used to find the influence on social media is discussed. Section 4 lists out the various shortcomings of the methods and in Sect. 5, we discuss the prospect into the future of identifying the influence on social media.

2 Background

A large number of methods have been proposed to find the influential users on the social network so far. Initially, researchers proposed methods which were based on the centrality [4, 5] of the nodes in the network, such as degree centrality, betweenness centrality, closeness centrality, eigenvector centrality, these methods mostly assumed that the network was static, as they never considered the interactions among the nodes of the network. Thus, to consider the dynamicity of the network various researchers started considering the interactions among nodes in the network [6, 7], such as α-centrality and β-centrality. Thus, link strength was determined based on the level of interactions among the nodes in the network to consider the dynamicity of the network. The link strength indicates the ratio of interaction with the concerned node as compared to the interaction with other nodes, which gives the weightage to the concerned link. The link strength between nodes u and v is the ratio of interaction between u and v to the sum of total interactions of u and total interactions of v.

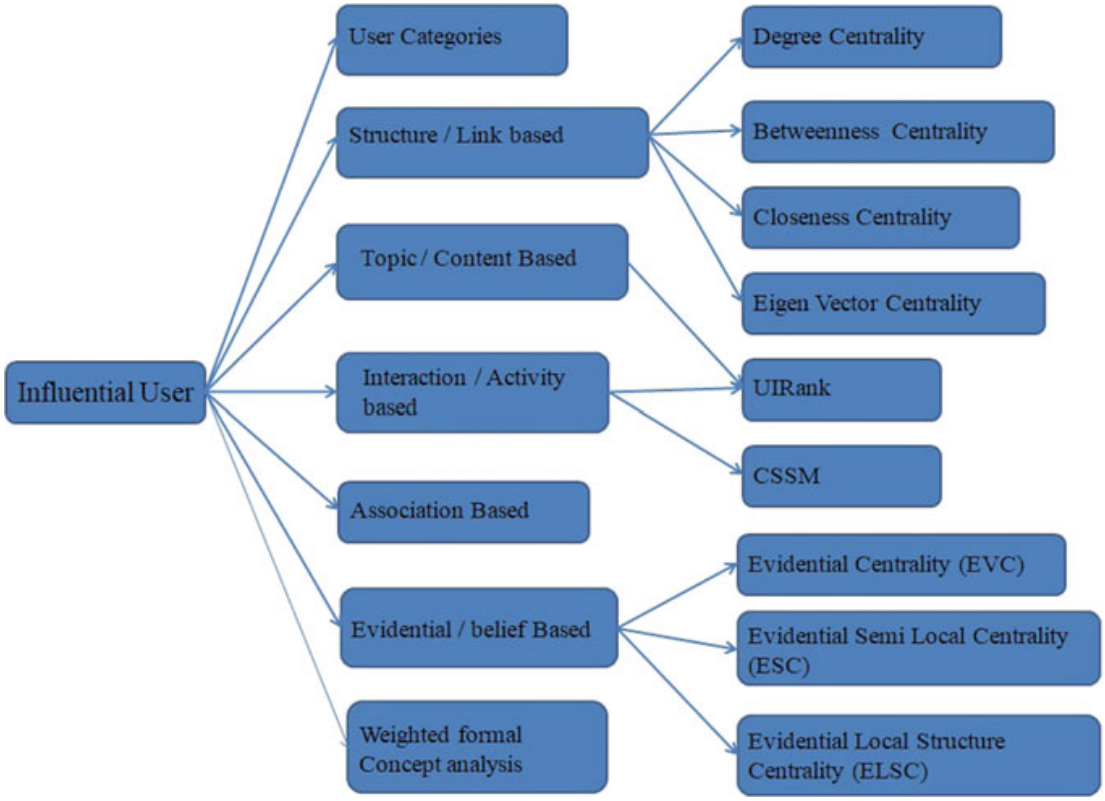


Fig. 1 Categorization of the various methods to find influential users in social networks

The influence of a node in the network can be determined based mainly on its link strength, its interest in the topic of discussion or its activity in the network. Figure 1 shows the categorization of the various methods which determine influential users on the social network.

3 Methods for Finding Influence on Social Network

In this section, we will discuss methods (Fig. 1) to find the influential users based on various parameters like the level of interaction among the nodes, interest of the user in the topics, etc.

3.1 Structure or Link Based Influence

These methods used the connectivity of the nodes in the network to find the influential nodes. However, with the increasing usage of the social media, the connectivity among the nodes kept changing at a frequent interval, thus making them more dynamic, which was not addressed in the link-based methods of finding influence [4, 5, 7].

3.2 User Categories Based Method to Optimize the Popularity of Twitter Messages

There are number of users on twitter who generate a million of content everyday however many of the contents go unnoticed or unread which can reduce the popularity of the applications among the users. This may happen because different categories of people will be interested in different topics/content. To detect the effect of the content in the tweet, on its popularity, the tweets were categorized on topics like brands, celebrities, community, entertainment, media, sport, places, and society [8]. The activities of the users were monitored for all these categories. The data was collected and five features-(1) number of mentions, (2) number of hashtags, (3) number of URLs, (4) the number of pictures, and (5) the sentiment of the tweet were collected, where the first four were directly derived from the twitter API but the fifth feature, the sentiment of the tweet, was extracted from the tweet using labMT word list. The sentimental analysis was done on the tweets and then segregated into neutral and positive or negative, 0 being neutral and 1 being extreme either positive or negative.

1. Analyzing the Influence of the content on the popularity of the user's tweets:
   The popularity of the user's tweets was determined based on the features of the content and how the popularity differs for a tweet of different user categories. Each feature of the content was ranked based on its importance for the popularity of each and every user categories.
   To minimize the sum of squared error in the prediction for finding the most popular content, LASSO (least absolute shrinkage and selection operation) [9] was used.
   From the results in [8], it was observed that the sentiment feature plays a very insignificant role towards the popularity of the tweets for every user categories, whereas when considering retweets and favorites, the feature-mentions were ranked high for favorite but had lower ranking for retweets.
   When considering the user categories for the most number of retweets, the feature -picture was ranked highest for communities and brands, whereas the feature-URLs was ranked highest for media and places.
2. Predicting the popularity of the tweets
   This work predicts the popularity of a tweet given its features. A prediction model was built using a regression model consisting of 50% of the tweets for the training and 25% to validate and another 25% for evaluating the model.
   There were three ways in which the prediction was done, Generic model, the category-specific model, and the smoothed model. In the generic model, the training was done on all the data, whereas in the category-specific model, the training was done based on the segregation of the categories of the tweets.
   The Generic model gave high prediction errors for different categories of users. In the category-specific model, the category to which, the tweet in the test set,

belongs was found out and then it was used in the prediction model. According to the arrived results [8], the category-specific model showed improvement in the prediction of the number of retweets, however, it did not outperform the generic model in many other cases, the reason for this may be because the test set considered was very small.

Thus, to get the best from both the models, the best elements of both the models were combined to get the smoothed model. This was achieved by considering the weighted average of both the generic model and the category-specific model. The smoothed model proved to be giving better results as compared to generic or category-specific model.

3.3 *Interaction or Activity Based*

1. Time slice-based interaction:

To calculate the level of interaction of nodes in the network, methods like indegree and outdegree centrality were used. High indegree of a node represented an inactive user as he/she consumes the information but does not help in the information dissipation.

The steps involved to calculate the influence of a user based on its interaction with its neighbors are [10]:

- Indegree and outdegree centrality
- Link strength
- Clustering value based on indegree and outdegree
- Calculating the influence

Let's take an example to illustrate it:

Step 1: The indegree for the nodes of the network from Fig. 2 at $t1$ are:

$A = 0,\ B = 2, C = 1,\ D = 2$ and $E = 2$

Thus, the outdegree for the nodes at time $t1$ are:

$A = 1,\ B = 1,\ C = 2,\ D = 1$ and $E = 1$

Thus, if there are users with high indegree and zero or low outdegree, it represents a very low activity.

Step 2: To find the influence of the nodes among friends of a node, the strength of the link that the nodes share across the network is calculated. If the link between nodes is found to be strong, then they can be influenced by their friends with a very little effort from the influencer. Thus, the link strength between the two friends 'a' and 'b' can be calculated by:

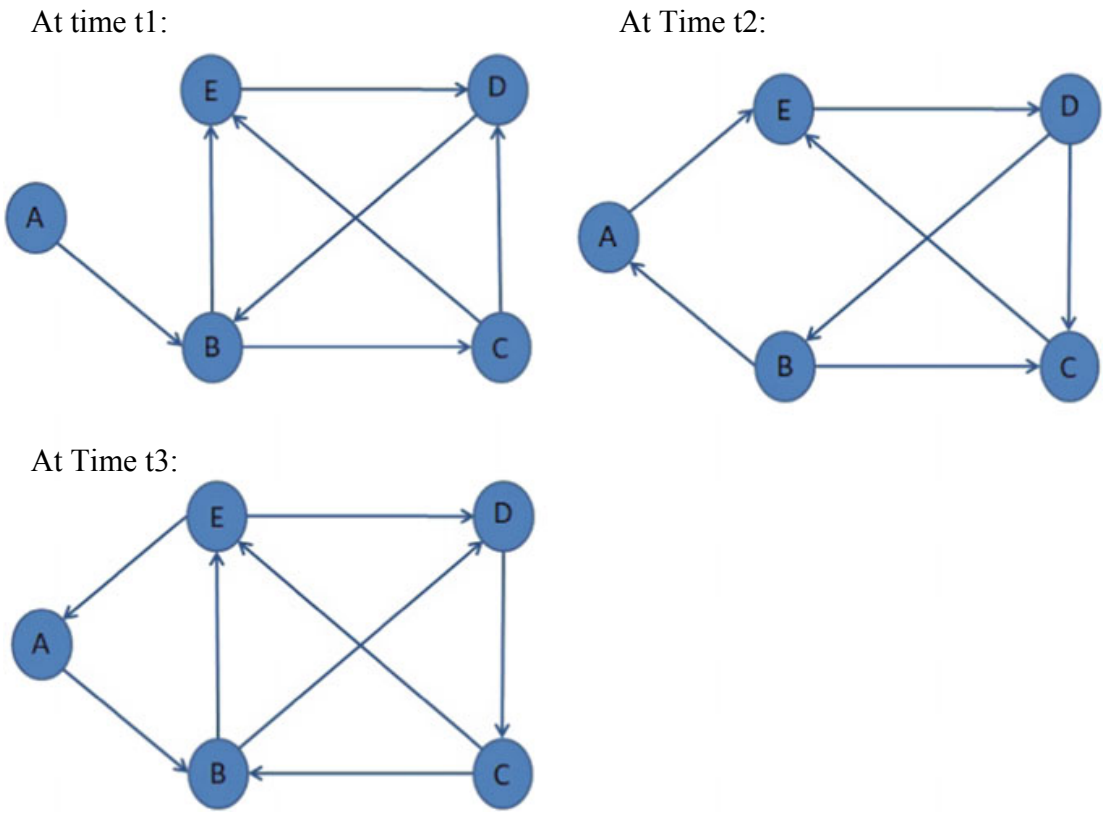


Fig. 2 Representation of nodes interaction in a network at the time slots t1, t2, and t3

$$L(b,a) = \frac{|\bar{\epsilon}(b,a)| + |\bar{\epsilon}(a,b)|}{d_o(b) + d_o(a)} \tag{1}$$

In the example that has been taken, there are three time periods that are considered as the time of interactions between the various nodes in the network.

Thus, the indegree and outdegree for the nodes across all the three-time slots are:

Indegree centrality:

$d_i(A) = 2$ $d_i(B) = 4$ $d_i(C) = 4$

$d_i(D) = 5$ $d_i(E) = 6$

Outdegree centrality:

$d_o(A) = 3$ $d_o(B) = 4$ $d_o(C) = 5$

$d_o(D) = 4$ $d_o(E) = 3$

Thus the link strength is calculated and the values are as follows:

Link strength between node A and B:

L(B,A) = 0.428

Link strength between Node A and E:

L(E,A) = 0.333

In a similar way, the link strength between various nodes is calculated. Thus it is inferred that the link strength between the node A and B and the nodes D and E, with the strength value of 0.428, seems to be stronger as compared to the other nodes.

Step 3: clustering value based on indegree and outdegree:

The connectivity of the node within a group of nodes is represented by the cluster value. The higher the cluster value of the node, the more connected it is, to a number of clusters and thus has more reachability to pass on the information to more number of people in the network.

The high indegree clustering value indicates that a number of messages are being transmitted from the neighboring nodes to the node in question, which implicitly means that the concerned node is involved in more number of discussions and thus there will be a high chance of him/her adopting other's behavior. The indegree clustering value is given by:

$$C_i(a) = \frac{\sum_{b \in S} d_i(b)}{d_i(a) * (d_i(a) - 1)} \tag{2}$$

From Eq. (7), the indegree clustering value for node A is 5, node B is 1.75, node C is 1.25, node D is 0.7, and node E is 1.25. It is evident that node A has high indegree clustering value and thus is connected to more number of clusters.

The number of messages transmitted from the concerned node to its neighboring nodes is indicated by outdegree clustering value. Thus a node with high outdegree clustering value has a high chance of being influential. The outdegree clustering value for a node is given by:

$$C_o(a) = \frac{\sum_{b \in S} d_0(b)}{d_o(a) * (d_o a - 1)} \tag{3}$$

After calculating the outdegree clustering values of the nodes in the network, A has a value of 1.167, B has 1.25, C has 0.55, D has 1, and E has 1.167. This clearly shows that the nodes A and E can be influencers in the network. Any node in the network with a high outdegree but a zero indegree value can be a spammer.

Step 4: Thus to find the influencers in a network, the values of the indegree centrality, outdegree centrality and indegree – outdegree clustering values will be used. Thus, from the above discussions, it is quite clear that the user who has high outdegree centrality with high outdegree clustering value will be influential. But the weightage of these parameters depends on the case. For example, maybe the user does not have high outdegree cluster value but he generates a lot of content which make him quite active and thus the interaction of the user with it neighbor will go high, which may make him influential. Thus, to normalize these parameters, two normalizing factors, α_1 and α_2 are introduced. The influencer of the network can be calculated by:

$$Infr(a) = [\tanh(d_0(a))] * (\alpha_1 d_0(a) + \alpha_2 C_0(a)) * \sum_{t\varepsilon S} L(b,a) \tag{4}$$

After calculating the influencer values for the nodes, it is found that the node B is most influential with an influencer value of 3.690 after assuming the value for both α_1 and α_2 as 0.5.

The users with high indegree centrality and high indegree cluster value are the users who can be influenced by its neighbors. Thus, they should not be mistaken for inactive users. To remove the inactive users in finding the influencer or the influenced user, the hyperbolic tangent is used to translate the outdegree value of the concerned node '*a*' to either zero or one.

The influenced value of a node is:

$$Infd(a) = [\tanh(d_0(a))] * (\alpha_3 d_i(a) + \alpha_4 C_i(a)) * \sum_{t\epsilon S} L(b,a) \tag{5}$$

Furthermore, the influenced value for each node in the above example is calculated and thus node B has the highest value with 4.087 with the assumption of α_3 and α_4 as 0.5. Thus, from the example network, it can be inferred that the node B creates influence on its neighboring nodes and it is influenced too by its neighbors.

The authors [10] have worked towards finding the influence of users using interaction among them in the network, which clearly avoids the problem of spammers or inactive users in the network.

2. Discovering influence using community scale-sensitive maxdegree (CSSM)

The community in a network is detected based on the interaction among the nodes in the network which represent similarity among the nodes present in the same community. The communities can be identified using various clustering algorithms.

Influence of a node depends on the parameters like the centrality based on the degree of the node, sum of neighbor's degree, and the attributes of the node [11].

The outdegree centrality (ODC) of the node is given by:

$$\mathrm{ODC}(x) = \frac{o(x)}{m-1} \tag{6}$$

where $o(x)$ is the number of neighbors adjacent to x and m is the total number of nodes in the network.

The sum of neighbor's degree (SND) is given by:

$$\mathrm{SND}(x) = \sum_{i=1}^{o(x)} o(y_i) \tag{7}$$

where y_i is the ith neighbor node of x. The value represented by SND denotes the number of followers of the node x at level 2.

The results obtained from ODC and SND are summed together for each community and they are finally given as the most influential nodes in the network.

The above method considers only the interaction among the nodes in the network, however, the interaction can be using a negative tweet, which will give an inaccurate result in identifying influence. This issue is rectified by considering the content of the interaction.

3.4 Content Based User Influence Rank (UIRank)

The influence of user on a microblogging site is calculated by a method called UIRank where the node's information disseminating ability (connectivity) and the contribution of the user's tweet (content) is considered [12].

The content generated by the user, that is, the tweet, retweet, and the comment is considered and the influence of the tweet is calculated by considering the retweet and comments on that tweet and is given by:

$$sp(u) = \sum_{t \in \text{Tweets}(u)} Rr(t) + Cr(t) \tag{8}$$

where $sp(u)$ is the probability that the user's u tweet propagates from user u to the neighbors of the fans of user u, $Rr(t)$ is the retweet to the read ratio of the tweet t and $Cr(t)$ is the comment to the read ratio of the tweet t.

The user's influence in the network is found out by finding the connectivity of the user in the follower relationship network. The connectivity of the user will decide the rate of diffusion of the user's tweet in his network and then his follower's network and so on until it reaches all the nodes in the connection. The rate at which the user's tweet is spread depends on the connectivity of the user in his follower relationship and then his fans influence their respective neighbor's network

The centrality of the node gives the importance of the node in the relationship network. The centrality of the node can be measured by degree centrality, betweenness or closeness centrality.

$Sa(u)$ is defined to measure the network influence of a node in the network and it is represented by:

$$sa(u) = C_d(u) + C_b(u) + C_c(u) \tag{9}$$

where $C_d(u)$, $C_b(u)$, and $C_c(u)$ are the degree centrality, betweenness centrality, and the closeness centrality of the user respectively.

The nodes disseminating ability is given by:

$$\pi_{uv} = \begin{cases} \frac{sp(u) + sa(u)}{u_{\text{out}}} & \text{if } u \text{ follows } v \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

Thus, all the nodes disseminating ability is calculated and put in the form of a transition matrix. Using a random walk method and the page rank algorithm, the UIRank is defined as:

$$IR(u) = \alpha \sum_{v \in \text{followers}(u)} (IR(v) * \pi_{uv}) + (1 - \alpha) \tag{11}$$

where α is a delay factor.

The performance of UIRank is compared to the other four influence finding methods like retweet Rank, outdegree rank, tunk Tank, and fans Rank. The results derived show that UIRank performs best on precision, recall, and F measure, as compared to the other four.

Further, the work can be improved by considering the topic of the tweet posted by the user to find the influence of the user on that topic and also finding the trust factor of the tweet posted by the user.

3.5 Based on Evidence Theory or Belief Based

Evidence theory or Dempster–Shafer theory (DST) works on the principle of reasoning with uncertainty. The theory combines, evidence from many sources and then arrive at a belief.

The evidential centrality (EVC) uses the nodes degree and the weight strength of each node in the network to find the influential nodes in the network [13]. However, EVC assumed a uniform distribution regarding the connectivity of the nodes, however, in the complex networks, the interconnectivity of the nodes in the network is unpredictable, and thus it may not be suitable for the complex weighted network. Thus, to overcome the assumptions of the EVC method, semi local structure centrality (ELSC) along with modified EVC [14] was used to find out the influential nodes in weighted networks.

The most common relationships in the twitter social network are follow, mention, and retweet. The belief function is used [15, 16] to overcome the data imperfection in the twitter dataset due to the limitation on the data extraction from twitter. It considers the strength of the link between the direct neighbors, information exchange, and propagation between the users and the assumption that if the user is connected to an influential node, becomes influential. Thus, the influence calculation has been taken up in two sections, the first section where the influence of the node on the directly connected nodes are calculated and the section calculates the influence of the user's neighbors in their respective network. The results of this

method show that it fetches better seed sets of influential users as compared to the previous methods.

3.6 Formal Weighted Concept Analysis

Formal concept analysis (FCA) [17] uses binary objects and attributes relationships to build the knowledge hierarchy which very well represents the relationship between the objects and attributes. A social network can be represented in terms of the formal context by representing the connectivity of the network where nodes indicate the objects and the attributes whereas the edges indicate the binary relationship between them.

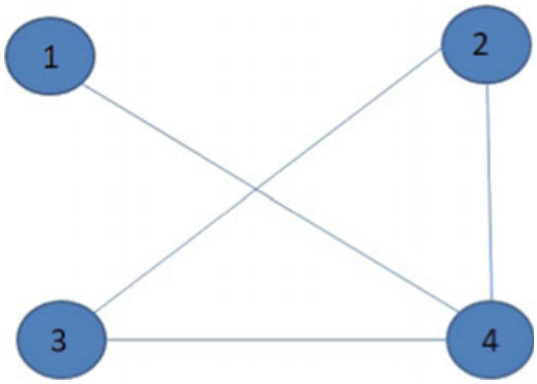
To calculate the important nodes using FCA includes three steps. First, the concepts are computed from the adjacency matrix of the formal concepts. Then, the weight of each node is calculated based on the concepts. Thus, to rank the nodes in the network based on its importance, the weighted formal concept Analysis (WFCA) [18] is used, where the hierarchical tree is generated from the cluster of nodes and then the concepts of the network are computed.

The adjacency matrix in Fig. 3b of the graph from Fig. 3a is considered as the formal concept where a row represents an object and a column represents an attribute. Thus, the weight of each node is calculated and then ranked according to this weight, given by

$$w_i = \sum_{k=1}^{n} \frac{O_{ik}}{A_{ik}} \tag{12}$$

where O_{ik} is the object number of the concept C_{ik} and A_{ik} is the attribute number of the concept C_{ik}

(a)



(b)

	1#	2#	3#	4#
1	0	0	0	1
2	0	0	1	1
3	0	1	0	1
4	1	1	1	0

Fig. 3 An example network **a** Graph G for the network. **b** Formal context of the network G

Once the weights of the nodes are calculated, they are ranked depending on their weights.

The adjacency matrix for the network in Fig. 3a is constructed and shown in Fig. 3b. The entries in the adjacency matrix are either 1 or 0 representing the presence of an edge between the nodes or absence of it, respectively.

Figure 4 represents the Hasse Diagram which contains all the concepts of the formal context. The weights, $W_{i,}$ of all the nodes are calculated using the equation 17 to find the rank of the nodes. From the example network of Fig. 3a, the weight of the node 1 is calculated by using the value of its attributes and objects as follows:

({4}, {1#}), ({4}, {1#, 2#}), ({4}, {1#, 3#}), ({4}, {1#, 2#, 3#,}). Thus, according to the equation 17, the weight of the node 1, W_1, is: $\frac{1}{1} + \frac{1}{2} + \frac{1}{2} + \frac{1}{3} = 2.33$. Similarly, the weight of node 4 is: $\frac{3}{1} + \frac{1}{2} + \frac{1}{2} = 4$.

After calculating the weights of each node in the network, they are ranked. So from the above example, using WFCA, node 4 is considered to be most influential among other nodes in the network.

4 Outcome of the Review

- Since the social network's size is increasing exponentially, thus to find an influential node, a sample of dataset representing the social network is

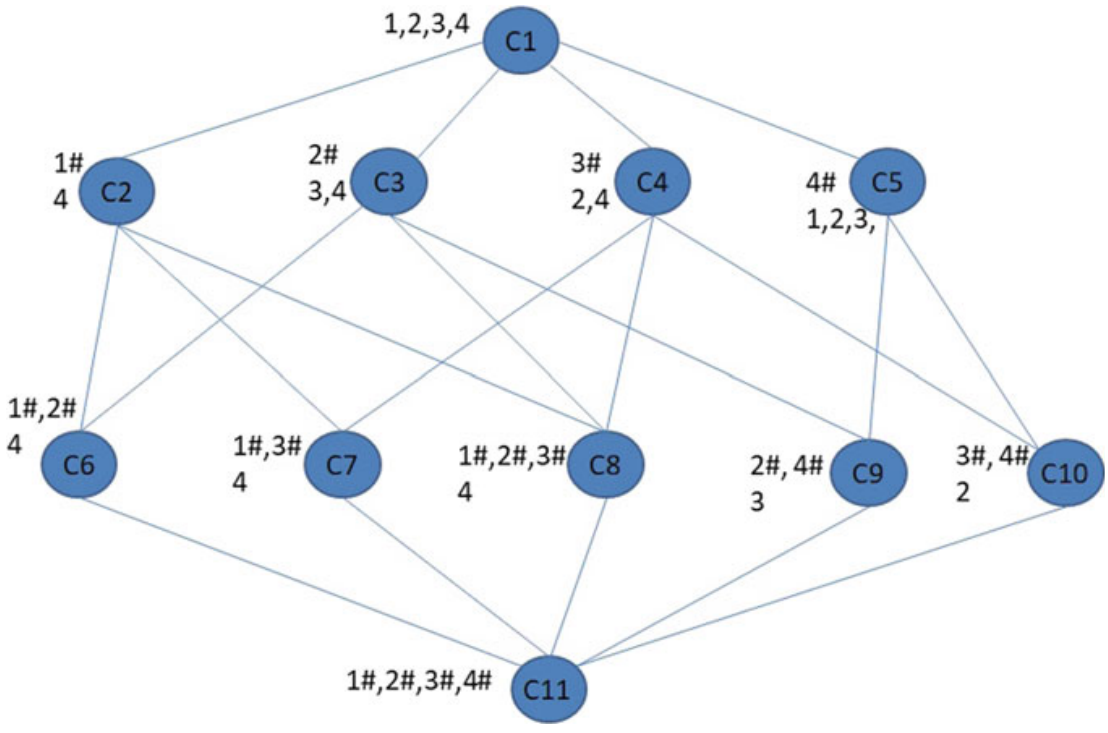


Fig. 4 Concept lattice (Hasse diagram) for the graph in Fig. 3a

considered. Thus, the accuracy may suffer as the result depends on the sample set that is chosen.

- The user's interest can be determined based on the activity of the user, like tweet or retweet. However, if a particular topic is not of a user's interest, he/she may not react to it at all or may give a negative comment. Thus for any business model trying to use influence on social media to promote their business, can't blindly use the user (influential) without knowing their sentiment regarding the product. For example, a user is a well-known wrestler who has won a gold medal at Olympics (thus influential), if an artificially flavored juice product is given to him to promote among his followers, he may not react to it at all.
- Influence of web users is determined by (1) connectivity of the nodes [4–6, 19, 20], (2) through user categories [8], (3) interaction-based [10, 21, 13], and (4) based on the content of the tweet as well as the connectivity [14]. It is observed that the two methods, the content of the tweet and the connectivity is considered, yields a better accuracy, thus there is a need to find the influence based on a method which considered all the above parameters and thus accuracy of the result will be improved.

5 Conclusion

The determination of influential users in the network can be done in many ways; however, the result of finding the influence by combing two or more methods is better as compared to determining by just one method as shown in [12, 14, 15]. Thus, to obtain an accurate and improved result, it is advisable to combine two or more appropriate methods. The study also reveals that one can create a fake influence by retweeting/commenting/following the already established, highly influential user on social media, without being actually influential. The author proposes to explore this issue in future work. A method to detect such fake influencers may be devised so that such users are excluded from the list of the influential user.

References

1. Sassine, R.: What is the impact of social media influencers? Retrieved on February 28th 2018, from http://digital-me-up.com/2017/04/06/impact-social-media-influencers/
2. Duiverman, C.: Does influencer marketing have an impact on actual sales? Retrieved on February 28th 2018, from https://thecirqle.com/blog/2017/9/5/influencer-marketing-impact-sales
3. Wong, K.: The explosive growth of influencer marketing and what it means for you. Retrieved on February 28th 2018, from https://www.forbes.com/sites/kylewong/2014/09/10/the-explosive-growth-of-influencer-marketing-and-what-it-means-for-you/#1227987752ac
4. Bonacich, P.B.: Power and centrality: a family of measures. Am. J. Soc. **92**, 1170–1182 (1987)

5. Bonacich, P.: Eigenvector-like measures of centrality for assymetric relations. Soc. Netw. **23** (3), 191–201 (2001)
6. Freeman, L.C.: Centrality in social networks conceptual clarification. Soc. Netw. **1**(3), 215–239 (1978–1979)
7. Lerman, K., Ghosh, R., Kang, J.H.: Centrality Metric for Dynamic Networks. University of Washington, Washington, DC (2010)
8. Lemahieu, R., et al.: Optimizing the popularity of Twitter messages through user categories. In: IEEE International Conference on Data Mining Workshop (ICDMW). IEEE (2015)
9. Tibshirani, R.: Regression shrinkage and selection via the lasso. J. R. Stat. Soc. Ser. B (Methodological) **58**, 267–288 (1996)
10. Rad, A.A., Benyoucef, M.: Towards detecting influential users in social networks. In: MCETECH (2011)
11. Hao, F., et al.: Discovering influential users in micro-blog marketing with influence maximization mechanism. In: 2012 IEEE on Global Communications Conference (GLOBECOM). IEEE (2012)
12. Jianqiang, Z., Xiaolin, G., Feng, T.: A new method of identifying influential users in the micro-blog networks. IEEE Access **5**, 3008–3015 (2017)
13. Wei, D., et al.: Identifying influential nodes in weighted networks based on evidence theory. Phys. A Stat. Mech. Appl. **392**(10), 2564–2575 (2013)
14. Gao, C., et al.: A modified evidential methodology of identifying influential nodes in weighted networks. Phys. A Stat. Mech. Appl. **392**(21), 5490–5500 (2013)
15. Ren, J., et al.: Identifying influential nodes in weighted network based on evidence theory and local structure. Int. J. Innov. Comput. Inf. Control **11**(5), 1765–1777 (2015)
16. Jendoubi, S., et al.: Two evidential data based models for influence maximization in Twitter. Knowl. Based Syst. **121**, 58–70 (2017)
17. Wille, R.: Restructuring lattice theory: an approach based on hierarchies of concepts. In: Ordered Sets. Springer, Dordrecht, pp. 445–470 (1982)
18. Sun, Z., et al.: Identifying influential nodes in complex networks based on weighted formal concept analysis. IEEE Access **5**, 3777–3789 (2017)
19. Chen, D., et al.: Identifying influential nodes in complex networks. Phys. Stat. Mech. Appl. **391**(4), 1777–1787 (2012)
20. Erlandsson, F., et al.: Finding influential users in social media using association rule learning. Entropy **18**(5), 164 (2016)
21. Li, Y.M., Shiu, Y.L.: A diffusion mechanism for social advertising over microblogs. Decis. Support Syst. **54**(1), 9–22 (2012)

Factex: A Practical Approach to Crime Detection

Rachna Jain, Anand Nayyar and Shivam Bachhety

Abstract The crime on roads is a major problem faced today by all the modern cities. Road Transport is the most common escape route for many criminals. Thefts and many other crimes remain unregistered and unsolved due to lack of evidence. Effective tracking of vehicles and criminals is still a big problem and involves plenty of resources. To evade such a condition, we have proposed a machine learning-based practical crime detection system using the text and face recognition techniques. Such systems will be proved useful in parking lots, toll stations, airports, border crossings, etc. In the proposed system, the text recognition involves extracting the characters present in the Indian number plates and the predicted output will be compared with the registered vehicle database. Simultaneously, Face recognition feature constitutes identifying criminal faces based on certain face regions and then mapping the respected coordinates with the criminal database. The proposed system presented in this research paper targets to deliver improvised outcomes considering the time constraints and accuracy with more than 85% successful recognitions in normal working conditions with the goal to accomplish the successful detection of crime using machine learning algorithms such as KNN, SVM, and face detection classifiers to present a practical real time detection.

Keywords Crime detection · Dlib · Haar cascades · KNN · License plate · Text recognition · Face recognition · OpenCV · SVM

R. Jain · S. Bachhety
Computer Science and Engineering, Bharati Vidyapeeth's College of Engineering, New Delhi, India
e-mail: rachna.jain@bharatividyapeeth.edu

S. Bachhety
e-mail: shivambachhety97@gmail.com

A. Nayyar (✉)
Graduate School, Duy Tan University, Da Nang, Vietnam
e-mail: anandnayyar@duytan.edu.vn

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_35

1 Introduction

The recent works in computer vision and machine learning, text recognition, and face recognition can be visualized as effective measures against rising criminal activities. Using video surveillance, license plate recognition has been an effective approach utilized in the past for traffic management and crime detection. With the integration of text recognition from number plate, an added feature of face recognition can prove to be a boon to the overall system. These techniques involve certain considerations of subjects, posture, emotions, and light for its smooth implementation. Text recognition [1] involves the basic steps of extraction, recognition, and identification of characters, whereas Face recognition [2] involves detecting faces, matching the input face images based on pre-stored face data points and its coordinates and then identifying it. The implemented method with a limited dataset can be seen as a milestone for better surveillance if extended to actual case databases [3]. In the proposed system, KNN is used as text recognition technique and HAAR Faces Classifier and SVM technique for the performing face recognition.

In India, the problems related to traffic are increasing day by day. A typical number plate in India has black foreground color and white background color for private cars. The number on the license plate consists of the two-digit letter showing "state code" followed by two numeral digits, followed by a single letter and then the last four digits. The system consists of the identification of number plate present in the car and then segmentation of all the characters on the number plate. The identification task is interesting because of the nature of the light. Noise in the image background and other environmental factors can result in low accuracy of output. In the proposed system, the edge detection method and mid-filtering are used for noise removal. The next aspect of face recognition consists of mining some significant features of the image and then transforming them for classification. It mainly involves geometric computation of facial features as regions. The face recognition process consists of making a database of faces with numerous facial images for each person. Next, is to detect faces present in the database and then train the face classifier. The last step is testing the classifier to recognize faces it was trained for.

This paper highlights a novel approach to recognize the license plate and criminal face in real time and then matches the output in the vehicle registration database and criminal record face database respectively. The output result produces an email alert to the respective concerned authorities with the license plate number and recognized face details from the database.

1.1 Organization of Paper

Section 2 elaborates the related works highlighting the contributions by several other authors in the corresponding domain. Section 3 highlights the proposed

system, methodology as well as the algorithm used for the live implementation of the proposed system. Section 4 gives a detailed overview of the results produced by the system. Section 5 concludes the paper with future scope.

2 Related Works

Various techniques are proposed and implemented in the past few years regarding license plate detection using visual image processing. Vuong et al. [4] focus on license plate detection based on features like color feature, edge detection, equilibrium, and other operators. Gou et al. [5], proposed license plate detection system in which the structure is regarded as a crucial factor to verify characters from number plate and extraction is done via projected segmentation. For most of the algorithms, Sobel as edge detection operator is frequently used and marked as the best method to transmit from grayscale image to gradient image. Ahmad Radmanesh [6] focused chiefly on the vertical edge detection via the Sobel detection, in comparison to the horizontal edge density of background objects. The template matching and neural networks proved to be an effective technique for character segmentation and recognition as used by Zhu Wei-gang et al. [7] but template matching feature has many limitations when compared to neural networks. After segmentation, the neural network method is expected to be the template match one.

The first approach established for proficiently demonstrating faces was using PCA. The aim of this method was to present a human face as a coordinate system. The vectors setup on the coordinate system were known as eigenfaces vector. Later, Turk and Pentland used this to develop an eigenface-based algorithm for face recognition [2]. SVMs as linear classifiers were first used by Osuna et al. [8] to feature the border between the resulting hyperplane and the training set examples. The ultimate goal of optimal hyperplane was to minimize the error in the classification of the unseen test patterns. Schneiderman and Kanade [9] labeled an object recognition algorithm and projected the results via Bayesian classifier. The result lies in computing the probability of a face to be present in the image based on the count of the frequency of occurrence of a series of patterns over the training images. Saraswathi et al. [10] proposed face recognition authentication process via feature extraction technique, i.e., Linear Discriminant Analysis (LDA). It generates facial regions and then classifies using Euclidean Distance. The experimental results were as high as about 93.7% using the LDA feature set. Extending the PCA to GPU power, Bhumika Agrawal et al. [11] with the help of NVIDIA CUDA performed computations for faster processing of face recognition on Principal Component Analysis algorithm. In the later phase, the work is compared to varied CPU implementations. Jian Yang et al. [12] used the one-dimensional, pixel-based error model based on regression analysis to characterize two-dimensional error image matrix, namely nuclear norm-based matrix regression (NMR) and the alternating direction method of multipliers (ADMM) to estimate the regression coefficients.

3 Proposed System

The present system is mostly based on license plate number recognition using video surveillance, but it has several limitations. The systems based on OCR are highly time-consuming lacks any facial identification for crime detection. The system proposed in this research paper is highlighted in Fig. 1 for observing and supervising real time traffic in the parking lots of private and public organizations, airports, toll booths, and border crossings. Apart from identifying any stolen vehicles, it is highly proficient to track escaping criminals. With minimal processor units and cameras, the system can be installed with much ease. A 24/7 all-time working solution can eliminate human dependency and prevent any crime in the case of negligence also. It can also help keep evidence of crime by capturing images and be used in future investigations. Other benefits of the proposed system include fake number plate detection, tracking criminal activities and better surveillance.

The proposed system has certain limitations in terms of light illumination, masked face identity or makeup, predicting accuracy, facial emotions, and expressions.

3.1 Methodology: Algorithms and Implementation

3.1.1 Text Recognition

K-Nearest Neighbors (KNN) Algorithm

KNN is used for classification of characters present on the number plate. In classification, the class of an object is calculated using the class of its neighbors as

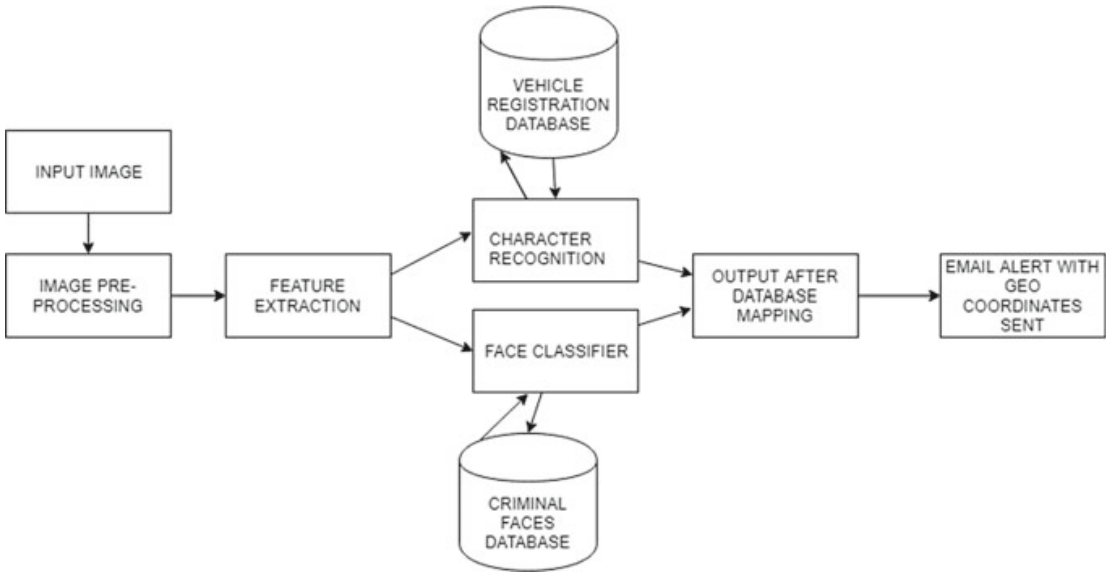


Fig. 1 Proposed method working flow diagram

shown in Fig. 3. It involves two main steps: 1. Character Segmentation [13] and 2. Character Recognition [14]. Characters including alphabets and digits are trained and then recognized by extracting features from the image. The characters are segmented first and then recognized by the algorithm as shown in Fig. 2 (Table 1).

Number Plate Recognition: Steps

Text extracted from the number plate comprises of various steps:

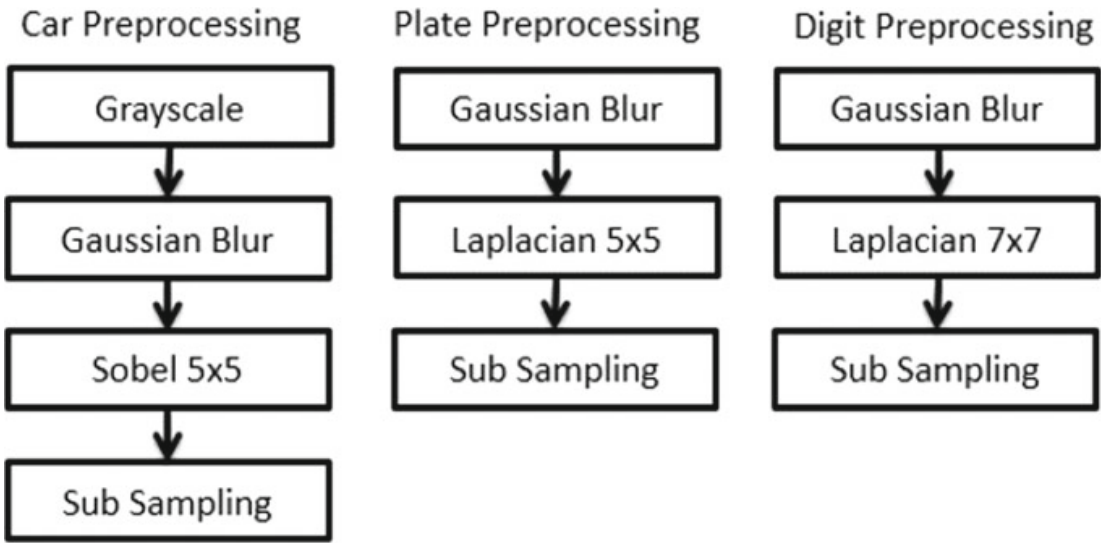


Fig. 2 KNN algorithm flowchart of steps in text recognition

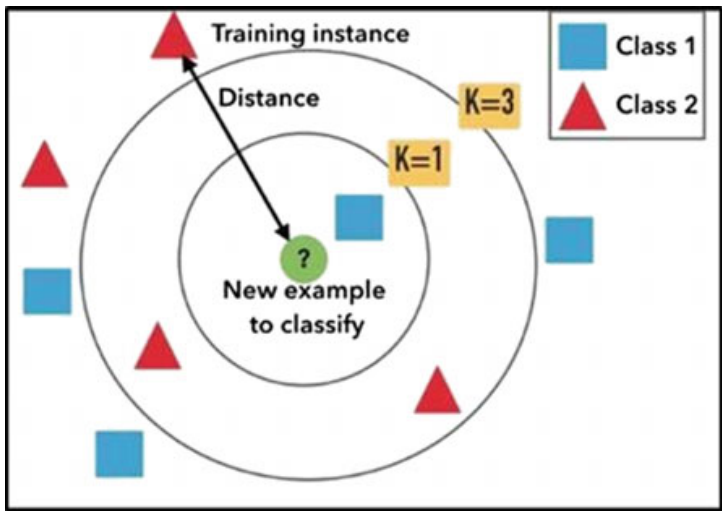


Fig. 3 Green circle is the sample for which classification has to be performed. It has to be assigned either of the two classes, i.e., either class 1 of blue squares or class 2 of red triangles. With $k = 1$, class 1 is assigned to it as its 1 nearest neighbor is a blue square. Now when $k = 3$, class 2 is assigned to the sample because there are 2 triangles in comparison to only 1 square inside the inner circle. *Source* https://cdn-images-1.medium.com/max/800/0*Sk18h9op6uK9EpT8

Table 1 k-NN Algorithm comparison

Pros	Cons
Output is not affected by those sets that are outside the selected boundary	Storing all the training data makes this algorithm computationally expensive
Assumptions are not made by KNN	More memory is required
One of the simplest algorithms	If the value of k is high, then output may be slow
Accuracy is high but not as high as some of the other better supervised learning algorithms	Output is affected by unrelated features and noise

- *Step 1* Capturing the image from a camera.
- *Step 2* Converting an original image to a grayscale [15] image.
- *Step 3* Converting the grayscale image to binary image and applying morphological transforms using OpenCV library for the ease of detection of a license plate from the image.
- *Step 4* Applying Gaussian blur filter [16] to smoothen the image.
- *Step 5* Listing all contours in the image to remove the boundary of the image that has the same color of intensity.
- *Step 6* Applying KNN algorithm to list all matching characters [17] in the image.
- *Step 7* Listing all possible plates that can be found in the image.
- *Step 8* Selecting the highest probability [18] license plate that is matched from the standard license plate measurement and extracting it.
- *Step 9* Reapplying/Repeating steps 2 to 8 to the number plate after extracting it so as to recognize the characters in the number plate. Recognition [19, 20] is done with the help of the KNN algorithm with the training data set. Recognized characters with KNN results in the output [6].

Face Recognition

For face recognition [21], dlib and OpenCV are used to spot all facial landmarks. Relevant regions of the face are represented by facial landmarks [22]. These regions are:

- Eyes
- Nose
- Jaw
- Mouth
- Eyebrows.

Detection of Facial landmarks is a subset problem of the shape prediction problem. In shape prediction problem, the aim is to identify and localize important points of interest.

In the problem of face recognition, facial landmarks are localized by identifying key points in the face. With problems such as face alignment, head pose estimation, face swapping, blink detection, etc., facial landmarks have been effectively used.

For detecting facial landmarks, the following two steps are used:

- *Step 1* In the image face is identified and localized.
- *Step 2* Key facial structures are detected in the face.

Step 1 Different ways of performing Face detection (Step 1):

- OpenCV's built-in Haar cascades.
- Pre-trained HOG + Linear SVM object detector.
- Deep learning-based algorithms designed for face localization.

Irrespective of the algorithm used, our main aim in this step is to find a closed bounding box around the detected face. This is achieved by finding the set of (x, y) coordinates that bound the face.

Step 2 No matter which face landmark detector is used the basic thing they all do is localizing key facial regions. These regions include:

- Right eyebrow
- Left eyebrow
- Right eye
- Left eye
- Mouth
- Nose
- Jaw

- In this method, images in which facial landmarks are manually labeled are used as training images. These specify (*x*–*y*) coordinates of regions surrounding each facial structure.
- Without feature extraction [23], from pixel intensities, the facial landmark positions are projected and a group of regression trees is trained.
- All this results in creating a detector that can detect facial landmarks in real time with good efficiency.

dlib's facial landmark detector.

In the proposed system, the dlib's facial landmark detector is used. It estimates the location of 68 points that locates various facial structures in the human face.

Figure 4 demonstrates the indexes of the 68 coordinates of the human face.

3.1.2 Face Recognition: HAAR Faces Classifier and SVM

Considering face recognition, the complete process can be divided into three main steps:

1. The first step involves finding a decent database of faces which contain several images of each person.

Fig. 4 Picturing the 68 facial landmark coordinates. *Source* https://content.iospress.com/articles/journal-of-intelligent-and-fuzzy-systems/ifs169436

2. The second step is to find faces in the database images and make use of them to get face recognizer trained.
3. The last step is to check face recognizer whether it identifies faces it was trained for or not.

Database

The database is created using webcam capturing images of an individual (Fig. 5).

Face Detection Classifiers

A classifier [24] is a computer program that finds out whether an image is a positive image (face image) or negative image (non- face image). We are using OpenCV [25] for this purpose.

Every file starts with the name of the classifier, it belongs to. For example, a Haar cascade classifier starts off as *haarcascade_frontalface_alt.xml.*

Haar Classifier

The Haar Classifier [4, 5] is an algorithm made by Paul Viola and Michael Jones which uses a machine learning-based approach.

It gets started by digging out Haar features [5] from each image as shown in Fig. 6.

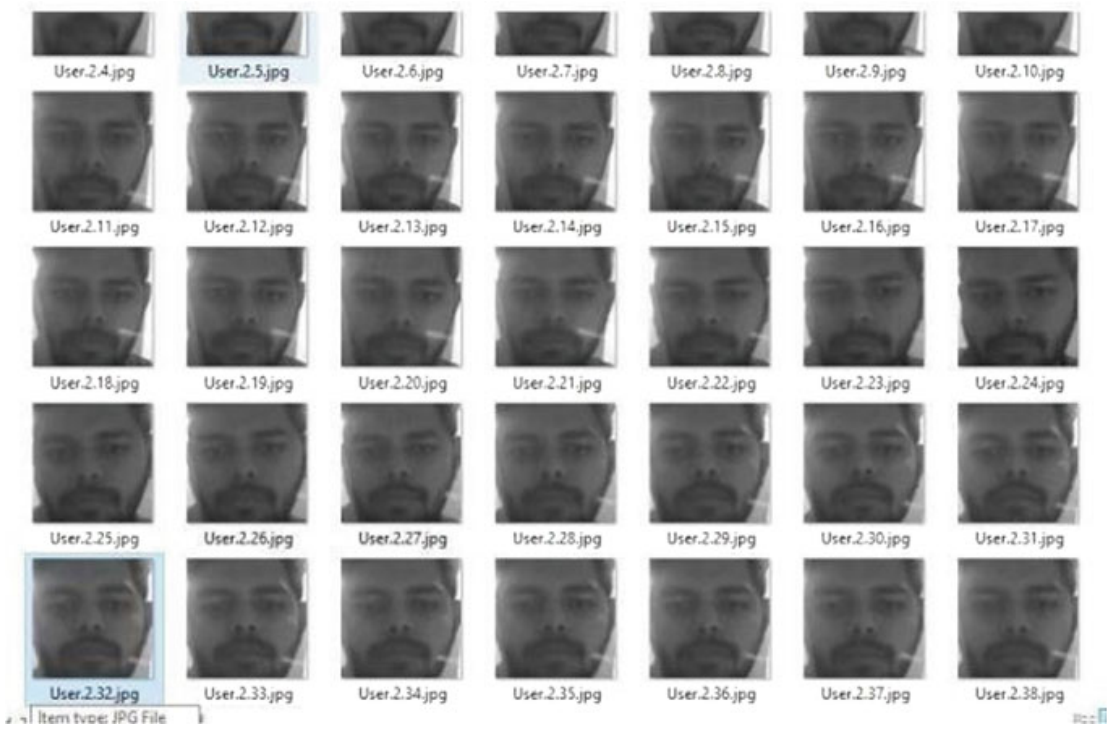


Fig. 5 Image database consisting of 50 images of individual people. Each image has an altered facial expression [JPG]

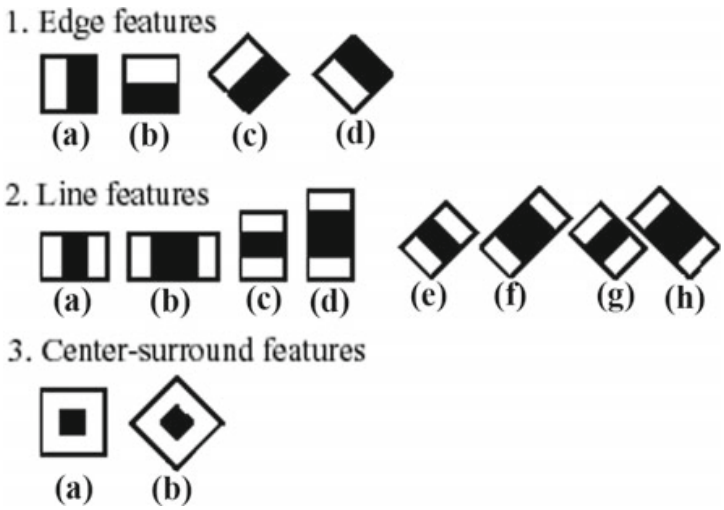


Fig. 6 A couple Haar-like edge, line and center-surround features. *Source* https://docs.opencv.org/2.4/_images/haarfeatures.png

A particular feature is determined by placing each window on the picture. This feature is denoted by a single value which is calculated by subtracting the sum of pixels below the white part of the window from the sum of pixels below the black part of the window [26, 27].

Now, plenty of features are determined by placing all possible sizes of each window on all possible locations of each image. In the end, generally the fact which is considered by the algorithm is non-face region [28, 29] covers most of the area in an image:

- Step 1: Adjusting Tolerance/Sensitivity [30].
- It is done with the tolerance parameter [31]. The default tolerance value is 0.6 and lower the numbers, stricter become face comparisons.
- Step 2: All the faces in an image are automatically found.
- Step 3: The facial features of a person in an image automatically located.
- Step 4: Faces in images are recognized and identified who they are.

4 Result Analysis

The following results were produced using Python programming language. Every detailed step from grayscale conversion to possible vector contours for number plate recognition is shown in the Fig. 9 for the number plate highlighted Fig. 7. Figure 8 shows the predicted output produced on the screen. The result of face recognition is shown in Fig. 10 with the identified faces marked with red squares and the respective names under it. The screen displays Unknown in case the name is not present in the database. The results are much accurate compared to real time objects when tested with different images.

Fig. 7 Actual image of number plate

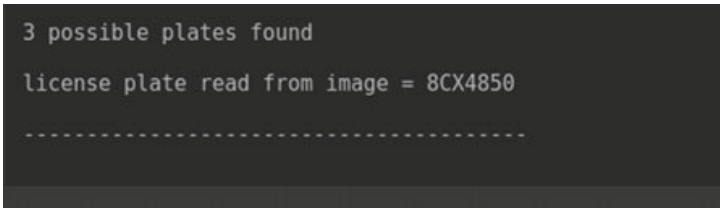


Fig. 8 Predicted output after recognition in python

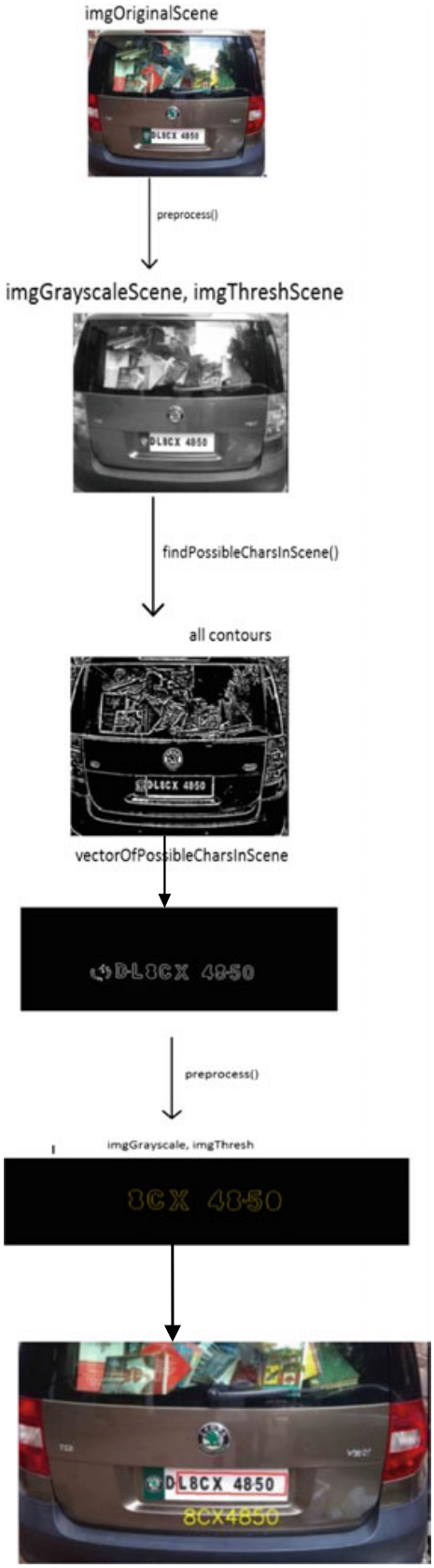


Fig. 9 Procedural steps with the final output

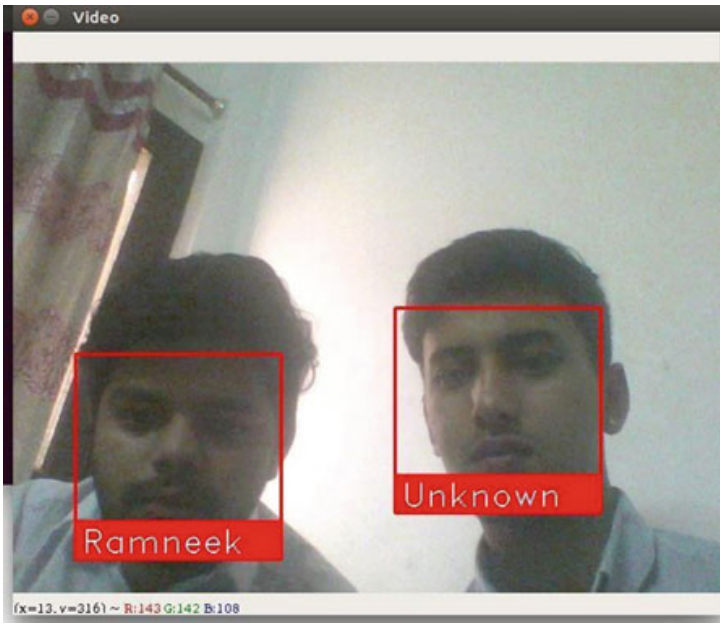


Fig. 10 Final output window of real time face recognition based on image dataset

5 Conclusion

The system for crime detected is improved with much more accurate results and higher efficiency. This system is complete and sophisticated in terms of crime detection. It uses the features of both text and face recognition. Segmenting the characters of the license plate and the English letters and numerals are trained as per the requirement of Indian number plate. Using Gaussian Blur filter in KNN algorithm makes it efficient to deal with all types of images for number plate detection irrespective of noise, intensity, and other factors. It generates multiple plate combinations and chooses one with the highest probability. Similarly, for face recognition, the self-built database is predicting much accurate results in terms of lighting conditions and the movement of persons. The localized facial features based on sensitivity and tolerance parameter in Haar Classifier can also be changed as per requirements. Also, the present values are fit enough to predict well in almost all conditions. The final implementation involves an effective tracking module in the form of a Python Email Script that sends an alert message to the concerned authorities and the nearby police stations and geo-coordinates of the place where the criminal activity took place including the number plate and criminal name detected after matching with their respective databases. The proposed system working with an accuracy of more than 85% successful detections in ambient light conditions.

6 Future Scope

For higher precision and accurate model, certain new features and techniques need to be incorporated to identify faces in all possible conditions of light, barriers in front of the face and aging of facial features with time. An additional phase of the exploration involves, developing a generic model to be installed at all traffic lights in the cities for operating it in dynamic moving traffic.

Acknowledgements This research project is a sole work of the authors mention in this paper and no organization has participated in its contribution. All experimentation performed involving human participants were in accordance with the ethical standards and their consent. There is no violation of copyrights in use of the images of the vehicle number plate. The authors declare that they have no conflict of interest [32, 33].

References

1. Du, S., et al.: Automatic license plate recognition (Alpr): a state- of-the-art review. IEEE Trans. Circuits Syst. Video Technol. **23**(2), 311 (2013)
2. Turk, M.A., Pentland, A.P.: Face recognition using eigenfaces. In: Proceedings of the IEEE, pp. 586–591 (1991)
3. Bachhety, Shivam, Singhal, Ramneek, Rawat, Kshitiz, Joshi, Kunal, Jain, Rachna: Crime detection using text recognition and face recognition. Int. J. Pure Appl. Math. **119**(15), 2797–2807 (2018)
4. Do, H.N., Vo, M.T., Vuong, B.Q., Pham, H.T., Nguyen, A.H.: Automatic license plate recognition using mobile device. In: 2016 International Conference on Advanced Technologies for Communications (ATC), pp. 268–271. IEEE (2016)
5. Gou, C., Wang, K., Yao, Y., Li, Z.: Vehicle license plate recognition based on extremal regions and restricted Boltzmann machines. IEEE Trans. Intell. Transp. Syst. **17**(4), 1096–1107 (2016)
6. Radmanesh, A.: A real time vehicle's license plate recognition system. In: Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance (AVSS'03), vol 4, pp. 159–167 (2015)
7. Wei-gang, Z., et al.: A study of locating vehicle license plate based on color feature and mathematical morphology. In: 6th International Conference on Signal Processing, vol 1, pp. 748-751 (2002)
8. Osuna, E., Freund, R., Girosi., F.: Training support vector machines: an application to face detection. In: Proceedings of the IEEE Conference Computer Vision and Pattern Recognition, June, pp. 130–136, (1997)
9. Schneiderman, H., Kenade, T.: Probabilistic modeling of local appearance and spatial relationships for object recognition. In: Proceedings of the IEEE Conference Computer Vision and Pattern Recognition, pp. 45–51 (1998)
10. Saraswathi, M., et al.: Evaluation of PCA and LDA techniques for face recognition using ORL face database. Int. J. Comput. Sci. Inf. Technol. **6**(1), 810–813 (2015)
11. GPU Based Face Recognition system for authentication, Bhumika Agrawal, et.al, © 2017 IJEDR|Volume 5, Issue 2|ISSN: 2321–9939
12. Yang, J., Luo, L., Qian, J., Tai, Y., Zhang, F., Xu, Y.: Nuclear norm based matrix regression with applications to face recognition with occlusion and illumination changes. IEEE Trans. Pattern Anal. Mach. Intell. **39**(1), 156–171 (2017)

13. Zhai, X., Bensaali, F., Sotudeh, R.: OCR-based neural network for ANPR. In: 2012 IEEE International Conference on Imaging Systems and Techniques Proceedings 2012 Jul 16, IEEE, Jul 16, pp. 393–397 (2012)
14. Chellappa, R., Ni, J., Patel, V.M.: Remote identification of faces: problems, prospects, and progress. Pattern Recogn. Lett. **33**, 1849–1859 (2012)
15. Yetirajam, M., Nayak, M.R., Chattopadhyay, S.: Recognition and classification of broken characters using feed forward neural network to enhance an OCR solution. Int. J. Adv. Res. Comput. Eng. Technol. (IJARCET) **1** (2012)
16. Almudhahka, N.Y., Nixon, M.S., Hare, J.S.: Automatic semantic face recognition. 2017 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017), Washington, DC, 2017, pp. 180–185
17. Lund, W.B., Kennard, D.J., Ringger, E.K.: Combining multiple thresholding binarization values to improve OCR output. Presented in document recognition and retrieval XX Conference 2013, California, USA, 2013. USA: SPIE (2013)
18. Satti, D.A.: Offline Urdu Nastaliq OCR for printed text using analytical approach. MS thesis report Quaid-i-Azam University: Islamabad, Pakistan. p. 141 (2013)
19. Bhansali, M., Kumar, P.: An alternative method for facilitating cheque clearance using smart phones application. Int. J. Appl. Innov. Eng. Manag. (IJAIEM) **2**(1), 211–217 (2013)
20. Roy, A., Ghoshal, D.P.: Number plate Recognition for use in different countries using an improved segmentation. In: 2nd National Conference on Emerging Trends and Applications
21. Zhang, K., Zhang, Z., Li, Z.: Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Process. Lett. **23**(10), 1499–1503 (2016)
22. Anagnostopoulos, C.N.E., Anagnostopoulos, I.E., Loumos, V., Kayafas, E.: A license plate-recognition algorithm for intelligent transportation system applications. IEEE Trans. Intell. Transp. Syst. **7**(3), 377–392 (2006)
23. Erdinc Kocer, H., Kursat Cevik, K.: Artificial neural networks-based vehicle license plate recognition. Proced. Comput. Sci. **3**, 1033–1037 (2011)
24. Bartlett, M.S., Movellan, J.R., Sejnowski, T.J.: Face recognition by independent component analysis. IEEE Trans. Neural Netw. **13**(6), 1450–1466 (2002)
25. Mita, T., Kaneko, T., Hori, O.: Joint Haar-like features for face detection. In: IEEE International Conference on Computer Vision, pp. 1550–5499 (2005)
26. Almudhahka, N.Y., Nixon, M.S., Hare, J.S.: Automatic semantic face recognition. In: 12th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2017), Washington, DC, pp. 180–185 (2017)
27. Chen, Z., Huang, W., Lv, Z.: Towards a face recognition method based on uncorrelated discriminant sparse preserving projection. Multimed. Tools Appl. **76**, 17669 (2017). https://doi.org/10.1007/s11042-015-2882-0
28. Dillak, R.Y., Dana, S., Beily, M.: Face recognition using 3D GLCM and Elman Levenberg recurrent neural network. In: International Seminar on Application for Technology of Information and Communication (ISemantic), pp. 152–156 (2016)
29. Winarno, E., Hadikurniawati, W., Rosso, R.N.: Location based service for presence system using haversine method. In: 2017 International Conference on Innovative and Creative Information Technology (ICITech), pp. 1–4 (2017)
30. Qi, X., Liu, C., Schuckers, S.: CNN based key frame extraction for face in video recognition. In: 2018 IEEE 4th International Conference on Identity Security and Behavior Analysis (ISBA), pp. 1–8 (2018)
31. Deshpande, N.T.: Face detection and recognition using viola- jones algorithm and fusion of PCA and ANN. Adv. Comput. Sci. Technol. **10**(5), 1173–1189 (2017)
32. Zohra, F.T., Gavrilov, A.D., Duran, O.Z., Gavrilova, M.: A linear regression model for estimating facial image quality. In: 2017 IEEE 16th International Conference on Cognitive Informatics and Cognitive Computing (ICCI*CC), pp. 130–138 (2017)
33. Dhawanpatil, T., Joglekar, B.: Face spoofing detection using multiscale local binary pattern approach. In: 2017 International Conference on Computing Communication Control and Automation (ICCUBEA), pp. 1–5 (2017)

Analysis of Classification Algorithms for Breast Cancer Prediction

S. P. Rajamohana, K. Umamaheswari, K. Karunya and R. Deepika

Abstract According to global statistics, breast cancer is the second of all the fatal diseases that cause death. It will cause an adverse effect when left unnoticed for a long time. However, its early diagnosis provides significant treatment, thus improving the prognosis and the chance of survival. Therefore, accurate classification of the benign tumor is necessary in order to improve the living of the people. Thus, precision in the diagnosis of breast cancer has been a significant topic of research. Even though several new methodologies and techniques are proposed machine learning algorithms and artificial intelligence concepts lead to accurate diagnosis, consequently improving the survival rate of women. The major intent of this research work is to summarize various researches done on predicting breast cancer and classifying them using data mining techniques.

Keywords Breast cancer prediction · Classification · Decision tree · Feature selection · K-NN · Random forest · SVM

1 Introduction

A malignant tumor developed from breast cells is termed as breast cancer, and it is the most commonly observed type of tumor in women. Almost 70% of deaths due to cancer arise in poverty-stricken and developing countries. Breast cancer affects about

S. P. Rajamohana (✉) · K. Umamaheswari · K. Karunya · R. Deepika
Department of Information Technology, PSG College of Technology, Coimbatore 641004, India
e-mail: monamohanasp@gmail.com

K. Umamaheswari
e-mail: umakpg@gmail.com

K. Karunya
e-mail: karunyakandimuthu@gmail.com

R. Deepika
e-mail: deepika.rajendran97@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_36

10% of women at some point during their lives. As of late, the death rate continues to expand, with 88% of survival following 5 years from determination and 80% following 120 months from conclusion. There exists a popular method, called knowledge discovery in databases (KDD) which is used by medical researchers to predict the disease outcome by discovering patterns and relationships among the various number of variables stored in the database as historical data. In order to maintain the natural working mechanism of the body, a balance must be maintained between the growth and death rate of cells. But sometimes there is an abnormal and rapid growth of cells which can lead to cancer. Among numerous types of cancer, breast cancer is the major cause for the death of the women worldwide [1]. It is proved that there is a reduction of 38–48% in the mortality rate. There might be several reasons for breast cancer, including age, bosom density, obesity, and changing food habits like alcohol, causing undesirable effects. It is evident from the recent statistics tha, currently the scenario has become worse. The major symptom is a lump observed underarm or in the mammary gland that continues to exist even after the menstrual cycle of a woman. Out of the lump formed, most of them are painless, some may give prickly sensation. The other common symptoms are redness, swollen lymph nodes, or thickening or puckering off the skin. Milk secreting and duct cells are the main cells that cause tumors by draining the milk into the nipple from lobules. A proportion of growth developed from fibrous or fatty tissue can lead to breast cancer. This kind of gene mutation related to cancer is very popular among woman. The treatment for breast cancer includes chemotherapy, hormonal therapy, and radiation therapy. The efficiency result of treatment is comparatively low and requires more attention for preventive measures and control in the current research world. Many breast cancer charities are organizing campaigns such as National Breast Cancer Awareness Month (NBCAM) on the eighth month of every year to bring about awareness of breast cancer in society. This annual international health campaign assists in raising funds for research in breast cancer.

Breast Cancer is categorized as follows:

(i) **Benign (Noncancerous)** The benign cases are those which are noncancerous and non-life threatening. But they could turn into cancerous cells easily. However, these cells could be easily separated.
(ii) **Malignant (Cancerous)** Malignant case leads to abnormal cell growth invading nearby tissues. It is life-threatening in most cases.

Many researchers proposed various data mining algorithms that are deployed for the diagnosis of breast cancer. Among which, feature extraction and classification algorithms employed in it lead to the design of an efficient system. These techniques provide a significant process for extracting the key features which can lead to a proper diagnosis. It is experimentally proven that machine learning and deep learning algorithms are efficient when compared to conventional approaches [2].

2 Breast Cancer Overview

The twentieth century is also called as the cancer century because hundreds of cancer types were discovered in this century. After decades of hard work in analyzing various types of cancers, doctors are now able to identify the causes of these diseases, preventive measures to adopt, and type of treatment to be given. Among all types of cancers existing, breast cancer is rampant among women, very rarely in men. Heredity is a major factor among multiple factors that can cause breast cancer. Nearly 15–20% of women identified with breast cancer has had a recorded occurrence of the same through their generations. In extremely rare cases, a gene called p53 is responsible for breast cancer. The severity of breast cancer was 16 times more than average in families having this type of gene. The number of families with this gene is about 100 all over the world. Researchers have noticed the double risk of breast cancer in individuals producing wet wax of ear glands than in those producing dry wax. Some of the most common symptoms and types are listed in Tables 1 and 2.

Table 1 Breast cancer factors with symptoms

Breast cancer risk factors	Symptoms
Age	• Redness • Swollen lymph nodes • Thickening or puckering of the skin • Scaling of nipple
Family history	
Genetics	
Breast cancer history	• Dimpled skin • Change in the texture of the skin • Breast or nipple pain
Exposed to radiation in the chest or face below the age of 30	
Race/ethnicity	
Being overweight	
Menstrual history	
Drinking alcohol	

Table 2 Breast cancer types

Types of breast cancer	Description
Metastatic breast cancer	Cancer cells break and spread from the original tumor to other parts by means of the lymphatic system or through the blood
Phyllodes tumors of the breast	These are rare case tumors which require surgery to reduce the risk as they can grow rapidly fast. Phyllodes tumors are malignant and borderline benign
Lobular carcinoma in situ	In this type, the growth of cells in milk producing glands is abnormal and increases the person's risk of developing invasive breast cancer
Inflammatory breast cancer	Instead of lump, this type of cancer is found with swelling and reddening in the area of the breast and spreads rapidly to other areas with symptoms worsening
Invasive lobular carcinoma	Also known as infiltrating lobular carcinoma. It starts in duct cells of milk carrying area and spreads beyond it

3 Related Works

There are several studies on breast cancer prediction based on machine learning algorithms.

A. A Novel Approach for Breast Cancer Detection Using Data Mining Techniques [2014]

Chaurasia and Pal [3] compared the performance of the supervised learning classifiers like Naive Bayes, decision tree, and SVM-RBF kernel and simple CART. The dataset that is used for classification is the Wisconsin breast cancer dataset that comprises 11 attributes. The attributes include sample code number, clump thickness, uniformity of cell size, cell shape, and marginal adhesion. The experiment's results proved that the SVM-RBF kernel has the maximum accuracy of about 96.84.

B. Artificial Neural Networks Applied to Survival Prediction in Breast Cancer [2000]

Lundin et al. [4] constructed the classification model using neural network and statistical analysis. The dataset that is used consists of the attributes like age, primary tumor size, axillary nodal status, mitotic count, and tumor necrosis. The dataset contains the details of 951 breast cancer patients, predicted 5-, 10-, and 15-year breast cancer by using ANN and logistic regression models. The accuracy obtained from the constructed model is about 94.35% which is a considerable accuracy.

C. Hybrid Bayesian Network Model for Predicting Breast Cancer Prognosis [2009]

Choi et al. [5] constructed the classification model using artificial Bayesian network. The dataset that is used consists of the attributes such as age at diagnosis, a clinical extension of the tumor, and the number of primary tumors. The dataset consists of the details of 2, 94,275 breast cancer patients. This dataset uses 15 attributes of which seven are primary attributes, seven derived, and one target attribute. The accuracy obtained from the constructed model using the artificial Bayesian network is 88.8%.

D. Predicting Breast Cancer Survivability: A Comparison of Three Data Mining Methods [2004]

Delen et al. [6] refined a prediction model with the help of ANN, decision tree, and logistic regression to envision breast cancer and to determine the survival rate by analyzing the SEER cancer incidence database. The SEER breast cancer dataset consists of 4, 33,272 records and 72 variables. The accuracy obtained from the model is 89.2%.

E. Using Three Machine Learning Techniques for Predicting Breast Cancer Recurrence [2013]

Ahmed et al. [7] analyzed the use of decision trees, SVM, and artificial neural network on breast cancer prediction. The dataset used contains population characteristics and includes 22 input variables. The sample dataset is the data collected from 1189 women. The experimental results found that the SVM produced the least error rate with greater accuracy of 95.7%.

F. Breast Cancer Prediction Using DT-SVM Hybrid Model [2015]

Sivakami [8] developed a model for the breast cancer prediction which uses the Wisconsin breast cancer dataset. The dataset encompasses 699 records, containing 458 belonging to the benign class and the remaining to malignant class. The data is collected from the needle extracts of patients' breasts. Before the prediction process, preprocessing must be done. Preprocessing fills up missing values with attribute's mean value in the dataset. The nine traits of the dataset include uniformity in size of the cell, the shape of cell, and thickness of clump. The implemented algorithm in this work is based on SVM and decision tree. The accuracy obtained from the classification model is 91%, and the error rate is 2.58. The number of precisely classified instances is 459 and imprecisely classified instances are 240.

G. Combining Bagging and Boosting [2007]

Kotsiant et al. [9] worked on various ensemble approaches such as bagging, boosting, and combination of both with a variety of base learners. They are C4.5, Naïve Bayes, OneR, and decision stump. These algorithms used datasets from the UCI repository dataset. The accuracy obtained is about 93.47%.

H. Prediction of Breast Cancer Using Support Vector Machine and K-Nearest Neighbors [2017]

Islam et al. [10] used the most common dataset, Wisconsin breast cancer (WBC) dataset obtained from the repository of UCI machine learning. This dataset consists of 699 instances, where the cases are labeled as either benign or malignant. Classification is performed using SVM and K-NN. The accuracy obtained from the model using SVM and K-NN is 98.57% and 97.14%, respectively.

I. Heterogeneous Classifiers Fusion for Dynamic Breast Cancer Diagnosis Using Weighted Vote-Based Ensemble [2014]

This work proposed an ensemble approach by combining various classifiers including decision tree with Gini index and information gain, Naïve Bayes, SVMn and memory based learner. The classification is decided on the basis of weighted voting. Different datasets are collected from the public repository. Classification accuracy is enhanced using different preprocessing techniques and feature selections. The experimental results conclude that proposed approach contributed to major improvement than other existing classifiers. The accuracy obtained from the ensemble model is about 97.2%, the precision is 100%, and the recall value is 98.60%.

4 Proposed Methodology

4.1 Data Source

The publicly available breast cancer database is used. The database [11] constitutes about 570 records. The dataset consists of numeric attributes. Most of the research papers referred to the 32 attributes present in the dataset for breast cancer prediction.

4.2 Data Preprocessing

The dataset applied in the proposed work is a clinical dataset which may contain many inconsistent, incomplete, or missing data. Such data reduces the accuracy of the model. Hence, preprocessing is an important step that needs to be carried out. To remove the inconsistencies in the dataset, several preprocessing techniques were applied. Also, normalization techniques were applied to handle the missing data. Thus, preprocessing techniques tend to increase the accuracy of the model constructed. The workflow of the proposed model is shown in Fig. 1.

4.3 Random Forest

For the classification of malignant and benign types of cancer, random forest algorithm can be employed. Based on various types of randomization, random forest has

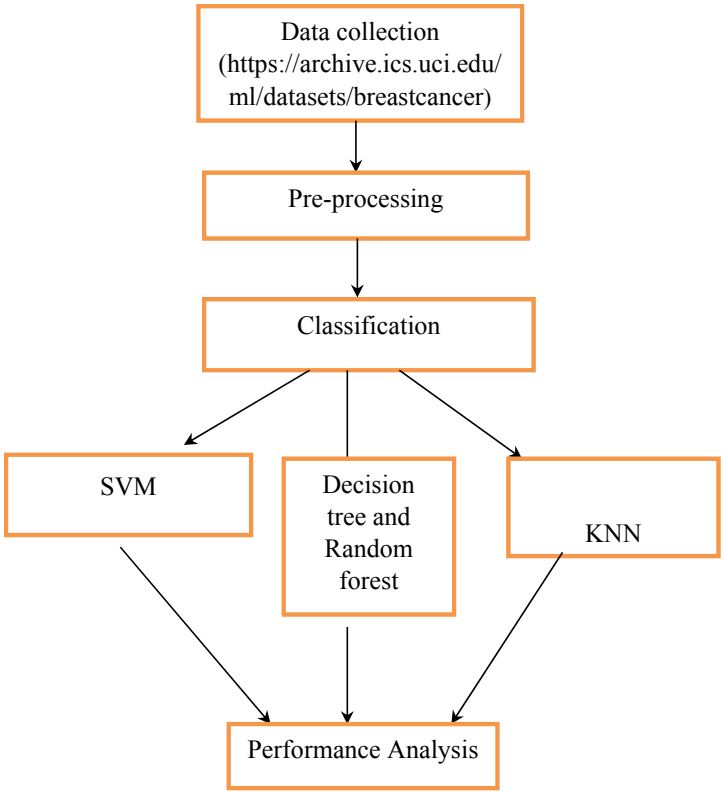


Fig. 1 Workflow of the proposed model

been built as it is an ensemble of decision trees. As the random forests are very flexible rather it is widely used [12]. This supervised learning algorithm creates forests using many trees. The accuracy depends on the number of trees. One of the major advantages with random forest is that it could be functioned as both regression and classification. It can also handle the missing values and it will not overfit in case of a greater number of trees [13]. The random forest takes the test features and predicts the outcome of the randomly created trees based on rules and then stores the result. The votes for each predicted target is calculated, since each tree results in different prediction. Finally, the target receiving high vote will be considered as the final prediction. The random forest processes a huge amount of data at very high speed. The random forest has each tree in binary structure form, which is created based on top-down approach. Generally, the convergence in the random forest algorithm is very fast. The most important parameters are the depth and number of trees. Increasing depth led to an increase in the performance [14]. Thus, random forest is considered as the best classification algorithm on the basis of the processing time and accuracy. This algorithm is implemented on the dataset and the accuracy obtained is 97.34%. The overall view of the random forest algorithm is laid out in Fig. 2.

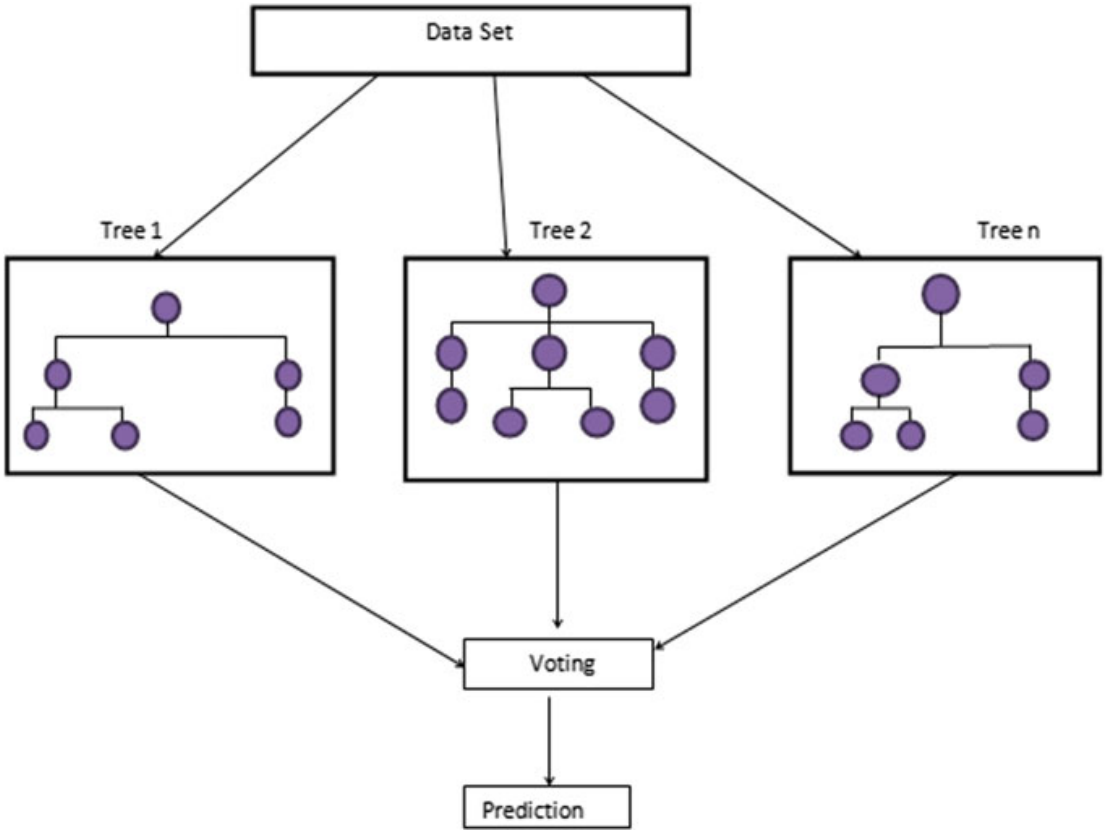


Fig. 2 Random forest

4.4 Support Vector Machine (SVM)

Another powerful supervised learning model is the support vector machine that examines the data for classification and regression analysis using the support of associated learning algorithms. SVM has been specialized for a certain range of problems and gained success in the field of pattern recognition specifically in bioinformatics and cancer diagnosis. In support vector machine, all the data points are drafted in an N-dimensional space [15]. SVM performs linear classification. In addition to this, it also performs nonlinear classification by mapping the inputs implicitly at the range of high-dimensional feature space. The main advantage of SVM is considered to be a unique technique called kernel trick, where lower dimensional space is converted to higher dimensional space and classified. In other words, a hyperplane or set of hyperplanes can be constructed with support vector machine constructs in a high- or infinite-dimensional space for the purpose of classification, regression, or outlier detection [16]. In general, the hyperplane with the largest functional margin achieves good separation as it lowers the generalization error of the classifier. Support vector machine gives an accuracy of 97% when employed on the dataset. The hyperplane used for classification is illustrated in Fig. 3.

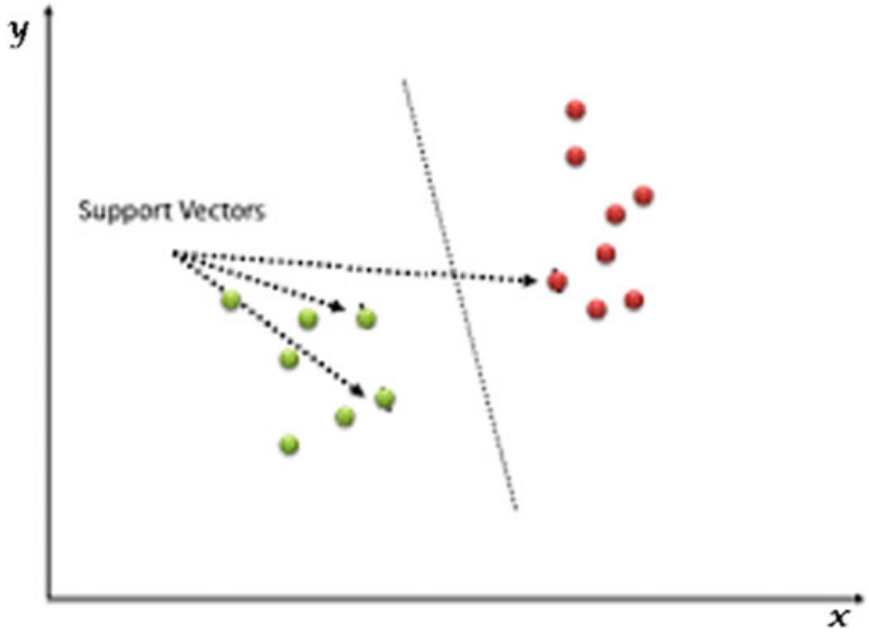


Fig. 3 Support vector machine

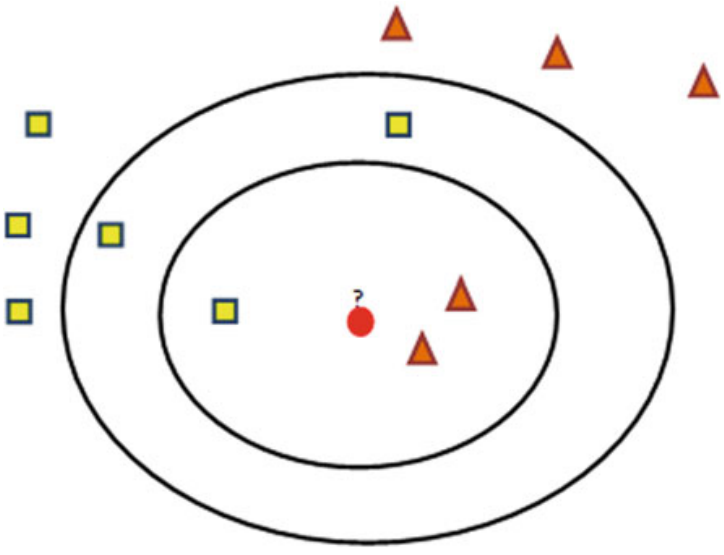
4.5 K-Nearest Neighbor

A nonparametric method, K-nearest neighbors algorithm (K-NN) is another algorithm to be utilized for various purposes like regression and classification. The output of the classification from k-NN is a class label which describes to which class or group it belongs to [17]. In K-NN, the membership is assigned based on the majority vote by the neighbors which is decided by the K value. In other words, each object has been assigned to one class that is most common among the neighbors. For example, if $K = 2$, then the query point is assigned to the class where two nearest neighbors belong to. Being the simplest machine learning algorithm, the explicit training step is not essential. In the training step, neighbors are chosen from the set of objects for which their corresponding classes are known. The algorithm is very sensitive to the local data [18]. Euclidean distance for continuous variables and Hamming distance for discrete variables are most widely used. However, the accuracy can be improved by using specialized algorithms like Large Margin Nearest Neighbor or Neighborhood Components Analysis. The K value is chosen based on the data. Choosing an appropriate K value is significant because that decides if the data is classified correctly. Heuristic techniques like hyperparameter optimization are used because the larger K value reduces the effect of noise but makes less distinct boundaries between classes [19]. The performance gets degraded as the noise in the data increases. The level of accuracy achieved from K-NN is about 95% with an appropriate K value. K-NN is shown in Fig. 4.

4.6 Decision Trees

These are classification algorithms where the attributes in the dataset are recursively partitioned. Decision trees contain many branches and leaf nodes. All the branches tell the conjunction of the attributes that leads to the target class or class labels [20].

Fig. 4 K-nearest neighbor



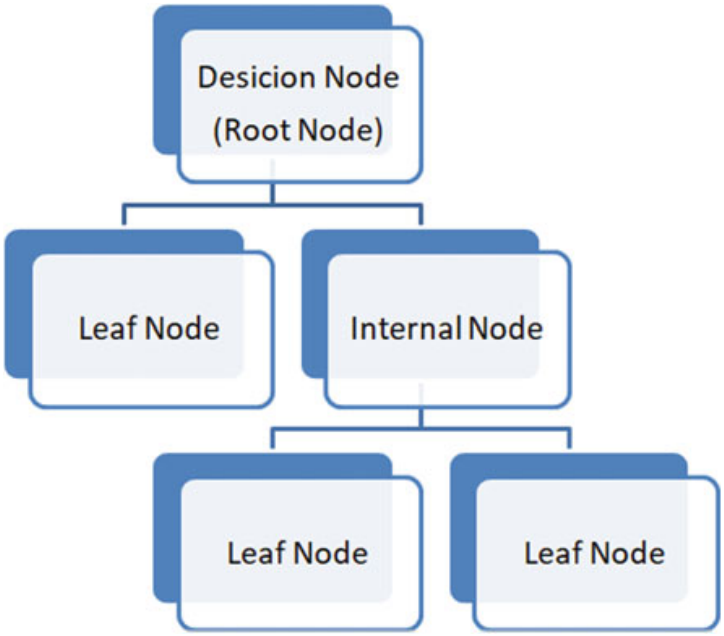


Fig. 5 Decision tree

The leaf nodes contain class labels or the target class that tells to which class tuple it belongs to [21]. There are various decision tree algorithms which can be used for classification of the data. Some algorithms include C4.5, C5, CHAID, ID3, J48, and CART [11, 22].

The Decision Tree can be built as

- The attribute splits decide the attribute to be selected.
- The decisions whether to continue for splitting or to represent as a terminal node is made.
- The assignment of the terminal node to a class.

Information gain, gain ratio, Gini index, etc. are the impurity measures to decide attribute splits done on the tree. After pruning, the tree is checked against noise and overfitting. As a result, the tree becomes an optimized tree. The main advantage of having a tree structure is that it is very easy to understand and interpret. The algorithm is also very robust to the outliers also. The structure of the decision tree is shown in Fig. 5.

5 Experimental Results

Initially, the entire dataset was preprocessed and normalized to handle all the missing values. The mandatory features are extracted using statistical methods. Similar to the experimental results shown in the table above, for the breast cancer prediction, random forests can give a better result that helps the people to get a systematic way of treatment and save their life and hence decreasing the mortality rate. Support vector machines are also equally good with an accuracy of about 91%.

Table 3 Performance analysis of various machine learning algorithms

Algorithm	Accuracy (%)	Precision (%)	Recall (%)
Support vector machine	91	82	81
Decision trees	90.6	83	79
K-nearest neighbor	87	78	76.8
Random forests	93.34	89	86.3

The decision tree and K-NN also give us a good result in predicting breast cancer with an accuracy of 90.6% and 87%, respectively. Hence, a good and accurate model can be built using the algorithms mentioned above. In conclusion, random forest has proven that it is very efficient for breast cancer prediction with an accuracy of 93.34% for diagnosis is exhibited in Table 3. It achieves the finest performance with respect to precision and recall.

6 Conclusion

Most of the Indian women die due to breast cancer [23]. Proper diagnosis of breast cancer is very important in the medical domain. Many models are built using machine learning and data mining methodologies to analyze huge voluminous data. But the challenge with the model is its accuracy and precision for the medical industry. The key for proper treatment and cure is significantly dependent on detecting breast cancer at its early stages [24]. This paper describes how several machine learning algorithms like random forests, decision trees, K-NN, and SVM are employed to model the exact diagnosis of breast cancer for an organized treatment which can save lots of life. The experimental results show the effectiveness of various machine learning algorithms and it proves to be efficient in breast cancer classification thus providing a major breakthrough in clinical diagnosis and treatment of the same. As mentioned earlier, the SVM, K-NN, decision tree, and random forest have their own advantages. Hence, the careful usage of the above will definitely lead to better results. Thus, the appropriate use of a proper algorithm eliminates the risk of death and alleviates the survival rate in women.

References

1. Ferlay, J., Héry, C., Autier, P., Sankaranarayanan, R.: Global burden of breast cancer. In: Li, C. (ed.) Breast Cancer Epidemiology, pp. 1–19. Springer, New York (2010)
2. Sharma, A., Kulshrestha, S., Daniel, S.: Machine learning approaches for breast cancer diagnosis and prognosis. In: Proceedings of the International Conference on Soft Computing and Its Engineering Applications, Changa, India, 1–2 December (2017)

3. Chaurasia, V., Pal, S.: Data mining techniques: to predict and resolve breast cancer survivability. Int. J. Comput. Sci. Mobile Comput. **3**(1), 10–22 (2014)
4. Lundin, M., Lundin, J., Burke, H.B., Toikkanen, S., Pylkkänen, L., Joensuu, H.: Artificial neural networks applied to survival prediction in breast cancer. Oncology **57**(4), 281–286 (1999)
5. Choi, J.P., Han, T.H., Park, R.W.: A hybrid Bayesian network model for predicting breast cancer prognosis. J. Korean Soc. Med. Inform. **15**(1), 49–57 (2009)
6. Delen, D., Walker, G., Kadam, A.: Predicting breast cancer survivability: a comparison of three data mining methods. Artif. Intell. Med. **34**(2), 113–127 (2005)
7. Ahmad, L.G., Eshlaghy, A.T., Poorebrahimi, A., Ebrahimi, M., Razavi, A.R.: Using three machine learning techniques for predicting breast cancer recurrence. J Health Med. Inform. **4** (124), 3 (2013)
8. Sivakami, K., Saraswathi, N.: Mining big data: breast cancer prediction using DT-SVM hybrid model. Int. J. Sci. Eng. Appl. Sci. (IJSEAS) **1**(5), 418–429 (2015)
9. Kotsianti, S.B., Kanellopoulos, D.: Combining bagging, boosting and dagging for classification problems. In: International Conference on Knowledge-Based and Intelligent Information and Engineering Systems, pp. 493–500. Springer, Berlin (2007)
10. Islam, M.M., Iqbal, H., Haque, M.R., Hasan, M.K.: Prediction of breast cancer using support vector machine and K-Nearest neighbors. In: 2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), pp. 226–229. IEEE
11. Barghout, L.: Spatial-taxon information granules as used in iterative fuzzy-decision-making for image segmentation. In: Pedrycz, W., Chen, S.M. (eds.) Granular Computing and Decision-Making, pp. 285–318. Springer, Cham (2015)
12. Winn, J., Shotton, J.: The layout consistent random field for recognizing and segmenting partially occluded objects. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 37–44. IEEE (2006)
13. Breiman, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001)
14. Bosch, A., Zisserman, A., Munoz, X.: Image classification using random forests and ferns. In: IEEE 11th International Conference on Computer Vision, 2007. ICCV 2007, pp. 1–8. IEEE
15. Marcano-Cedeño, A., Quintanilla-Domínguez, J., Andina, D.: WBCD breast cancer database classification applying artificial metaplasticity neural network. Expert Syst. Appl. **38**(8), 9573–9579 (2011)
16. TP, L., Parthiban, L.: Abnormality detection using weighed particle swarm optimization and smooth support vector machine. Biomed. Res. (0970–938X), 28(11) (2017)
17. Mirkes, E.: KNN and Potential Energy (Applet). University of Leicester, Leicester (2011)
18. Breast Cancer Organization: http://www.breastcancer.org/symptoms/
19. Han, J., Pei, J., Kamber, M.: Data Mining: Concepts and Techniques. Elsevier, Amsterdam (2011)
20. Devi, R.D.H., Devi, M.I.: Outlier detection algorithm combined with decision tree classifier for early diagnosis of breast cancer. Int. J. Adv. Eng. Technol. **VII**(II), 98 (2016)
21. Cuingnet, R., Rosso, C., Chupin, M., Lehéricy, S., Dormont, D., Benali, H., Colliot, O.: Spatial regularization of SVM for the detection of diffusion alterations associated with stroke outcome. Med. Image Anal. **15**(5), 729–737 (2011)
22. Bhargava, N., Sharma, G., Bhargava, R., Mathuria, M.: Decision tree analysis on j48 algorithm for data mining. Proc. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **3**(6),1114–1119 (2013)
23. Wolberg, W.H., Mangasarian, O.: UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine
24. Blockeel, H., Struyf, J.: Efficient algorithms for decision tree cross-validation. J. Mach. Learn. Res. **3**, 621–650 (2002)

Real-Time Footfall Prediction Using Weather Data: A Case on Retail Analytics

Garima Makkar

Abstract Be it a retailer, producer, or supplier, the weather has a substantial effect on each one of them. Climate variability and weather patterns have become critical success factors in retail these days. As a matter of fact, weather forecasting has become a $3 billion business now. One of the main reason behind this surge is the capability of the forecasters to sell weather-related information to businesses who then strategize their various decisions regarding inventory, marketing, advertising, etc. accordingly. Hence only those retailers who stay "ahead of the game" will be able to enjoy huge sales while others who do not would face the consequences. Various studies regarding change in consumer behavior occurring due to the change in weather conditions have shown that even a degree change in temperature affects the store's traffic and reflect the growing importance of predictive analytics in this domain. However, these studies incorporate only the historical weather statistics into account. In this paper, we will propose our methodology for footfall analytics to see how the changes in weather conditions will impact the retail store's traffic and thereby retailing value chain, using real-time weather forecasts and footfall data. This analysis provides a platform for retailers to make evidence-driven decisions and strategize their business plan which would help them to deepen the customer involvement and to get efficiency in the planning process.

Keywords Footfall · Retail · Retailers · Real time · Weather

1 Introduction

Climate decides what shall we eat and wear, where shall we spend our holiday, how shall we commute, and even what shall we do every single day. All in all, weather affects four basic purchasing decisions: where, when, what, and in what quantity to purchase [1]. Be it a producer, supplier, or retailer, the weather has a substantial

G. Makkar (✉)
Tata Consultancy Services, Bangalore, India
e-mail: garimamakkar93@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_37

effect on each one of them. From the availability of products and logistics to purchaser demand patterns and sudden spikes, the groundwork for and mitigating the consequences of weather is a real trouble in all aspects of retail sector. Weather conditions affect whether customer would visit brick and mortar outlets or would shop online, making this another important reason for retailers to regulate their operations in order to avoid retail risks associated with different climate patterns. Hence, for retail success, meteorology is as important as geography is. All this have led to the growing importance of weather variability in the retail domain.

Traditionally, it was believed that the retail sector does not get affected much by climate, but a series of recent studies have concluded that retail sector is a weather-sensitive sector. In other words, weather plays an important role in sales of many product categories, store's footfall and store types. Many store managers often blame climate for poor store traffic/sales, but only some of them are able to manage the weather-related risk given that most of the suppliers offer diversified products thus alleviating the effect of weather on store's traffic. Incorporating the effect of weather on consumer behavior demand was considered as a real challenge in the past. For a long time, retailers have been aware of the fact that weather affects the footfall in the stores, but until now, nothing much was being done by most of the retailers to manage such weather-related risks. The rising variability in weather these days have stimulated new interest in studying the relationship between footfall and weather.

Footfall, being a bread and butter for any retail store manager, is directly linked to drive retail sales and improve conversion rates—the more people who come to the store, the more chance they will buy something. While the conditions like heavy rainfall, snowstorm coming, heat and dry conditions, etc., keep the tendency to impact the footfall drastically. So it becomes important to know the footfall peaks and troughs which could help retail managers to optimize in-store operations and manage their resources effectively. Thus we propose a methodology for predicting the footfall of customers in different retail stores given real-time weather forecasts and historical footfall data using a supervised analytical methodology. This paper allows retailers to find the threshold probabilities that will trigger warning and require necessary decisions to decrease inventory waste. Also, this methodology would predict favorable climatic conditions which could be taken care of before-hand in order to achieve retailer's motive of maximizing sale/footfall in real time. Thus, this analysis can be used by retailers to make evidence-driven decisions and strategize their business plan which would help them to deepen the customer involvement and to get efficiency in the planning process.

This paper is divided into the following sections: Sect. 2 highlights the literature survey and critical research gaps. Section 3 defines our problem statement. In Sect. 4, the methodology followed has been proposed. The results are highlighted in Sects. 5 and 6. The conclusion is explained in Sect. 7.

2 Literature Review

This section gives a brief about some of the works that have been done related to retail and weather so far. Generally speaking, these researches can be divided into two groups: (1) theories about the nature of the impact and (2) theories about the magnitude of the impact. Here, impact refers to the impact of weather on consumer activities in a retail store. So the former group explains why and how the change in climate brings the change in retail shopping behavior while the latter explains the extent of the weather change on activities such as consumer spending. The following is the description of some of these theories explaining the consequences which retail sector faces because of the change in the climatic conditions. First, we explain the theoretical work linking weather stimuli to shopper's behavior. Second, we review some of the empirical studies done in this context. Lastly, we review the analytical models predicting the footfall for retail firms during different climatic conditions.

2.1 Theoretical Work

There are different ways by which meteorological factors can affect both psychology and physiology of consumers. Overall, these researches have confirmed that weather has the tendency to influence an individual's mood. For instance, Sanders and Brizzolara [3], studied the effects of various climatic factors such as temperature, biometric pressure, relative humidity, precipitation, and wind speed on a one-dimensional mood rating scale. They concluded that sunshine and biometric pressure together have the strongest impact on mood. Following the same methodology but different mood scales, other researchers like Persinger and Levesque [2], Howarth and Hoffman [4] and Miranda-Moreno and Lahti [5], etc. have reported that rain tends to decrease the sense of comfort of pedestrians impacting consumer's behavior negatively. Similarly, studies by Cunningham [6] and Parrott and Sabini [7] have suggested that sunshine affects customer's mood to visit a retail outlet in a positive way. Experimenters like Sherman et al. [8] and Keller et al. [9], etc., have found that two major factors responsible for moderating psychological effects of weather are the season and the amount of time spent outside and thus pointing toward a positive association between mood and shopping intentions. Robert and John [10] and Gardner and Hill [11] etc., are some of the other contributors of theoretical work done in this context.

2.2 Empirical Studies

All these behavioral instances which are explained in above subsection generally comes from investigational and fieldwork. However, there is also empirical evidence demonstrating the impact of weather on different aspects of retail sector, but the number of works is limited. For example, Steele [12], has presented a paper about weather effects on sales of a departmental store and found that factors like snow, rain, and extreme temperature are responsible for making shopping at store less attractive and thus, affects sales and store traffic negatively. Agnew and Thornes [13], conducted a survey study on the weather sensitivity of food and beverages sector and concluded that in terms of number of customers, the bad weather affects the hypermarkets located in urban areas more adversely as compared to small retail stores located in rural stores. Kirk [1], concluded that about 35% of inquiries in his study would avoid shopping at low temperatures, 37% would avoid going shopping when it is raining, and 30% would not go shopping in excessive heat. Murray et al. [14], further investigated that high temperature is correlated positively to emotions as well as to greater propensity to consume.

2.3 Analytical Work

Various studies regarding change in consumer behavior occurring due to the change in weather conditions have shown that even a degree change in temperature affects the store's traffic, reflecting the growing importance of predictive analytics in this domain. Most of the work on footfall analytics based on weather data has been done using multiple regression analysis. Agnew and Palutikof [15], implemented separate regression models for each month of the year to examine the impact of temperature, precipitation, and sunlight on the total U.K. retail sales and sales of some products categories. Their approach concluded that the sign of the effects of weather variables and percentage of sales volatility is fluctuating between months and product categories. Another regression analysis by Starr [16], concentrating on effect of temperature on the U.K. retail sector, showed that temperature has dual effects, i.e., current and lagged effect in retail. They concluded that weather effect depends on sales assortment and thus, is not same for all retail types. After a year, Parsons [17], examined the correlation between shopping store attendance and weather on a daily basis using a multiple regression model. It was seen that temperature and rainfall are the most tangible weather variables having negative impact on number of visitors while variables like humidity, sunshine hours, etc., show insignificant effect on the store's traffic. However, this analysis was being done for the colder half of the year only. Also, Bahng and Kincade [18] carried out a simple regression analysis to uncover the associations between temperature and apparel retail sales in the U.S. and concluded that weather defines the start and the end of a

season, with summer and winter selling season lasting up to 20 weeks. And any unusual change in weather can delay the beginning of each of these selling seasons.

Thus all in all various researches have been conducted to give prominent result for the effect of weather change on a retail store but all of them links retail and consumer activities to historical weather events. And none of them forecast this relationship in real time. While this paper tends to analyze the correlation between weather and retail shopping behavior (i.e., footfall) in real time and in unique manner.

3 Problem Statement

Weather, being an unmanageable factor, tends to influence shopper's purchasing decisions and causes the revenue to move in any direction. In order to manage the considerable revenue swings caused by fluctuating weather patterns, retail distribution needs to understand the relationship between weather patterns and customer activity. Basically, there are two types of risk which can affect retail performance and retailer's abilities to fulfill customer demands: the risk of understocking and the risk of overstocking. If they miscalculate these, retailers end up losing millions through disappointed customers and incorrect stock levels. Thus, impacting the costs and efficiencies of whole retail business.

It is in this context that the present analysis aims to examine, quantify, and predict the impact of weather factors on retail traffic across various stores located at different regions of the U.S. The aim here is to create an early warning decision-making system relating the retail sector to real-time weather events, preparing both retailers and customers against any weather-related risk. It should be noted that this experiment focuses on weather as an important factor affecting footfall of these retail stores.

4 Analytical Methodology to Predict Footfall

Our experimentation contributes to the field of retail distribution system by giving significant instance of association between meteorological phenomena and retail business. This section provides a recommendation on how to make weather information useful to retailers, through expected footfall levels. The first subsection gives a detailed description about our dataset. Followed by data preparation and experiment analysis.

4.1 Data

For our analysis, the data has been taken from two different sources: (1) OpenWeatherMap, which is a global geographical platform providing various types of earth observation data. 2) Kaggle, also known as "Home for Data Science", is a stage for analytics and predictive modeling competitions. To carry out our experiment we are considering three datasets namely, Footfall/Store data, Key dataset, and Real-time weather API data. The following is the description of each of these datasets:

(a) *Footfall data:* This is a customer's footfall data for 44 hypermarkets (all located in the United States) where footfall may be affected by the change in weather conditions. The footfall dataset provides us detailed information about each hypermarket/store, which can be used to find some interesting facts. For example, The column "station_nbr" corresponds to an id for one of the 10 weather stations covering these 44 locations. It should be noted that some of the stores are nearby and hence share the same weather station. Also, another column called "X3 h" measures the amount of rainfall (in inches) that occurred near each store on a daily basis. Similarly, there are other columns like "temp", "snowfall", "pressure", etc. describing the type of weather conditions near each store location.
(b) *Key data:* The key dataset comprises mainly of two columns: (1) "station_nbr" representing weather stations and (2) "store_nbr", which is an id showing one of the store locations. These columns are then mapped together in such a way that stores nearby would come under the same weather station.
(c) *Real-time weather API data:* This dataset gives us the 5-day weather forecast, i.e., how the weather conditions would be in the coming 5 days. Using an API key, the required dataset would be generated in JSON file format. And for our analysis, we converted this extracted dataset into csv format. The weather API data tells upcoming climate conditions helping retailers to plan their inventory stock accordingly. It should be noted that this data includes weather-related information every 3 h.

The entire dataset captures information about footfalls happening each day in each store for the complete 2 years. And with the count of shoppers, the retailers can easily find out the sales taking place in each store at the time of different climatic conditions. Retailers use various metrics to maximize their sales, with thinking behind making a purchase massively influencing where items are kept in a store, availability of customer's background data and shopper's every move observed as they move around stores. Amid all these factors, the most important factor from the perspective of retailers is "Footfall", which is the count of shoppers who has visited a hypermarket during a given period of time. Thus, using the above-mentioned datasets, we applied real-time footfall analytics to forecast the number of people visiting each store at the time of different climatic conditions.

4.2 Data Exploration and Preparation

In today's world, data comes from various industries, varied sources, and in various formats. The industries can be telecom, insurance, etc. while data sources can be open-source data providing sites, sensor data, social media data, etc. and formats can be .csv, .xls, .txt, etc. Each of these datasets would be having its own set of challenges and thus would be requiring a different series of steps for data preprocessing. Some of the problems which various analysts face while dealing with data include the presence of missing data, outliers, or operational challenges while capturing data and various such other challenges. Due to all these difficulties, it is believed that data preprocessing usually takes 70% of the total time devoted for building the entire model. Since it is the quality of inputs that decide the quality of output so once the hypothesis is built, it is recommended to spend a lot of time and effort performing this step (i.e., exploration, cleaning, and preparation of data).

Keeping this in mind, the second step we did is exploration and analyzing the data chosen for our analysis. Following are the steps performed for understanding, cleaning, and preparing the final dataset which was then used for building our predictive model:

(a) *Variable Identification*: Under this, the first step is to identify independent (Input) and target (Output) variables and then identify the category and datatype of all the variables. The following Fig. 1 shows this variable identification for our dataset:

(b) *Missing Value Treatment*: Missing value/data is defined as the value that is not present for the variable under interest. Being a symbol of messiness in the data, if missing values are not handled properly then it could lead to invalid conclusions about the data. The absence of data can cause several problems like

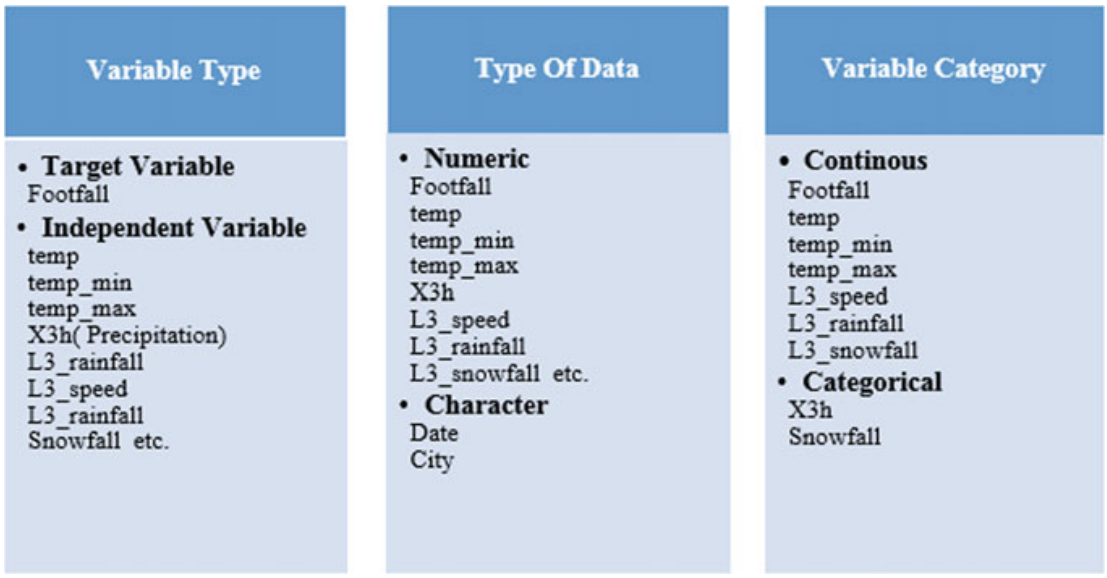


Fig. 1 Example of variable identification (modified and appended from Analytics Vidhya blog [19])

reduction in statistical power of the study, producing biased estimates, complicating the analysis of the study, etc. Thus in order to avoid these distortions, the treatment of missing values is considered as a necessary step in analyzing the data. The dataset which we considered also contained missing values for some of the columns like, "temp_max", "temp_min", "X3 h", "pressure", "snowfall", etc. But different treatments were being done for each of them. For example, missing data in column "snowfall" was handled by the sum of snowfall happened in past 3 days. Similarly, the maximum temperature in the past 3 days was used to replace the missing values present in the column "temp_max."

Hence, by following different treatment methods for each of the above-mentioned columns, we removed all the missing values from our dataset.

(c) *Bivariate Analysis*: This analysis helps to understand the hidden insights present in the dataset. Basically, bivariate analysis is a way of finding the relationship between two or more variables present in the underlying dataset. And since visualization makes understanding of data easier so we used stacked bar charts for the demonstration purpose. The following figures (Figs. 2, 3, 4) show few relationships existed among variables in our dataset:

Due to the presence of detailed information about each store, the footfall data thus can be used to extract numerous useful statistics as shown above.

(d) *Feature Engineering*: The art of extracting more statistics from existing data is known as feature engineering. So it is like making already existing data more useful from analysis point of view. For instance, consider the case of predicting

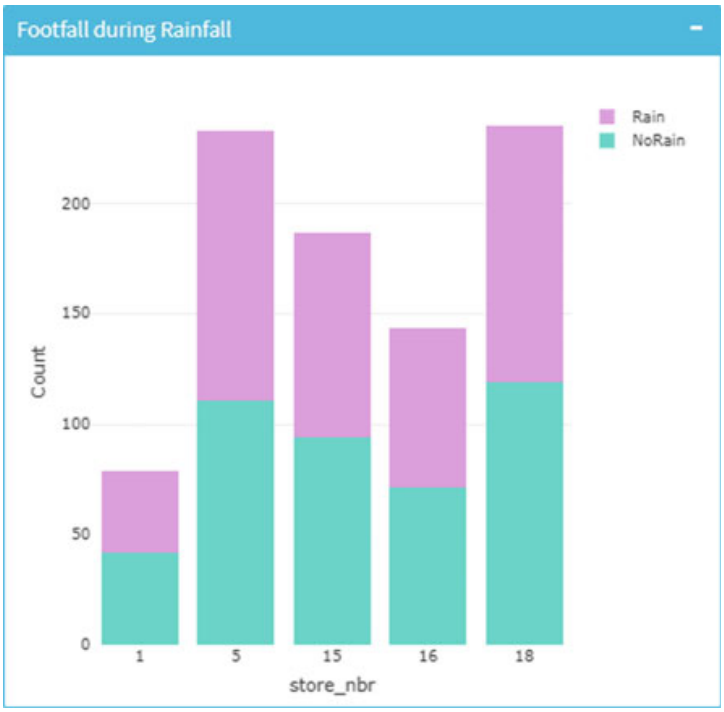


Fig. 2 Footfall during rain and no rain

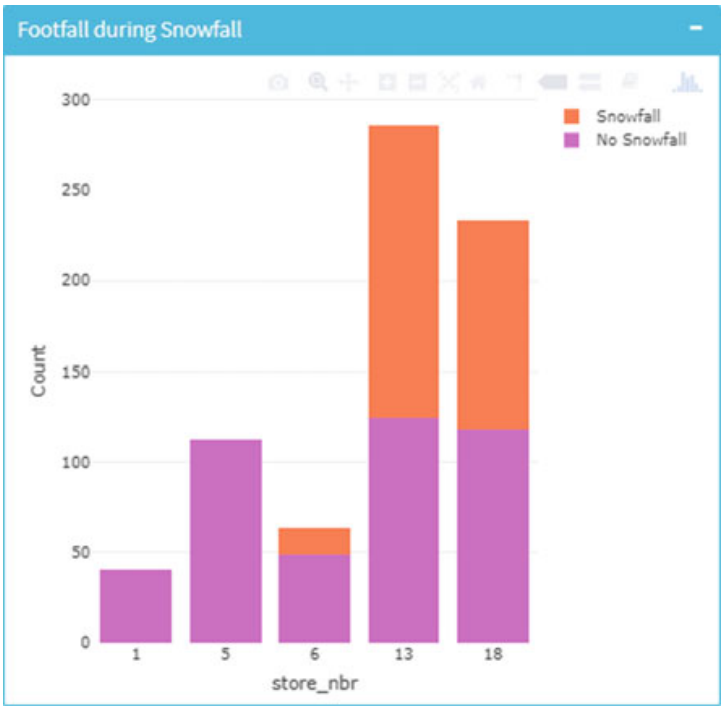


Fig. 3 Footfall during snow and no snow

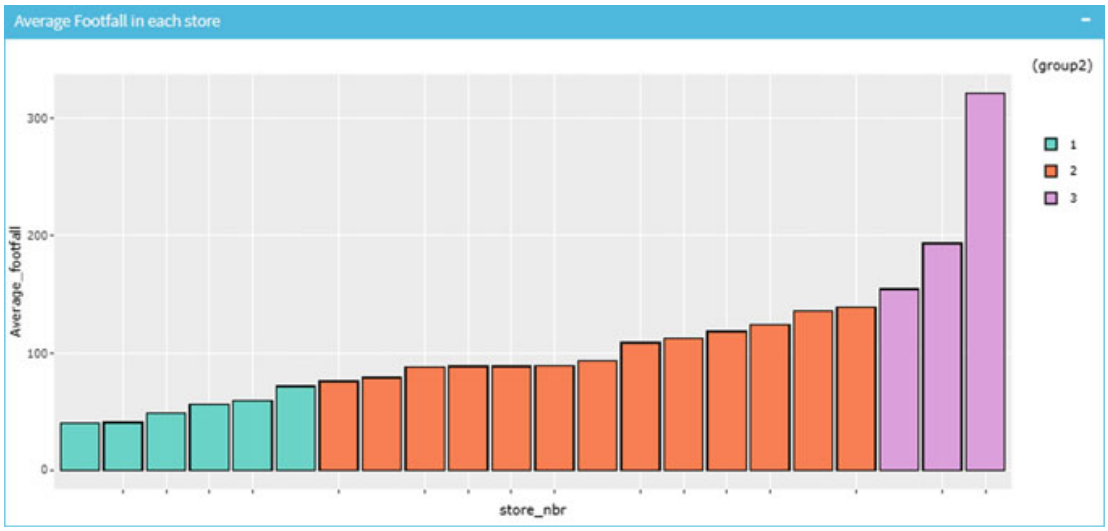


Fig. 4 Average footfall in each store

footfall in a shopping mart on the basis of dates. Here, using dates directly may not yield useful statistics from the data under consideration. This is because count of people visiting a shopping mart is likely to get less affected by the day of the month than the day of the week. And the latter information is present implicitly in the data which needs to be taken out for achieving reliable results. Thus, this activity of taking out information from data is known as "feature Engineering." This step contains two substeps: (1) Variable Transformation and

(2) Feature Creation. In our analysis, we created some new variables from already existing variables. For example, "Date" variable was split into three other new variables—"Day", "Month", and "Year." Also from input variables "temp_max" and "temp_min", new columns called "L3_temp_max" and "L3_temp_min" were being generated. This step was being done thinking that these newly formed variables may have better relationship with the dependent variable and thus could impact the power of prediction remarkably.

In the next section, we will discuss how these steps are combined together so as to predict the footfall based on real-time weather API and store dataset.

4.3 Random Forest as a Supervised Detection Technique

With limited information about climate variability and weather patterns, we solve our problem statement using a supervised machine learning algorithm called random forest. Being a part of ensemble learning algorithm, random forest works by generating multiple models over training dataset where the output of each model is then combined to produce a stronger model.

In the present analysis, this algorithm is applied for regression purpose helping us to map both linear as well as nonlinear relationships existing between the variables. The concept here is to decorrelate the various trees that are produced from separate bootstrapped training data samples. After which, the mean of output of all the trees is used to reduce the overall variance which eventually would help to avoid overfitting. It should be noted that the tree-based algorithms are famous for building models with high stability, accuracy, and easy interpretability.

4.4 Footfall Prediction Based on Weather with Random Forest Technique

In the current study, random forest algorithm has been applied to forecast footfall in different hypermarkets during different climatic conditions. Because of continuous nature of our target variable, regression analysis was being done using random forest. Machine learning algorithms like Support Vector Machine (SVM) etc., could have also been used to solve this problem statement. But the reason for choosing random forest over these techniques is the ease with which it tells the important variables contributing toward regression (and classification) analysis and the relative importance of each of them based on their depth of location in the tree. The following steps explain how random forest for regression works in predicting store's traffic given the weather data:

Each tree is generated as follows:

(i) Random Sample Selection: Divide dataset into train and test. Using around 2/3rd of the training dataset, each tree is trained. This 2/3rd selection is done randomly with replacement and is the training set for that particular tree.
(ii) Random Variable Selection: Each tree is trained using some number of predictor variables (say *m*) which are chosen randomly out of all the independent variables present in the dataset. For regression, the default value of *m* is total number of predictors divided by 3.
(iii) Splitting of node and Parameter Tuning: Each split is done by examining all the variables one by one and then the best split is picked using methods like Gini Index, Chi-square, or Information Gain, etc. Also to get better accuracy important parameters like "mtry", "ntree", etc. are tuned accordingly.
(iv) Calculation of "Out of bag" (OOB) error rate: Calculate the error rate of remaining 1/3rd data for each tree. Aggregate the errors for all the trees to get overall OOB error rate.
(v) Averaging the dependent variable: On the remaining data, each tree is generated giving output equal to some value of dependent variable. Then the average of all these values is taken as the predicted probability.

In the next section, we have leveraged aforesaid approach to showcase some footfall prediction scenarios. There is lot of potentials to extend our methodology to different sectors, different scenarios related to weather and technical approaches.

5 Result and Experimentation

After preprocessing the raw data as explained in Sect. 4.2, we have our dataset ready which could now be used for building the model. With "footfall" as a target variable, we applied random forest algorithm steps (as explained in Sect. 4.4) to carry out regression analysis for the underlying problem statement. It should be noted that the optimal value of parameters such as "mtry" and "ntree" has been found using fivefold cross-validation technique. Following results are obtained from the application of random forest for prediction of footfall in a retail firm:

(i) *Variable Importance Plot*: One of the advantage of random forest technique over other machine learning algorithms is that it tells which input parameters are important for predicting the target variable. Figure 5 shows the top five important parameters for predicting footfall in our model. This plot, called variable importance plot, tells which independent variables contribute more for the variation of target variable. Such information is used to choose the right set of features so as to make target predictions more accurate. It is important to note that these variables does not tell whether there would be more or less

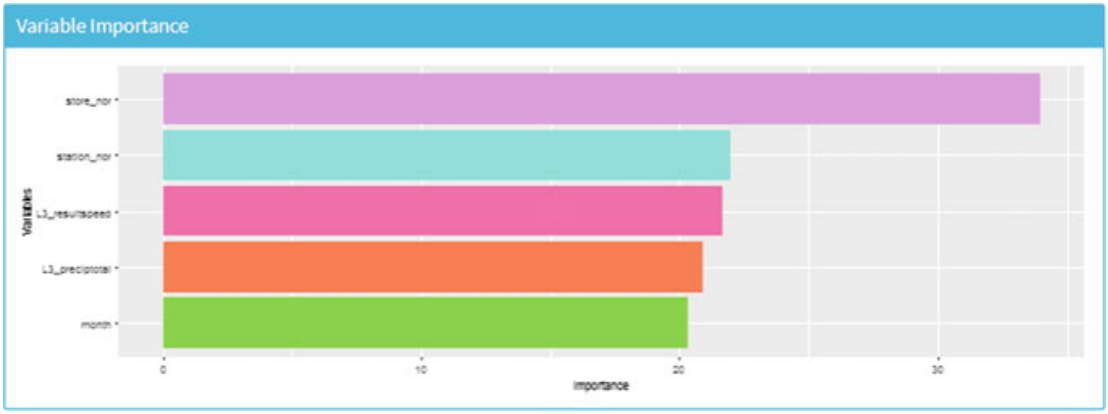


Fig. 5 Variable importance plot

footfall, but tell which all variables are important from prediction perspective. It can be seen that apart from store type, the weather variables such as precipitation and wind speed appear to be the most important parameters in our model.

(ii) *Root Mean Square Error*: The Root Mean Square Error (RMSE) is a performance metric measuring the difference between actual and predicted values of the model. This measure tells us how well our model is able to predict the test outcomes. For our model, RMSE is equal to 11% indicating the good fit of our model, given the spread of our dataset.

Using these two performance metrics, we were able to predict the expected footfall given the weather conditions for all 44 retail stores for the test dataset.

6 Live Demo Model

The above section explains our study for predicting the footfall given the weather conditions near every 44 retail stores. It should be noted that this procedure is based on historical weather data of the U.S. Now to carry out the same experiment in real time, we will make use of 5-day weather forecast data obtained using an API key. Thus, by integrating this new information we are able to predict the expected number of people in each store for the upcoming 5 days over a google map. This live demo model showcase two things: (1) current weather conditions of all the 44 stores present at different locations in the U.S. on a google map and (2) expected footfall curve for the 5 days using which retailers can strategize their activities accordingly. Figure 6 shows this dashboard which we have built using RShiny. It should be noted that the output from this model would keep on changing as the inputs are subjected to real-time factors.

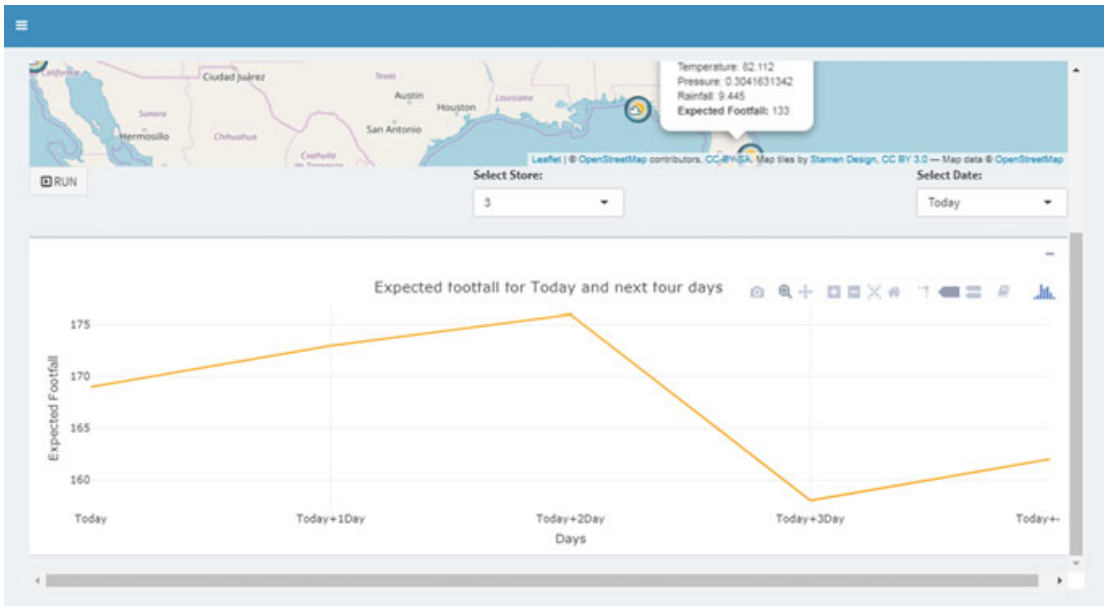


Fig. 6 Real-time demo model

7 Conclusion

The effect of climate variability and weather patterns on retail store's traffic using a supervised machine learning algorithm has been investigated in this paper. Most of the retailers are unaware of the weather consequences on footfall or believe that climate exposure is unmanageable. But, even though weather cannot be managed, the revenue strategies can easily be managed which will secure retailers against any adverse weather conditions. Keeping this in mind, we developed an early warning decision-making system associating retail sector with changing weather events, preparing retailers as well as shoppers against any weather-related risk in this sector. Our methodology put weather in a business context and provides actionable insight into customer behavior, which results in increased footfall levels and improved profitability. Thus, this experimentation is applicable universally and enables comprehensive scenarios of daily footfall traffic to be explored using real-time weather API data, allowing retail strategies to be targeted and implemented effectively.

The major implications of this analysis are as follows:

1. This analysis provides a platform to researchers, practitioners, and decision-makers for sharing and examining how non-catastrophic weather events affect the retail's footfall and the impact it will have in the future(real time).
2. It uses disaggregated data at a large scale and unlike past studies, the lagging effect of weather variables on store's traffic has also been taken into account while building our predictive model.

3. Apart from providing valuable insights about the drivers of footfall, this application can be used as a tool for evaluating how successful the previous events in each of these stores were after taking account of various weather factors.
4. Also, measuring customer's foot traffic and engagement would enable retailers to ensure the right number of staffs as well as products that are required during the impetus of global climate variation.

References

1. Kirk, B.: Better business in any weather. Res. Rev. **12**(2), 28–34 (2005)
2. Persinger, M.A., Levesque, B.F.: Geophysical variables and behavior: XII. The weather matrix accommodates large portions of variance of measured daily mood. Percept. Motor Skills **57**(3), 868–870 (1983)
3. Sanders, J.L., Brizzolara, M.S.: Relationships between weather and mood. J. Gen. Psychol. **107**, 155 (1982)
4. Howarth, E., Hoffman, M.S.: A multidimensional approach to the relationship between mood and weather. Br. J. Psychol. **75**(1), 15–23 (1984)
5. Miranda-Moreno, L.F., Lahti, A.C.: Temporal trends and the effect of weather on pedestrian volumes: a case study of Montreal, Canada. Transp. Res. Part D Transp. Environ. **22**(2013), 54–59 (2013)
6. Cunningham, M.R.: Weather, mood, and helping behavior: quasi experiments with the sunshine samaritan. J. Personal. Soc. Psychol. **37**(11), 1947 (1979)
7. Parrott, W.G., Sabini, J.: Mood and memory under natural conditions: evidence for mood incongruent recall. J. Personal. Soc. Psychol. **59**(2), 321 (1990)
8. Sherman, E., Mathur, A., Smith, R.B.: Store environment and consumer purchase behavior: mediating role of consumer emotions. Psychol. Mark. **14**(4), 361–378 (1997)
9. Keller, M.C., et al.: A warm heart and a clear head: the contingent effects of weather on mood and cognition. Psychol. Sci. **16**(9), 724–731 (2005)
10. Robert, D., John, R.: Store atmosphere: an environmental psychology approach. J. Retail. **58** (1), 34–57 (1982)
11. Gardner, M.P., Hill, R.P.: Consumers' mood states: antecedents and consequences of experiential versus informational strategies for brand choice. Psychol. Mark. **5**(2), 169–182 (1988)
12. Steele, A.T.: Weather's effect on the sales of a department store. J. Mark. **15**(4), 436–443 (1951)
13. Agnew, Maureen D., Thornes, John E.: The weather sensitivity of the UK food retail and distribution industry. Meteorol. Appl. **2**(2), 137–147 (1995)
14. Murray, K.B., et al.: The effect of weather on consumer spending. J. Retail. Consum. Serv. **17** (6), 512–520 (2010)
15. Agnew, M.D., Palutikof, J.P.: The impacts of climate on retailing in the UK with particular reference to the anomalously hot summer of 1995. Int. J. Climatol. J. R. Meteorol. Soci. **19** (13), 1493–1507 (1999)
16. Starr, M.: The effects of weather on retail sales (2000). https://papers.ssrn.com/sol3/papers.cfm?abstract_id=221728
17. Parsons, A.G.: The association between daily weather and daily shopping patterns. Australas. Mark. J. (AMJ) **9**(2), 78–84 (2001)
18. Bahng, Y., Kincade, D.H.: The relationship between temperature and sales: sales data analysis of a retailer of branded women's business wear. Int. J. Retail Distrib. Manag. **40**(6), 410–426 (2012)
19. https://www.analyticsvidhya.com/blog/2016/01/guide-data-exploration/

Normal Pressure Hydrocephalus Detection Using Active Contour Coupled Ensemble Based Classifier

Pallavi Saha, Sankhadeep Chatterjee, Santanu Roy and Soumya Sen

Abstract The Brain plays an imperative role in the life of human being as it manages the communication between sensory organs and muscles. Consequently, any disease related to brain should be detected at an early stage. Abundant accumulation of cerebrospinal fluid in the ventricle results to a brain disorder termed as normal pressure hydrocephalus (NPH). The current study aims to segment the ventricular part from CT brain scans and then perform classification to differentiate between the normal brain and affected brain having NPH. In the proposed method, firstly few preprocessing steps have been carried out to enhance the quality of the input CT brain image and ventricle region is cropped out. Then active contour model is employed to perform segmentation of the ventricle. Features are extracted from the segmented region and Ensemble classifier is used to classify CT brain scan into two classes namely, normal and NPH. More than hundreds of CT brain scans were analyzed during this study; area of ventricle has been used as a measure of feature extraction. Experimental results disclosed a significant improvement in case of ensemble classifier in comparison to Support Vector Machine in terms of its performance.

Keywords NPH · Ensemble · ACM · CSF

P. Saha · S. Chatterjee
Department of Computer Science and Engineering, University of Engineering and Management, Kolkata, West Bengal, India
e-mail: pallavisaha321@gmail.com

S. Chatterjee
e-mail: chatterjeesankhadeep.cu@gmail.com

S. Roy
Department of MCA, Future Institute of Engineering and Management, Kolkata, West Bengal, India
e-mail: santanuroy84@gmail.com

S. Sen (✉)
A. K. Choudhury School of Information Technology, University of Calcutta, Kolkata, India
e-mail: iamsoumyasen@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_38

1 Introduction

Recent studies have revealed biomedical image processing as an application in the detection of a wide range of diseases. It has a great contribution which allows physicians to perform an accurate diagnosis. The human brain is composed of cerebrum, brainstem, and cerebellum. Apart from these parts, there is a part called ventricle which is stuffed with a vivid and colorless fluid regarded as cerebrospinal fluid (CSF). Shape and size of the ventricle depend on the amount of CSF present in the ventricle. Accumulation of excessive amount of CSF in ventricle gives rise to a brain disorder known as NPH. This paper focuses on the detection of NPH brain disease by segmentation [4–6] of ventricle using ACM [2] and then estimates the area of ventricle which is then sent as input to the Ensemble classifier. Ensemble classifier performs classification of the CT brain scan and predicts as either normal brain or affected brain having NPH. This classifier comprises of three weak learners viz. Discriminant, Decision tree, and *K*-th nearest neighbor (KNN). The performance of Ensemble classifier is then compared with SVM and the best classifier is traced out. Features form the principal element in the training stage as well as in the testing stage. The efficiency of a classifier vastly depends on the extraction of features. At times, the best classifier fails to produce proper results if weak features are employed. In the field of image processing, a vast amount of work has already been undertaken by various researchers in the past and brought remarkable changes in the medical history. Classification of brain diseases has been achieved through several methods like Support Vector Machine (SVM), Random Forest, Artificial Neural Network skilled with Genetic Algorithm (GA) termed as ANN-GA, and ANN trained with particle swarm optimization (ANN-PSO). Ensemble classifier has also been applied in the detection of brain disease. Clangphukhieo et al. [1] designed an algorithm for the segmentation of the ventricular part from CT brain image. Initially, CT brain image is normalized and region of interest is estimated by applying gray level profile analysis followed by the Bayesian segmentation. This segmentation enables us to demarcate among the three intensities viz. white matter, gray matter, and CSF and at last the area of CSF is determined. Experimental results disclosed a minimum error of 3.14% and a standard deviation of 1.41 has been achieved by the proposed algorithm. Butman et al. [3] introduced an algorithm to calculate the size of the ventricle in the brain MRI examinations. This has been attained by firstly integrating the different serial images in order to enhance the SNR. Active Contour Model and fast marching techniques were applied to perform segmentation and later deformable registration was used to propagate the segmentation. Zhukov et al. [7] put forward a method based on the concept of a decision tree to estimate the credibility of power systems. Ensemble classifier is used for the classification of the system to determine whether it is in a safe state or not and aggregation is done based on boosting model and random forest model. Shenbagarajan et al. [12] proposed an efficient method to classify an MRI brain image into three types viz. normal, noncancerous, and cancerous. This method involves four phases namely preprocessing steps, segmentation, feature extraction

followed by the classification. Active Contour Model (ACM) is engaged for segmentation and then Artificial Neural Network (ANN) trained with Levenberg–Marquardt (LM) algorithm is used for classification. An enormous amount of research work has been performed in the past in the field of biomedical image processing. Still, it is difficult to meet the growing requirements. In the current study KNN, Decision tree, and Discriminant forming together the Ensemble classifier is implemented to perform classification between a normal brain and affected brain having NPH. The current paper is arranged as follows. Section 2 describes the contrast adjustment, segmentation, and classification model. Section 3 represents the proposed system laid out for the classification of the brain between a normal brain and affected brain having NPH. Section 4 reflects the depiction of experimental results in terms of its performance calculated from the confusion matrix.

2 Methodology

2.1 Image Enhancement and Cropping

Image enhancement forms one of the preprocessing steps that are extensively employed in image processing [15]. The main purpose behind image enhancement is to improve the intelligibility of data in images for viewers. Contrast is an essential factor in the calculation of the quality of an image.

In the current study, the contrast has been adjusted by adjusting the intensity values accordingly in the input CT brain scans to retrieve information efficiently from it. This, in turn, brightens up the image and then the ventricular part is cropped out from the input image.

2.2 Active Contour Models

Active Contour Model (ACM) also termed as snakes is well known in the field of computer vision and is widely applied in applications such as segmentation, edge detection, etc. ACM is classified into two types namely: Edge-based ACM and Region-based ACM [2, 16]. In order to hold back the contour during development for recognizing the boundary of the foreground object, edge-based model makes use of the gradient of the image.

The region-based model utilizes mathematical information regarding the regions lying both inside as well as outside the curve for contour development. In the proposed method, region-based model has been implemented for segmentation which focuses to encourage the curves to arrive at the boundaries of the input CT brain scans.

2.3 Region-Based ACM for CT Brain Segmentation

The region-based model has undertaken with the consideration that the pixel regions of the input image are mathematically similar. It works well on noisy, hazy images, and images having disconnected regions, etc. In the analysis of CT brain scans, region-based model believes the global properties like the length of contour and CT scan image pixel regions are opposite to the local properties like a gradient. The energy minimizing function can be written as:

$$\ln P(I_s|p) = \iint_A {}^0 I_S(x, y)\mathrm{d}A$$

$I_S(x, y)$ represents the intensity value at the pixel location (x, y) in the CT brain image and integral provides the sum total of area A confined by the curve p.

2.4 Feature Extraction Based on Area

The selection of the feature extraction technique plays a leading role in the performance of a classifier [8–11, 17]. In both the training and testing phase, the feature is the crucial element. In the current study, the area of the ventricle in the segmented region forms the measure of feature extraction. Extracted features become the dataset and thereafter enable us to classify between the normal and affected NPH brains.

2.5 Ensemble Classifier

Support Vector Machine (SVM) is a model that is widely applied for classification and regression analysis. The application of SVM in the field of image processing is an example of linear discrimination. The basic working principle of this model is that it divides the space into which the CT scan brain images are assigned into two classes by identifying a separating hyper plane. In 2D space, a line defines the boundary and hyperplane in higher-dimensional space. The inclination towards the application of SVM is that it employs the basis of fundamental risk reduction, which focuses to locate a hyperplane that decreases the area between training classes. The pixels that are nearest to the separating boundary describes the optimal separating hyperplane (OSH) produced by an SVM.

The SVM classifier is trained with the color features extracted from the pixels of several CT scan brain images to differentiate between two regions (classes) namely CSF and non CSF. In the testing phase, an unseen CT scan brain image is selected, then each and every pixel of the testing image is sent as input to the trained SVM

classifier and it assigns a class or a region to each of the pixels in the image as a response. This model is an age-old model and various researchers have adopted this classifier in the field of biomedical image processing to dig out some new innovations, here we have adopted this model and performed a comparison with the proposed model to determine the best classification model in terms of its performance.

2.6 Ensemble Classifier

The foundation of machine learning is the ensemble methods. The amount of diversification among the different learners that exists in the ensemble affects the efficiency of the ensemble to a huge extent. Ensemble learning employs various classification models together to achieve superior predictive performance during testing whereas this level of performance is never achieved by any classification model alone. This model works by integrating many poor learners namely Discriminant, Decision Tree, and KNN. Figure 1 describes the common workflow of ensemble classifier. Once the outcome from each of the weak learners is attained, an aggregation method such as AdaboostM1, Gentleboost, Logitboost, etc., is applied to forecast the final result for the test data.

3 Proposed System

Experiments have been carried out on more than 100 slices of CT brain scans that were gathered from the Apex Hospital situated at Kolkata in West Bengal. Initially, the contrast of the input CT brain image [13, 14] was enhanced and a part of the image was cropped out for later processing steps. Then segmentation of the ventricular part from CT brain scan was employed by using ACM. Features were extracted from the segmented region based on the area of that particular region. These extracted values of area for different CT brain scans together formed the dataset. These dataset were utilized in both the training and testing phase.

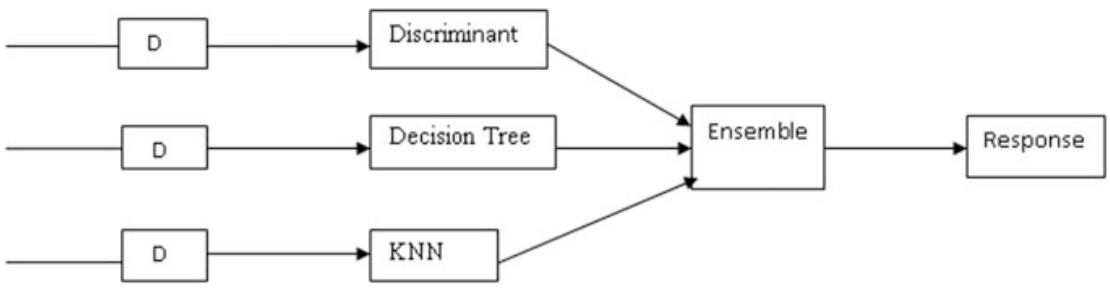


Fig. 1 Represents the workflow of Ensemble classifier

Further on, the extracted dataset is divided into two parts; former part is used for training the classification model and the later part is used during the training stage, depending on the value of data it is classified into either of the two classes namely: normal brain and affected brain having NPH.

In the present study, depending on the training data three weak learners such as Discriminant, Decision Tree, and KNN were trained and the predictions of each model were combined together by using few methods namely; AdaboostM1, Gentleboost, etc. Simultaneously, the training data is also sent to SVM model to train the model and prediction for the test data is noted. In the testing phase, test data was fed as an input to the three classification models and results are recorded and the efficiency of the ensemble classifier is determined and compared with SVM in terms of its performance. The different performance metrics which has been determined from the confusion metrics are as follows (Fig. 2).

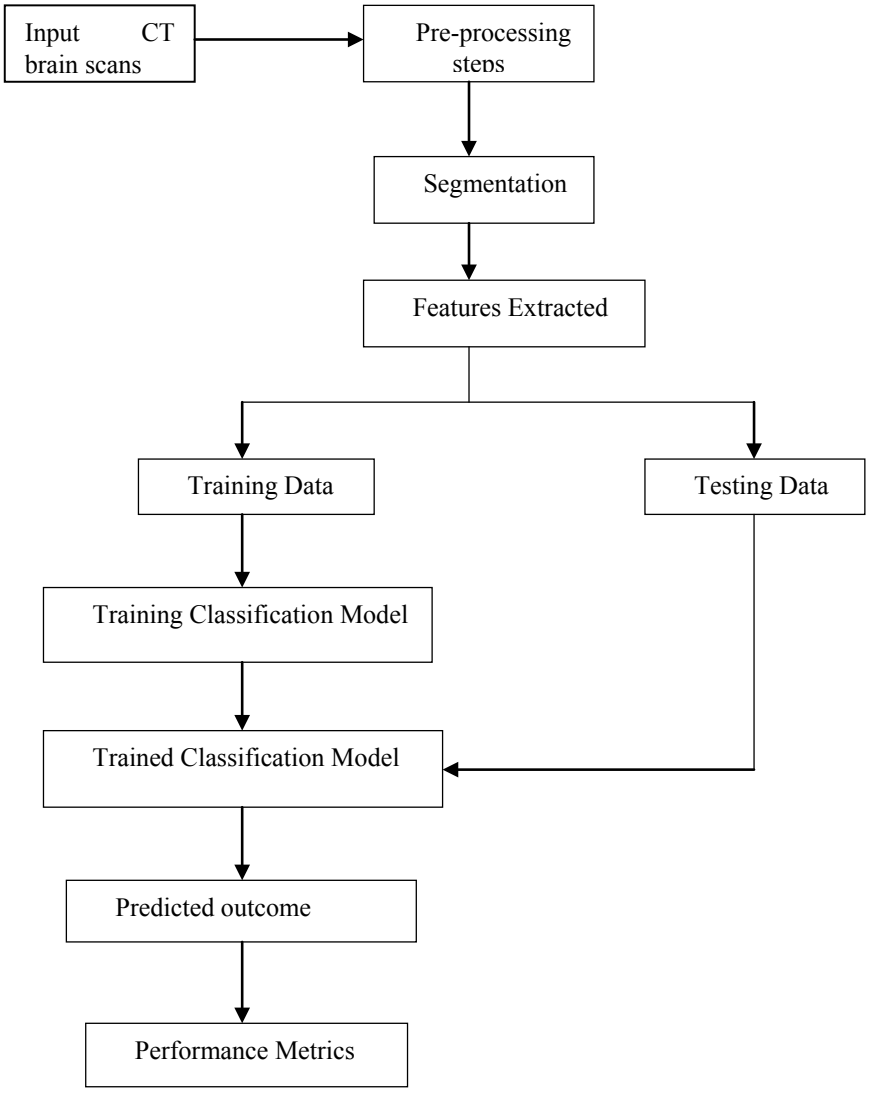


Fig. 2 Depicts the flowchart of the proposed system

$$\text{Accuracy} = \frac{\text{tp} + \text{tn}}{\text{tp} + \text{fp} + \text{fn} + \text{tn}}$$

$$\text{Precision} = \frac{\text{tp}}{\text{tp} + \text{fp}}$$

$$\text{Recall} = \frac{\text{tp}}{\text{tp} + \text{fn}}$$

$$F\text{-measure} = 2 * \frac{\text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

In the above equations, tp represents true positive, tn represents true negative, fp represents false positive and fn represents false negative for testing the efficiency of the classification model.

4 Experimental Results and Discussion

The simulations were carried out on more than 100 CT brain scans by using the proposed method. The resolution of each brain scan is 512 × 512. The dataset comprises of 100 instances among which 60 are normal brain and 40 are affected brain having NPH.

Figures 3a and 4a depicts the input CT brain image, Figs. 3b and 4b represent the enhanced images of the input image. Figures 3c and 4c reflect the cropped part of the input enhanced image, Figs. 3d and 4d depict the segmented ventricle portion from the input CT brain image. Since CSF is stuffed in the ventricle, area of the ventricle can be used to determine the amount of CSF present in the brain which in turn helps to diagnose the condition of the patient. On further calculation, it has been found that the area of the ventricle in Fig. 3d to be around 268.687 mm and Fig. 4d to be 7764.5 mm. Radiologists have also manually calculated the area of ventricle for Fig. 3a which was 268.254 mm.

Moreover, Radiologists have also informed that the CT brain scan to be the healthy and normal brain. The estimated area of the ventricle in a CT brain scan of Fig. 4a is quite greater than the normal brain henceforth it is considered to be affected brain having NPH. Moreover, Radiologist has also determined manually the area of the ventricle in Fig. 4a to be 7764.21 mm. In addition, Radiologist has also intimated that the corresponding CT brain scan to be affected brain having NPH. Table 1 epitomizes the performance of Ensemble and SVM classifier in terms of accuracy, precision, recall, and *F*-measure.

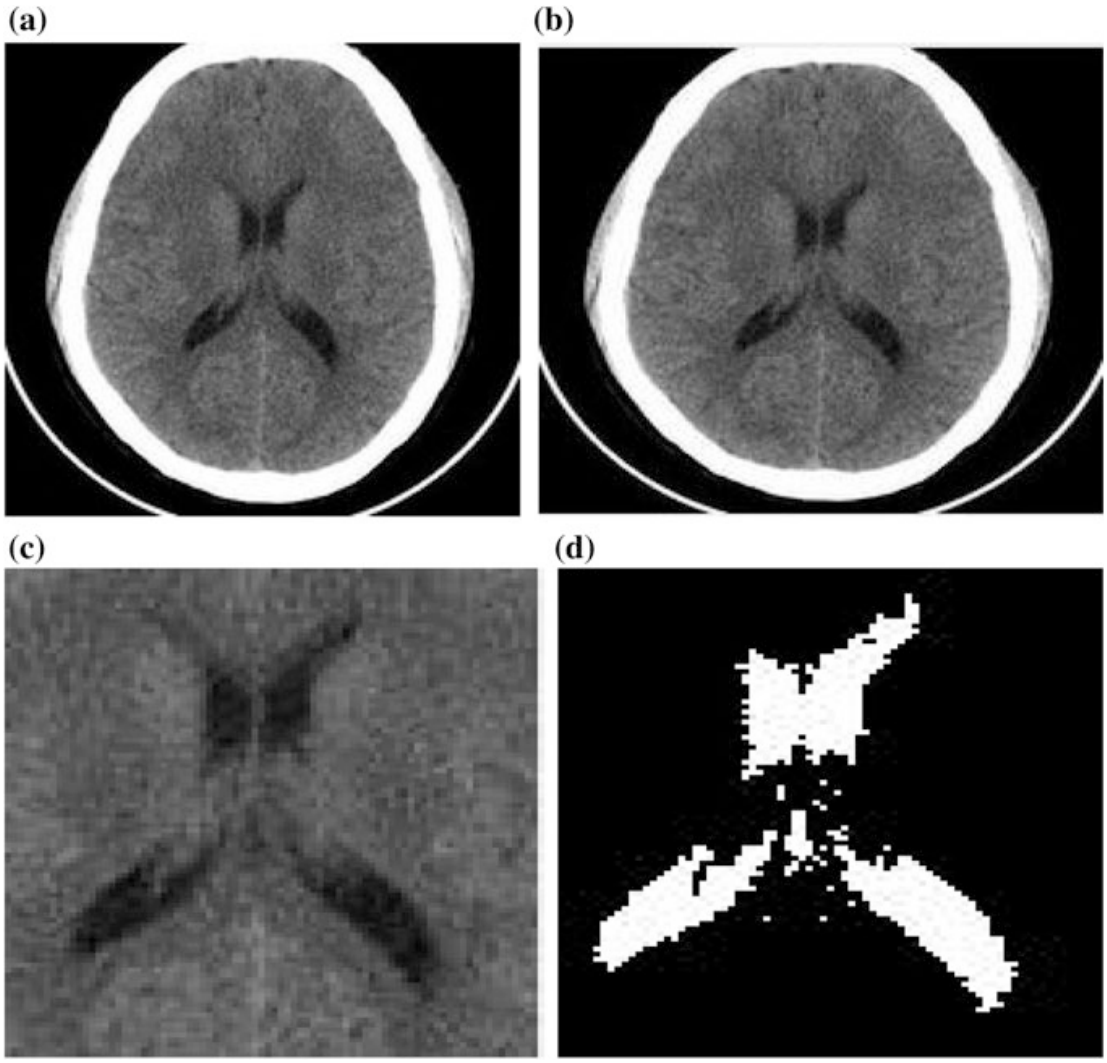


Fig. 3 **a** Source image. **b** Contrast adjustment. **c** Cropped image. **d** Segmented image

Table 1 reports that the accuracy obtained by the ensemble classifier is 93.06% in our proposed method while it achieves precision of 95.65%, recall of 94.28%, and *F*-measure of 94.96%. It has been observed that the performance of SVM is poor in comparison to the Ensemble classifier, achieved an accuracy of 87.12, precision of 88.40, recall of 91.42, and *F*-measure of 89.88.

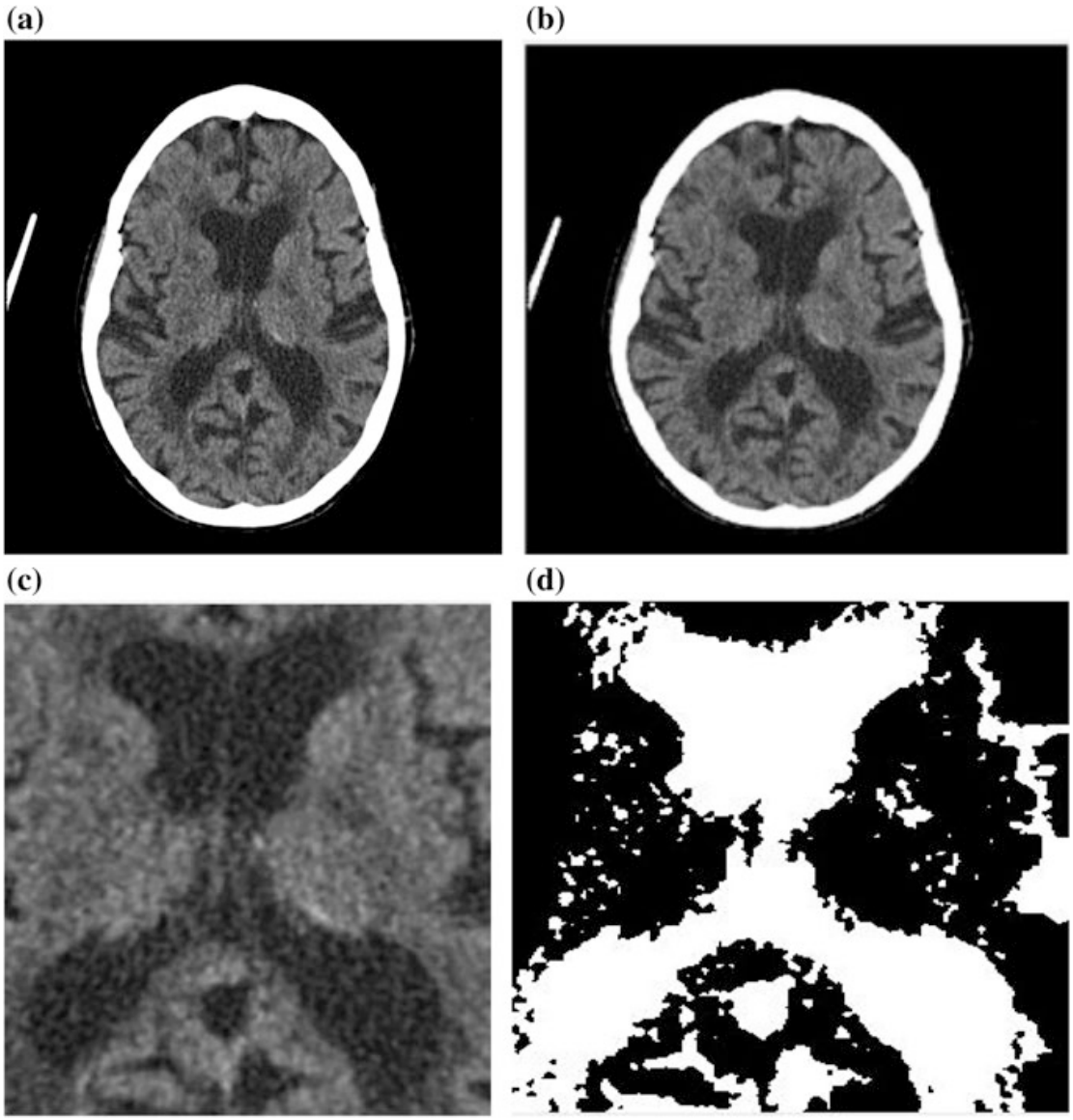


Fig. 4 **a** Source image. **b** Contrast adjustment. **c** Cropped image. **d** Segmented image

Table 1 Performance of the Ensemble classifier (in %)

Metrics	Ensemble	SVM
Accuracy	93.06	87.12
Precision	95.65	88.40
Recall	94.28	91.42
F-measure	94.96	89.88

5 Conclusion

Since the human brain controls many other parts of our body, hence proper functioning of the brain should be ensured primarily. Human brain forms the central organ of the nervous system, so any diseases related to the brain should be diagnosed at an early stage. In the current work, a powerful technique has been put forward for the detection of NPH in the CT brain scans. Experimental results disclosed that ensemble classifier produces better and satisfactory results in comparison to SVM by achieving an accuracy of 93.06%, precision of 95.65%, recall of 94.28%, and *F*-measure of 94.96%.

Acknowledgements We acknowledge Apex Hospital, Kolkata for their support to carry out this research work.

References

1. Clangphukhieo, B., Aimmanee, P., Uyyanonvara, B.: Automated segmentation of a ventricle boundary from CT brain image based on Naive Bayes Classifier. In: Proceeding of 7th International Conference on Computing and Convergence Technology (ICCCT2012), Seoul, South Korea, pp. 1168–1173 (2012)
2. Mostaar, A., Houshyari, M., Badieyan, S.: A novel active contour model for MRI brain segmentation used in radiotherapy treatment planning. Electron. Physician **8**(5), 2443–2451 (2016)
3. Butman, J.A., Lingurarn, M.G.: Assessment of ventricle volume from serial MRI scans in communicating hydrocephalus. In: 5th IEEE International Symposium on Biomedical Imaging, Nano to Macro, pp. 49–52 (2008)
4. Chakraborty, S., Chatterjee, S., Dey, N., Ashour, A.S., Ashour, A.S., Shi, F., Mali, K.: Modified cuckoo search algorithm in microscopic image segmentation of hippocampus. Microsc. Res. Tech. **80**(10), 1051–1072 (2017)
5. Aung, P.T.T., Khaing, A.S., Tun, H.M.: MR brain image segmentation using region based active contour model. Int. J. Sci. Technol. Res. **5**(06), 92–97 (2016)
6. Karimia, H., Esfahanimehrb, A., Moslehb, M., Mohammadianjadvalghadamc, F., Salehpourc, S., Medhatia, O.: Persian handwritten digit recognition using ensemble classifiers. In: The International Conference on Advanced Wireless, Information, and Communication Technologies (AWICT 2015)
7. Zhukov, A., Tomin, N., Kurbatsky, V., Sidorov, D., Panasetsky, D., Foley, A.: Ensemble methods of classification for power systems security assessment. Appl. Comput. Inform. **15**, 45–53 (2019)
8. Chatterjee, S., Dey, N., Shi, F., Ashour, A.S., Fong, S.J., Sen, S.: Clinical application of modified bag-of-features coupled with hybrid neural-based classifier in dengue fever classification using gene expression data. Med. Biol. Eng. Comput. **56**(4), 709–720 (2018)
9. Bijalwana, A., Chand, N., Pilli, E.S., Rama Krishna, C.: Botnet analysis using ensemble classifier. Recent Trends Eng. Mater. Sci. **8**, 502–504 (2016)
10. Chakraborty, S., Chatterjee, S., Ashour, A.S., Mali, K., Dey, N.: Intelligent computing in medical imaging: a study. In: Dey, N. (ed.) Advancements in Applied Metaheuristic Computing, pp. 143–163. IGI Global, Hershey (2018)

11. Farhan, S., Fahiem, M.A., Tauseef, H.: An ensemble-of-classifiers based approach for early diagnosis of Alzheimer's disease: classification using structural features of brain images. Comput. Math. Methods Med. (2014). https://doi.org/10.1155/2014/862307
12. Shenbagarajan, A., Ramalingam, V., Balasubramanian, C., Palanivel, S.: Tumor diagnosis in MRI brain image using ACM segmentation and ANN-LM classification techniques. Indian J. Sci. Technol. **9**, 1–12 (2016)
13. Xia, Y., Ji, Z., Zhang, Y.: Brain MRI image segmentation based on learning local variational Gaussian mixture models. J. Neurocomput. **204**(C), 189–197 (2016)
14. Chakraborty, S., Chatterjee, S., Dey, N., Ashour, A.S., Shi, F.: Gradient approximation in retinal blood vessel segmentation. In: 2017 4th IEEE Uttar Pradesh Section International Conference on Electrical, Computer and Electronics (UPCON), pp. 618–623. IEEE (2017)
15. Kaur, S., Ritika, R.: Contrast enhancement techniques for images—a visual analysis. Int. J. Comput. Appl. **64**(17), 20–25 (2013)
16. François, L., Besson, S.J., Fadili, J., Aubert, G., Revenu, M., Saloux, E.: Region-based active contour with noise and shape priors. In: IEEE International Conference on Image Processing; Atlanta, GA, pp. 1649–1652 (2006)
17. Chatterjee, S., Dey, N., Sen, S.: Soil moisture quantity prediction using optimized neural supported model for sustainable agricultural applications. Sustain. Comput. Inform. Syst. (2018). https://doi.org/10.1016/j.suscom.2018.09.002

Question–Answer System on Episodic Data Using Recurrent Neural Networks (RNN)

Vineet Yadav, Vishnu Bharadwaj, Alok Bhatt and Ayush Rawal

Abstract Data comprehension is one of the key applications of question-answer systems. This involves a closed-domain answering system where a system can answer questions based on the given data. Previously people have used methods such as part of speech tagging and named entity recognition for such problems but those methods have struggled to produce accurate results since they have no information retention mechanisms. Deep learning and specifically recurrent neural networks based methods such as long short-term memory have been shown to be successful in creating accurate answering systems. This paper focuses on episodic memory where certain facts are aggregated in the form of a story and a question is asked related to a certain object in the story and a single fact present is given as answer. The paper compares the performance of these algorithms on benchmark dataset and provides guidelines on parameter tuning to obtain maximum accuracy. High accuracy (80% and above) was achieved on three tasks out of four.

Keywords Episodic memory · LSTM · Memory networks · Question answering system · Recurrent neural networks

V. Yadav · V. Bharadwaj · A. Bhatt (✉) · A. Rawal
Research and Innovation Lab, Tata Consultancy Services, Bangalore, India
e-mail: alok.bhatt@tcs.com

V. Yadav
e-mail: vineet.y@tcs.com

V. Bharadwaj
e-mail: vishnu.bharadwaj@tcs.com

A. Rawal
e-mail: ayush.rawal@tcs.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_39

1 Introduction

Question answering is a complex NLP task which involves understanding of the meaning of text and after understanding, ability to reason about important facts [1]. This becomes tricky as computer does not infer like humans given text data and a set of questions to answer. The problem is very specific and involves answering factual questions, for example, given a paragraph on the US Administration and being asked: "Who is Donald Trump?".

A lot of interest is being shown by organizations in creating question–answer systems that can accurately answer simple questions without the use of a lot of training data. The focus is on obtaining the best possible approach to certain problems in the benchmark bAbI dataset using LSTM.

The bAbI dataset consists of total 20 tasks. These tasks are related to support facts about a question for which an algorithm should provide an answer. These supporting facts are single supporting fact, e.g., Mary traveled to office. Where is Mary? Two or three supporting facts, e.g., "John is in the playground" and "John picked up the football". Where is the football?

This problem involves learning from a large data which is in episodic form and has facts which are related to the asked question. An attempt has been made to solve this using a modern neural network architecture which uses the facts to train and is able to produce the specific answer.

2 Related Work

We have seen the emergence of QA-based strategies in the field of language understanding. QA is easy to evaluate, especially for those questions which have true/false or yes/no answers. These questions should be unambiguous which a human can understand and answer. Memory network proposed by Weston et al. [2] is a new network architecture called memory network which consists of a memory component m and four components I [Input feature map], G [Generalization component], O [Output feature map], and R [Response component] which are potentially learned.

Another work on Large-scale Simple Question Answering with Memory Networks by Bordes et al. [3] proves that question answering systems can handle millions of data points very well and that training on two related datasets enhances overall accuracy. It involves the process of generating the most similar candidates through n-gram matching and getting the best candidate through finding cosine similarity.

A work on end-to-end memory network by Sainbayar Sukhbaatar et al. [4] which can be seen as an extension of RNN search having multiple computational steps known as hops per output symbol. This model can be considered as a

continuous form of memory network which can be trained end-to-end with less supervision on an input–output pair. It also performs well on various tasks from question–answer task to language modeling.

3 Scope of the Work

Four tasks are selected and are mentioned below to build question–answer models. The selection was done with the objective of having unique unrelated contexts.

Task1: Single supporting fact
Task4: Two logical arguments
Task7: Counting
Task17: Positional Reasoning

The detailed description of these tasks is provided alongside model results. We have also conducted error analysis on Task 1 to study the algorithm in depth.

4 Our Approach

A deep learning Recurrent Neural Networks (RNN) architecture is designed which uses the concept of LSTM. The data used to train the models is the bAbI 10,000 sample dataset.

RNN is the preferred architecture for sequential information. In a standard neural network, there is an assumption that all inputs and outputs are independent of each other. Practically this assumption may not always hold true. To predict the next word in a sentence it would help to know the words appearing previously. RNN are known as recurrent neural networks because they perform the same task on every element of a sequence, where output dependents on computations of previous steps. RNN has memory which captures information about the computations of previous steps. In theory, RNN has the potential to use information present in long sequences, but practically it is only applicable for looking back at few steps (Fig. 1).

LSTM belongs to RNN family which has the capability of persisting information in longer sequence, which RNN lacks in implementation [5]. Key element of LSTM is known as cell state. This cell state keeps track of long-term sequential information [6]. LSTM neuron has three gates input, output, and forget. These gates regulate the interaction of information with RNN data flow (Fig. 2).

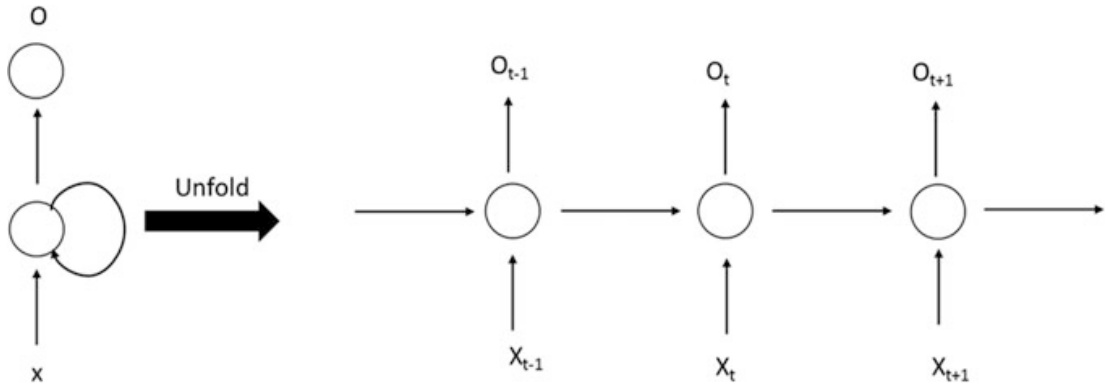


Fig. 1 Recurrent neural network architecture after unfolding the loop

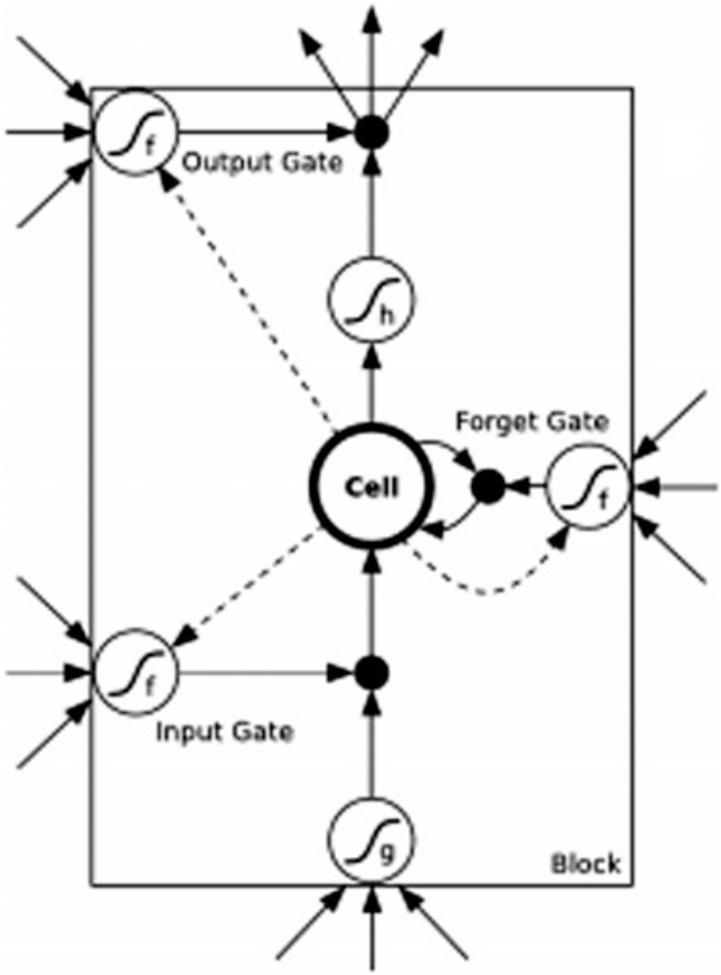


Fig. 2 LSTM cell with input, output, and forget gate which helps in passing specific information to the next level

4.1 Solution Architecture

The solution architecture used incorporates the memory network and has certain connections that help to emphasize the most appropriate words to pick the answer.

This consists of mainly four layers:

1. Embedding layer: It converts each word of dictionary to a vector.
2. LSTM: This is a type of RNN which is used for sequence data and overcomes short-term memory of regular RNN. Memory network is being used in this layer.
3. Dropout layer: It will shut off few neurons from a layer based on user-defined criterion and thus prevents model overfitting.
4. Dense layer: It is a layer with softmax activation [7], which will produce answer for a specific question by choosing the option with highest probability (Fig. 3).

First each question and story gets converted to embedding. After this, allow the network to follow the question based on story and give attention to certain most relevant words in the question; dot product between both the embedding was taken [8].

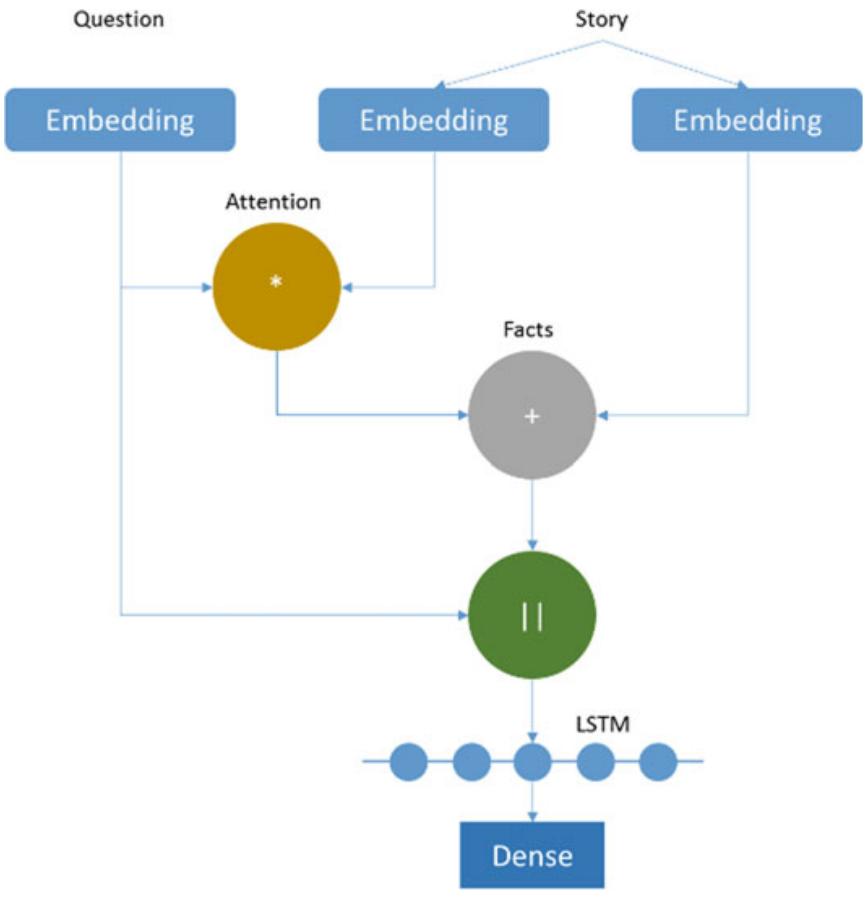


Fig. 3 Solution architecture involving embedding layers to dense layer that produces answer for a question given a story

This produces a vector which will have more weights to the words which are more relevant based on story and questions. This forms the network memory. This attention gets combined with the story embedding to find out facts about the question. These facts are then combined with questions and this information is fed into LSTM network to produce answer for a question.

4.2 *Memory Network*

This is the key component of the solution architecture and it controls information flow using gates. Memory networks [9] architecture enables them to store all the state from a given input, which means that queries can be asked and the memory is able to go many steps back to find the most suitable portions that need to be given greater "attention" as illustrated in the figure below (Fig. 4).

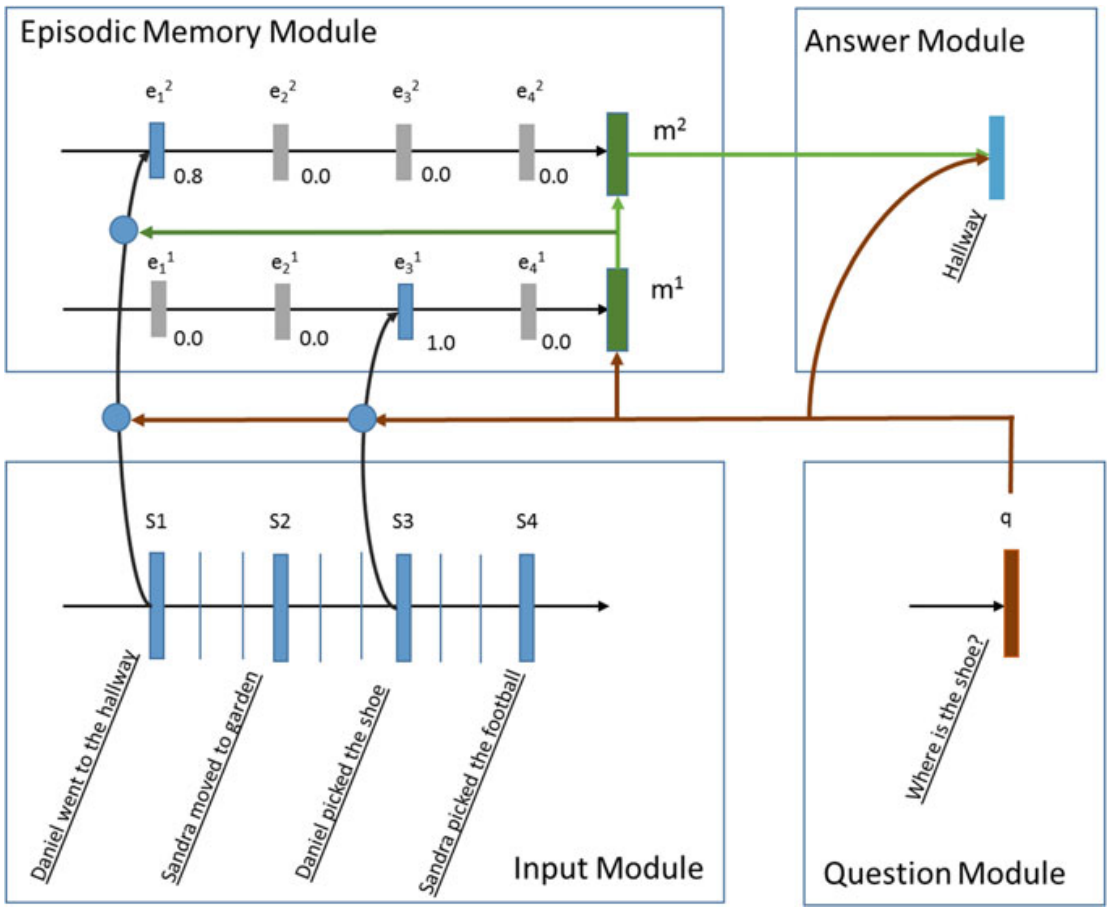


Fig. 4 Illustration showing memory network showing how relevant portions are being given more attention

4.3 Task Description and Model Results

This memory network was trained using different tuning parameters such as embedding size and LSTM size described below.

Embedding Size: This is the size of the Word2vec—word convert to a numerical vector.
Latent Size: Number of neurons in the LSTM layer.
Batch Size: Number of inputs processed together.
Epochs: Number of passes of forward and backward propagation in the deep network.

Changing dropout rate did not have any significant impact on model accuracy. Below are descriptions of each task and the results for different combinations and with 30% dropout for every model.

Task 1: This task consists of a story where answer is based on single supporting fact. For example,
Story: Dan went to the bedroom. Sunny moved to the bathroom. Q: Where is Dan? A: bedroom.
Results for the task (Fig. 5, Table 1).

Epochs were not contributing to increased accuracy in a significant manner. Small latent size with less number of epochs is producing the best accuracy.
Task 4: Answer for this task will depend upon understanding two logical arguments. For example,

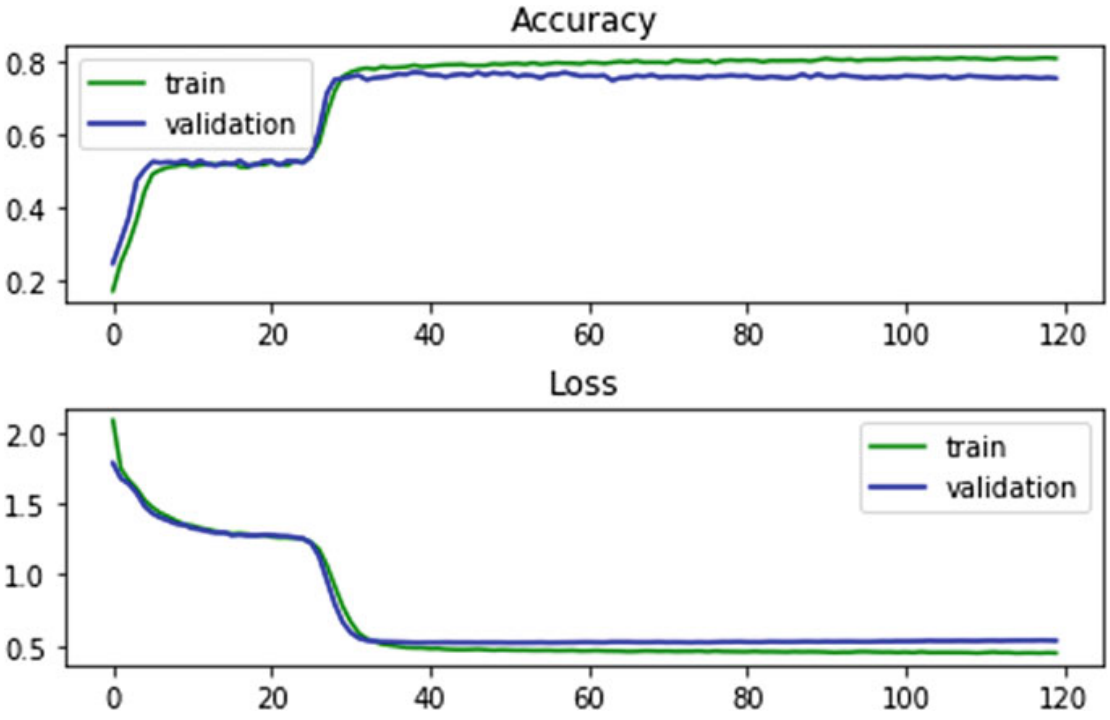


Fig. 5 Accuracy and loss for train and validation set w.r.t. epochs for Task 1

Table 1 Combinations of different parameters for Task 1 which shows changing architecture is having little impact on validation accuracy

Embedding size	Latent size	Batch size	Epochs	Validation accuracy (%)
32	32	64	120	77.98
32	64	64	300	74.83
32	128	64	120	76.2

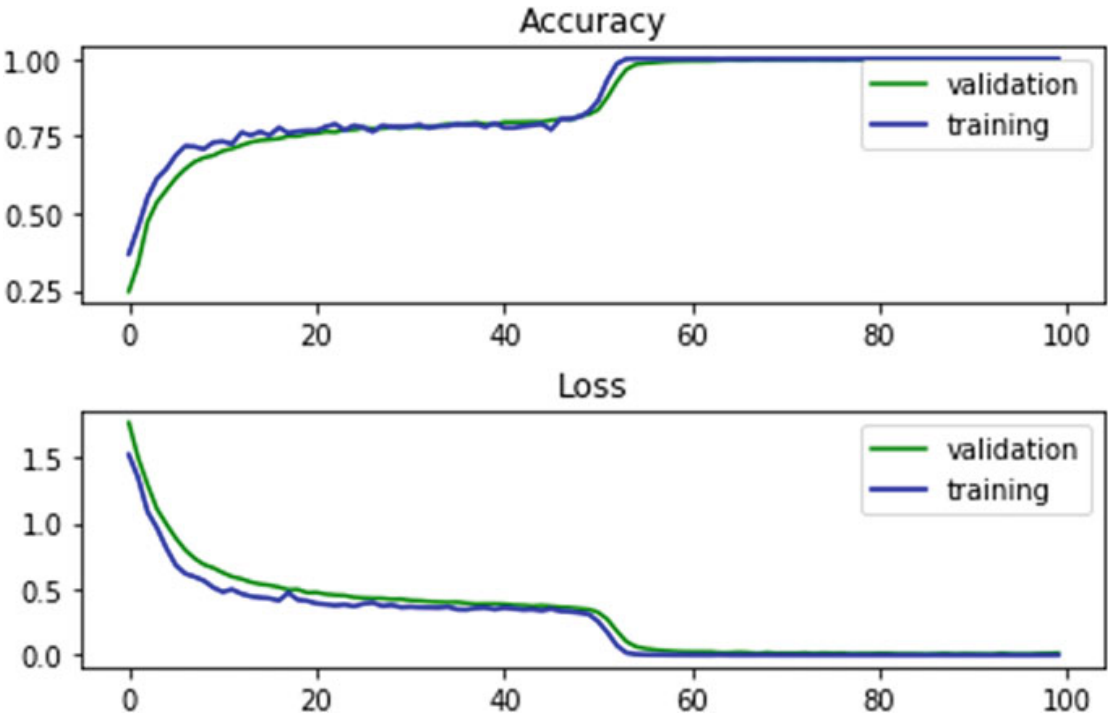


Fig. 6 Accuracy and loss for train and validation set w.r.t. epochs for Task 4

Table 2 Combinations of different parameters for Task 4

Embedding size	Latent size	Batch size	Epochs	Validation accuracy (%)
16	16	32	100	85.24
32	32	32	100	97.01
48	48	48	100	99.64

Story: The bathroom is west of the bedroom. The study room is east of the bedroom.
Q: Bedroom is west of what? A: study room.
Results for the task (Fig. 6, Table 2).

Increasing the embedding size is the key to obtain more accuracy on this task. Embedding is capturing more semantic relationship between words so increasing it is making model more robust. Even simple models give high accuracy on this task.

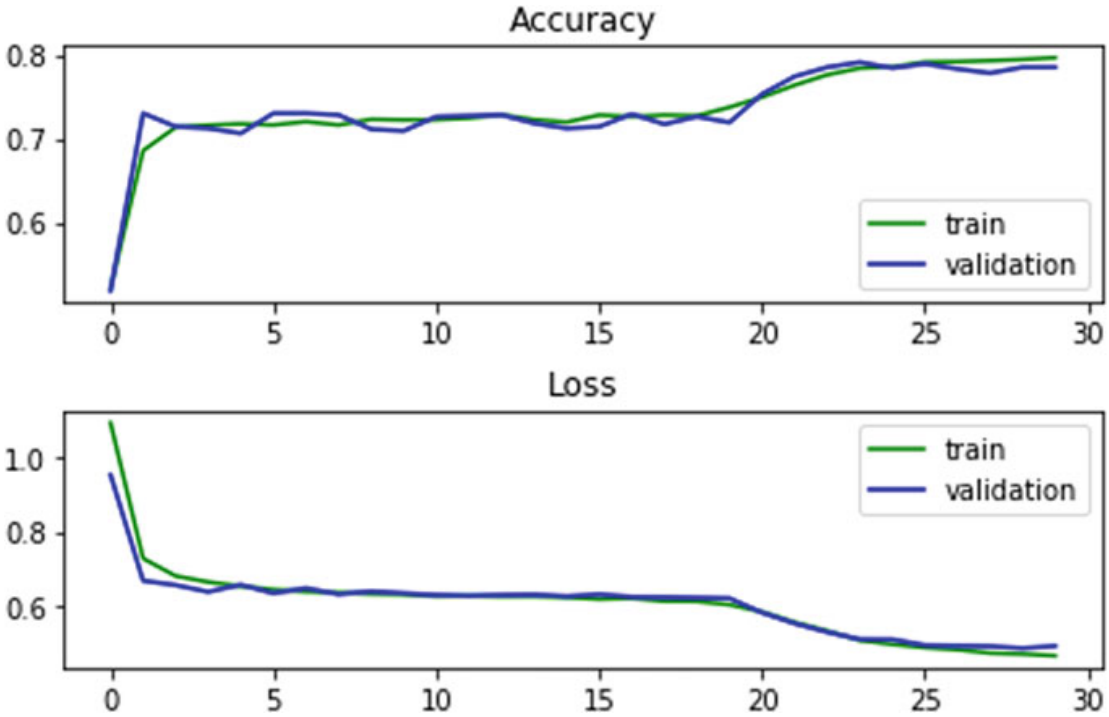


Fig. 7 Accuracy and loss for train and validation set w.r.t. epochs for Task 7

Table 3 Combinations of different parameters for Task 7

Embedding size	Latent size	Batch size	Epochs	Validation accuracy (%)
64	32	64	30	79.35
64	48	64	50	78.00
64	64	96	100	79.26

Task 7: This is a counting task, where counting of objects is to be done based on understanding the statements. For example,
Story: Jonny went to the garden. Daniel went to the bedroom. Sam picked up the apple there. Jonny got the milk there.
Q: Sandra is carrying how many objects? A: one.
Results for the task (Fig. 7, Table 3).

Other tests showed that embedding size was not contributing much to accuracy so the focus was on latent size and batch size. Overall we observe that models with low hyperparameter values are performing as well as those with high values.

Task 17: This is a task where the question relates to the position of an object based on given statements. For example,
Story: The red triangle is to the left of the yellow rectangle. The red triangle is below the pink square.
Q: Is the pink square to the right of the yellow rectangle? A: no
Q: Is the yellow rectangle to the right of the pink square? A: yes
Results for the task (Fig. 8, Table 4).

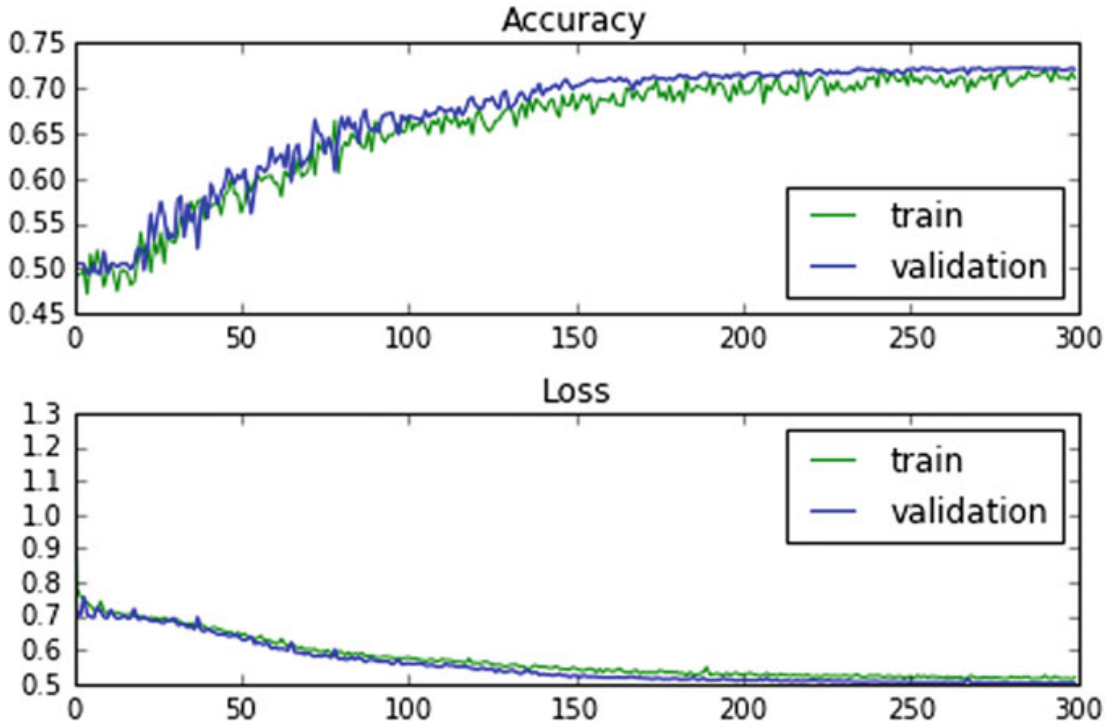


Fig. 8 Accuracy and loss for train and validation set w.r.t. epochs for Task 17

Table 4 Combinations of different parameters for Task 17 which shows increasing epochs and latent size does not improve accuracy much

Embedding size	Latent size	Batch size	Epochs	Validation accuracy (%)
32	64	64	300	70.70
32	128	64	300	72.10
32	300	64	300	72.30

Using different combinations of hyperparameters did not improve the accuracy of this task much with maximum accuracy of 72.3%. This is because in this task it becomes very hard for the memory network to exactly pinpoint the exact lines from which the solution gets extracted as there are several objects present in several positions.

5 Error Analysis

In this section, we analyze output on the basis of different criteria. This would help us in limitation and scope planning of output. The shortcomings which we have identified post-process will also help us to improve training models and algorithms.

Table 5 Accuracy rates by distance of relevant sentence

Distance	Accuracy (%)
1	100
2	100
4	65
5	55
7	20

Currently, this has only been done on the single supporting fact case. The following are a list of few experiments.

5.1 Position of Relevant Sentence from Answer

Our assumption is, if relevant sentence far from the question, then the algorithm has to disambiguate between lots of sentences (Table 5).

For example, in the case below, we see the question present at distance of 1 and 5, respectively, and therefore see a difference in accuracy.

Sandra went to the office
Daniel went to the hallway
Where is Daniel? *Prediction:* ***Hallway*** *Answer:* ***Hallway***
Daniel went to the office
John moved to the hallway
Where is Sandra? *Prediction:* ***Hallway*** *Answer:* ***Office***

We see that by the time the distance becomes seven sentences the accuracy reduces significantly since the question answering system has to look at much longer distances.

5.2 POS (Part of Speech) Tagging Cleanup

POS tagging will help in identification of content word which contains noun, verb, adverb, and adjectives. Also, it would help in ignoring the functional word and stop word like preposition and determiner (a, an, the, etc.). It would help in data cleanup process. The following sentences show input sentence and clean sentence. These extra words may increase the noise and reduce the discriminative power of the model.

Example: Daniel journeyed to the hallway gets converted to Daniel journeyed Hallway after excluding prepositions and determiners.

With POS tagging for 10 separate bootstrap tests, on an average the accuracy was **77.8%** with a while with ignoring unnecessary POS tagged keywords, we got average **82.8%** accuracy on test.

5.3 Number of Nearest Relevant Statement/Answer

In Babl dataset, the question is followed by two answers or statement. There are three different conditions.

(a) Relevant sentence is not present in the previous two sentences.

John moved to the garden
Sandra moved to the bedroom
Where is **Mary**?

Here Mary word is not present in previous two relevant statements before the question "Where is Mary?".

(b) One relevant sentence is present in two different sentences.

Mary went to the bedroom
Daniel went to the garden
Where is **Mary**?

Here Mary is present only once in previous two statements before the question "Where is Mary?"

(c) Two relevant sentences are present in the previous two sentences.

John moved to the bedroom
John journeyed to the hallway
Where is John?

Here John word is present in both previous two statements before the question, "Where is John?".

Here, in Case a, algorithm has to look at relevant statement which is at longer distance from the question, so accuracy would be low.

In Case b, the algorithm has just identified answer from one relevant statement so this would give more accurate results.

Table 6 Accuracy rates for the three possible cases

Number of nearest relevant statement/answer	Accuracy
0	0.64
1	0.8
2	0.6

In Case c, the algorithm has to disambiguate probable answer from two relevant statements. So accuracy is expected to be lesser (Table 6).

6 Conclusion

In this paper, we have utilized recurrent neural network for typical question answering system. Generally, the question answering system explores feature engineering tasks like parsing, pos tagging, and tasking. We have explored how features like token sequences are important for question answering system. We have changed deep learning parameters like epoch, embedding size, latent size, and batch size and identified batch combination to generate the highest accuracy. We also have performed error analysis, where we identified how salient features like (1) presence of stop words, (2) number of relevant statement/answers for different questions, and (3) distance between relevant statement and questions can improve the algorithm in identification of current answers for respective question.

It has been shown that each task will have its own specific architecture with different parameters driving accuracy improvements in each context. So a one size fits all approach which cannot be used in the question and answering systems if we use this framework. Three of the four tasks described have good accuracy (75% plus) which shows we can create very accurate systems. Further steps would involve trying a radically different architecture on the tasks or performing NLP modifications on the inputs where accuracy obtained is low.

References

1. Wang, Z., Yan, S., Wang, H., Huang, X.: An overview of Microsoft deep QA system on Stanford WebQuestions benchmark. Technical report, Microsoft Research (2014)
2. Weston, J., Bordes, A., Chopra, S., Rush, A.M. van Merrienboer, B., Joulin, A., Mikolov, T.: Towards Ai-Complete Question Answering: A Set of Prerequisite Toy Tasks (2014). https://arxiv.org/pdf/1502.05698.pdf
3. Bordes, A., Usunier, N., Chopra, S., Weston, J.: Large-Scale Simple Question Answering with Memory Networks. https://arxiv.org/abs/1506.02075,2015
4. Sukhbaatar, S., Szlam, A., Weston, J., Fergus, R.: End-To-End Memory Networks (2015). https://arxiv.org/pdf/1503.08895.pdf
5. Sutskever, I., Vinyals, O., Le, Q.V.: Sequence to Sequence Learning with Neural Networks (2015). https://arxiv.org/pdf/1409.3215.pdf
6. Murdock, J.W.: Decision Making in IBM Watson Question Answering. Ontology Summit (2015). Web presentations https://researcher.watson.ibm.com/researcher/view_person_pubs.php?person=us-murdockj&t=1

7. Su, V.: Solving the Prerequisites: Improving Question Answering on the bAbI Dataset (2015). http://cs229.stanford.edu/proj2015/333_report.pdf
8. Kumar, A., Ondruska, P., Iyyer, M., Bradbury, J., Gulrajani, I., Zhong, V., Paulus, R., Socher, R.: Ask Me Anything: Dynamic Memory Networks for Natural Language Processing (2015). arXiv preprint arXiv:1506.07285
9. LSTMs and Dynamic Memory Networks for Human-Written Simple Question Answering. https://cs224d.stanford.edu/reports/zack.pdf

Convoluted Cosmos: Classifying Galaxy Images Using Deep Learning

Diganta Misra, Sachi Nandan Mohanty, Mohit Agarwal and Suneet K. Gupta

Abstract In this paper, a deep learning-based approach has been developed to classify the images of galaxies into three major categories, namely, elliptical, spiral, and irregular. The classifier successfully classified the images with an accuracy of 97.3958%, which outperformed conventional classifiers like Support Vector Machine and Naive Bayes. The convolutional neural network architecture involves one input convolution layer having 16 filters, followed by 4 hidden layers, 1 penultimate dense layer, and an output Softmax layer. The model was trained on 4614 images for 200 epochs using NVIDIA-DGX-1 Tesla-V100 Supercomputer machine and was subsequently tested on new images to evaluate its robustness and accuracy.

Keywords Convolution neural network (CNN) · Softmax · Dropout · Galaxy type

1 Background and Introduction

The galaxy classification problem was given initially by Edwin P. Hubble in 1926 [1] and later extended by the scientist: de Vaucouleurs [2]. galaxy classification is a major task involved with all domains of analytical astronomy with major space research agencies relying on automatic classification of galaxies into their corresponding correct classes to further understand the properties and possibilities. Galaxy shapes form the basis of classification of galaxies. Three major shapes of galaxies are noticed

D. Misra (✉) · S. N. Mohanty
Kalinga Institute of Industrial Technology, Bhubaneswar, India
e-mail: mishradiganta91@gmail.com

S. N. Mohanty
e-mail: sachinandan09@gmail.com

M. Agarwal · S. K. Gupta
Bennett University, Greater Noida, India
e-mail: ma8573@bennett.edu.in

S. K. Gupta
e-mail: suneet.banda@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_40

by scientists: elliptical, spiral, and irregular. As we know, the orbits of planets are elliptical due to gravitational force [3]. Similarly, galaxies which are a group of stars also form shapes similar to symmetric curves due to mutual gravitation between stars. Stars are in constant motion with respect to each other due to gravitational pull which causes galaxies to rotate [4]. Irregular galaxies are an exception from above circular shapes and do not have any symmetric shape. This is due to a lot of elemental hydrogen or dust in these galaxies. As we also see on earth dust particles do not form any regular shape, thus the reason behind the shape of irregular galaxies.

Generally, there are three major classes of galaxy, which is described as follows.

1.1 Elliptical Galaxies

A galaxy which has the ellipsoidal shape and a smooth nearly featureless image [5]. Moreover, the spherical form in $3D$ appears to us as a circular shape in $2D$. These are classified into $E0$ which are perfectly circular to $E7$, which are most flattened. Such type of galaxy is the brightest at center and brightness diminishes moving away from the center [6] (Fig. 1).

1.2 Spiral Galaxies

These have three major components: bulge, disk, and halo. Bulge is the central portion which consists of old stars [7]. Arm is the linear portion of stars which are in circular motion around bulge and is made of dust and younger stars. Halo is the spherical part around bulge and covers some part of the disk. Arms emerge directly from bulge (ordinary spiral) or from a bar of material around bulge (barred spiral) [8]. They are also further classified into lower case letters: a, b, c., on how tightly the arms are bound to bulge. The category has most tightly bound arms (Fig. 2).

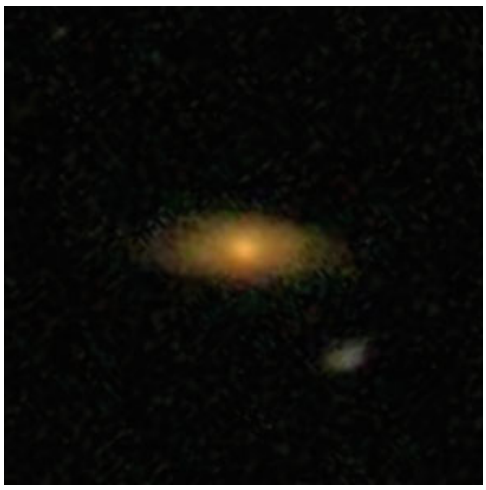


Fig. 1 Sample elliptical galaxy image

Fig. 2 Sample spiral galaxy image

1.3 Irregular Galaxies

These are classified into Irregular I or Irregular II [9]. First category has a lot of elemental hydrogen and young stars. Irregular II has a lot of dust that makes the distinct stars not clearly visible [8] (Fig. 3).

We have made an effort to use deep learning and train the system using some existing labeled images. After this, if the machine is fed with new galaxy images, it will classify the new image into the correct type of galaxy with some degree of accuracy.

In this paper, we approached the problem of galaxy classification based on galaxy images into its three major classes elliptical, irregular, and spiral by applying a Deep Convolution Neural Network (CNN) Architecture. We applied the deep CNN comprising of four hidden layers, one flatten layer, and another dense layer using a Softmax Activation function. The network architecture used max pooling and dropouts wherever required and Tanh activation function was applied in every layer. The model was trained for 200 epochs on the dataset and then was subsequently tested on real-life images in batches of 64 for all the 3 categories. The model was

Fig. 3 Sample irregular galaxy image

highly accurate in its predictions and outperformed existing models and conventional machine learning models in this regard. The rest of the paper is organized in the following chronological order.

In Sect. 2, literature survey of related past work in this domain of galaxy classification is discussed. Followed by Sect. 3 which extensively elaborates about the proposed work and is divided into three corresponding subsections which are introduction to CNN, dataset description, and proposed architecture. Section 4 embeds the experimental setup and results obtained in forms of interactive visualizations and graphs. Concluding, Sect. 5 summarizes the whole paper and provides the conclusion of the research.

2 Related Work

Kormendy and Bender [10] explained about $S0$ galaxies which are intermediate between $E7$ (elliptical) and Sa (true spiral). They have a bulge and a disk, but no spiral so differs from both elliptical and spiral galaxies and are called lenticular galaxies. Authors make a parallel classification of $S0$ galaxies to Sa, Sb, Sc and give them names $S0a$, $S0b$, $S0c$. This classification is done based on B/T ratio, i.e., bulge divided by total light. This value decreases from a to c for both spiral and lenticular galaxies. Buta et al. [11] proposed a new classification of galaxies by the name Comprehensive de Vaucouleurs revised Hubble-Sandage (CVHRS). In this, authors classified galaxies in notations of the form: Sab-A Sab galaxy that is closer to Sa than to Sb, Sab-A Sab galaxy that is closer to Sb than to Sa and so on for other galaxies. Shamir described the automatic classification of galaxy images into elliptical, spiral, or edge-on galaxies [12]. For the experimental purpose, authors used manually classified images to extract image features and discussed the Fisher score. Moreover, test images are classified using nearest weighted neighbor using the Fisher score as an important parameter. The author found automatic classification into elliptical, spiral, or edge with 90% accuracy [12].

In 2013, M. Martin et al. cite9 have proposed galaxy classification algorithm based on Naive Bayes and random forest. The achieved accuracy was about 91% for the Naive Bayes and 79% for random forest classifier. In [13], authors have proposed a classification method based on supervised machine learning with non-negative matrix factorization for images of galaxies in the Zsolt Frei Catalog [14] and achieved the accuracy about 93%. In 2017 [15], authors proposed a new automated machine supervised learning astronomical classification scheme based on the non-negative matrix factorization algorithm with the accuracy of about 92%. Kim and Brunner [16] describe that most star-galaxy classifiers use reduced information from catalogs, this requires careful feature extraction and selection. With the latest advances in machine learning which use deep CNN allows the machine to automatically learn the features directly from the images and minimizes the need for human input. Authors present a star-galaxy classification framework using (ConvNets) directly on galaxy images pixel values.

In contrast to above mentioned approaches, we have used the deep convolution network to improve the accuracy. The details about the proposed model are discussed in the following sections.

3 Proposed Work

3.1 Dataset

In this research project, we classified galaxy Image data into its three corresponding major classes—elliptical type, spiral type, and irregular type using a Deep Convolutional Neural Network (CNN) architecture. The dataset containing the galaxy Images were downloaded from Kaggle [17] and NASA Hubble-Space Gallery Websites [18]. The dataset was categorized into three classes with images kept in two main folders: training and validation. The training folder is further subdivided into three subfolders for three classes: spiral, elliptical, and irregular. The validation folder is also similarly having three subfolders with same names as the classes. The number of images in these folders is listed in Table 1.

Initially, we had only 11 images for the Irregular class. So, we used the Augmentor Package of Python to perform Image Augmentation on those 11 images and generated 1615 augmented images.

Convolution Neural Network (CNN)—A deep learning framework used mostly for object detection and image classification.
Softmax—An activation function used for multiclass categorization.
Dropout—A regularization technique in neural networks to reduce overfitting.
Deep Learning—Deep learning (also known as deep structured learning or hierarchical learning) is part of a broader family of machine learning methods based on learning data representations, as opposed to task-specific algorithms.
Galaxy—A cosmological cluster of stars systems and planets.

3.2 Brief Discussion About Convolutional Neural Network

Convolution Neural Networks (CNN) [19] is the state-of-the-art deep learning model for object recognition, image segmentation, classification, and analysis; but are not only restricted to this scope of problems and fare very well in many other aspects

Table 1 Different classes with the number of images for training, validation, and testing

Classes	Total images	Training set	Validation set	Testing set
Spiral	1464	1000	400	64
Elliptical	1464	1000	400	64
Irregular	1686	1232	390	64

including generating new images as used in Generative Adversarial Networks (GAN) [20]. CNNs are very robust and are very efficient in understanding the features in an image by applying simple convolution formula on the input features using random initialized weights.

The pictorial representation of the architecture of the convolutional neural network is given in Fig. 4.

We built our Convolutional Neural Network (CNN) model in Python using the Keras framework. The CNN architecture comprised of one Input Convolution $2D$ layer followed by four hidden layers, one penultimate dense layer, and finally one output layer. We used a filter size of 3×3 in each layer. All the images were resized

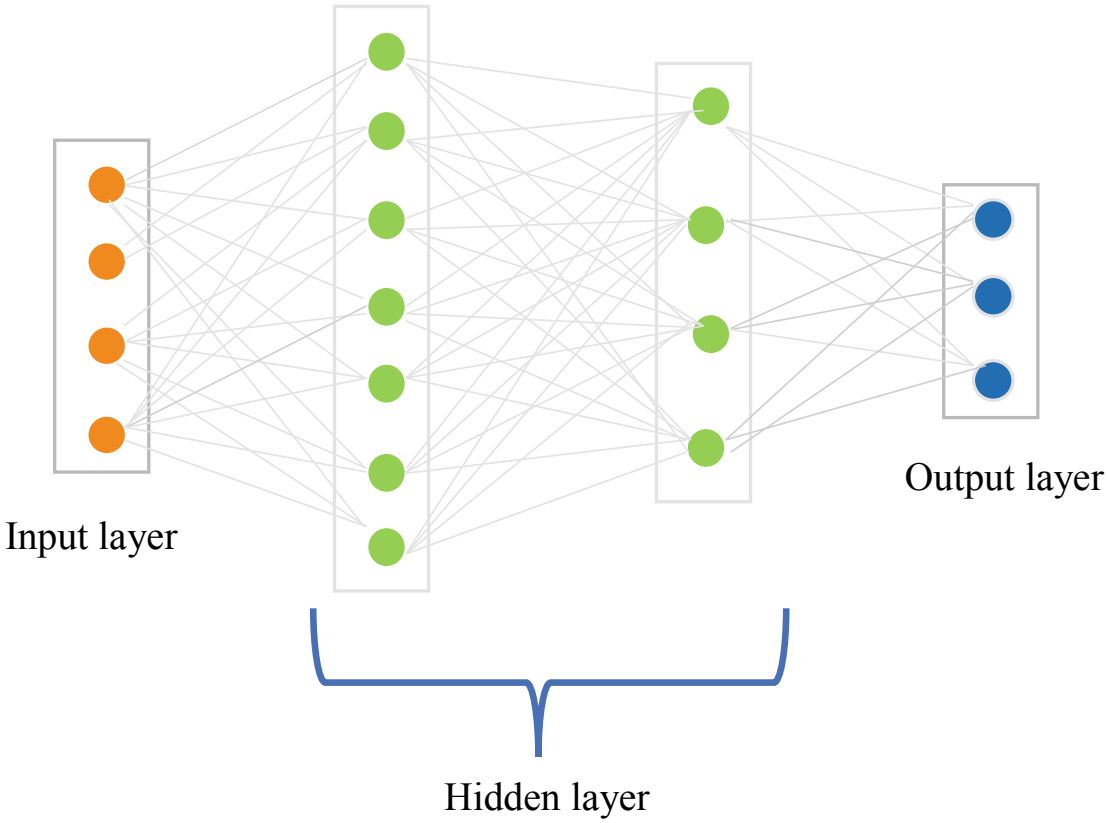


Fig. 4 CNN architecture Layman's visualization

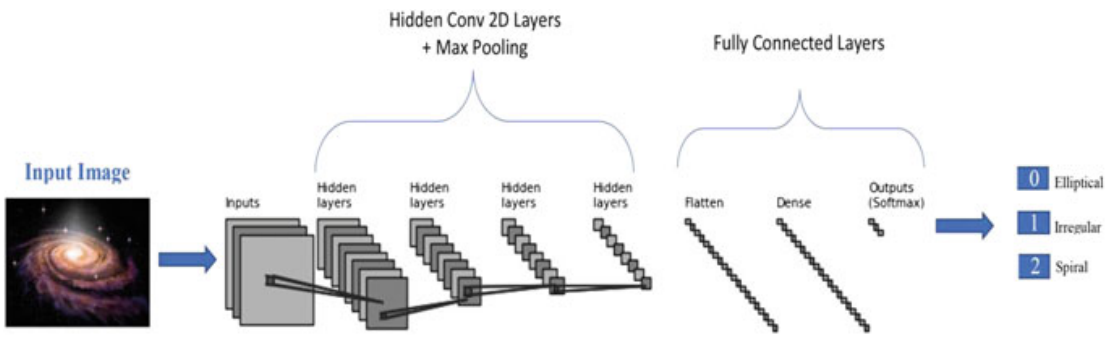


Fig. 5 The architecture of convolution neural network for galaxy classification

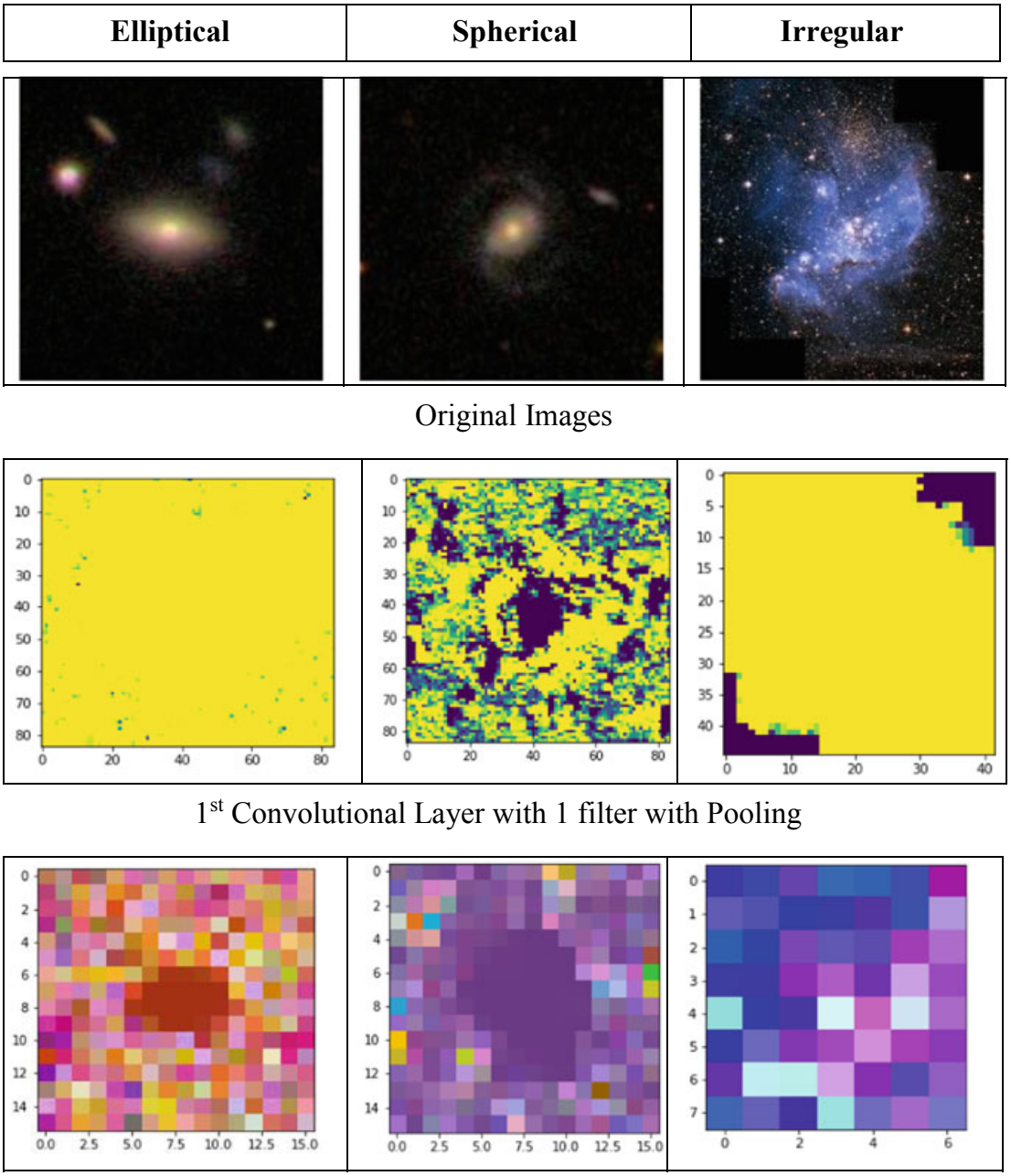


Fig. 6 Intermediate feature map representation with and without pooling

to 256 × 256 pixels. The batch size used was 64 and was trained for 200 epochs with 10 timestamps per epoch. We used dropout regularization after the hidden layers and before the output layer. The detailed architecture diagram is shown in Fig. 5.

After training the CNN model, we saved the weights of the model in weights .h5 file as we could now perform testing easily as many times we needed using the already trained model weights. This step was also performed as training took nearly 1.5 h to complete and it was not feasible to train the model every-time for testing. Thus, this time was saved by making the weights .h5 file for testing. The output images after the execution of intermediate operation of CNN is depicted in Fig. 6.

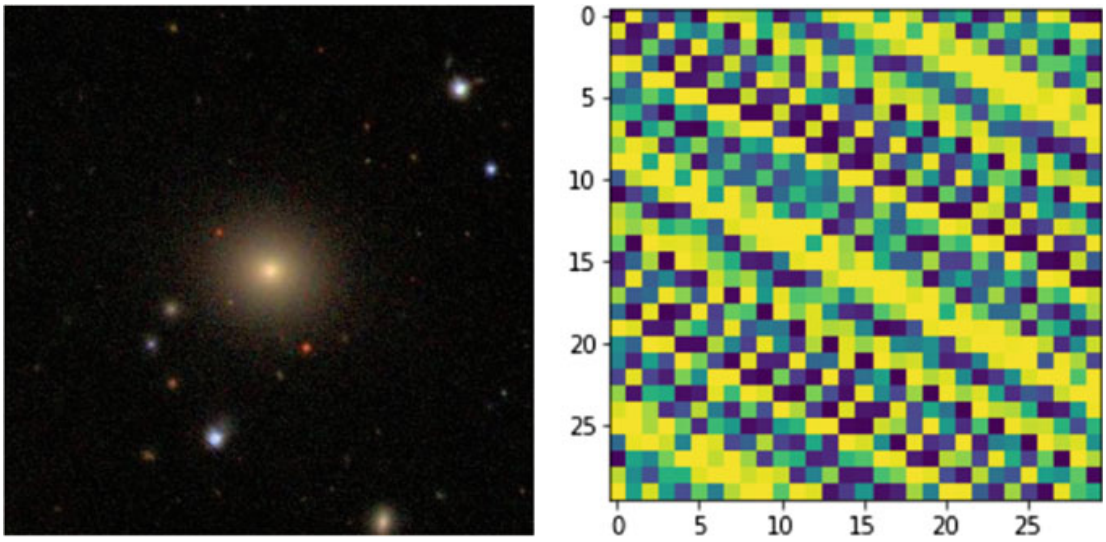


Fig. 7 Feature map representation of hidden layer for an elliptical galaxy image

The hidden layer feature map representation for a sample elliptical galaxy image from the training data is shown in Fig. 7.

4 Experimental Setup

We also trained the model on the Nvidia-DGX-1 (8X Tesla V100) Supercomputer servers having 5120 Nvidia Tensor Cores and a computing power of 960 TeraFlops. The python code was written on Spyder 3.0.0, Jupyter 1.0.0, and Python 3.5.2. using Keras 2.1.3 framework.

After 200 epochs, we obtained a training accuracy of 99.20% and a validation accuracy of 98.45%. After training and saving the model weights, we tested the model and obtained a test accuracy of 97.3958%. The model accuracy and loss curve are depicted in Figs. 8 and 9, respectively.

Upon testing, the results were obtained in the form of a list with each image having labels as 0, 1 and 2 where 0 stands for elliptical type, 1 stands for irregular type and 2 stands for spiral type. The sample output for 64 elliptical images is shown below:

array([0, 0, 0,

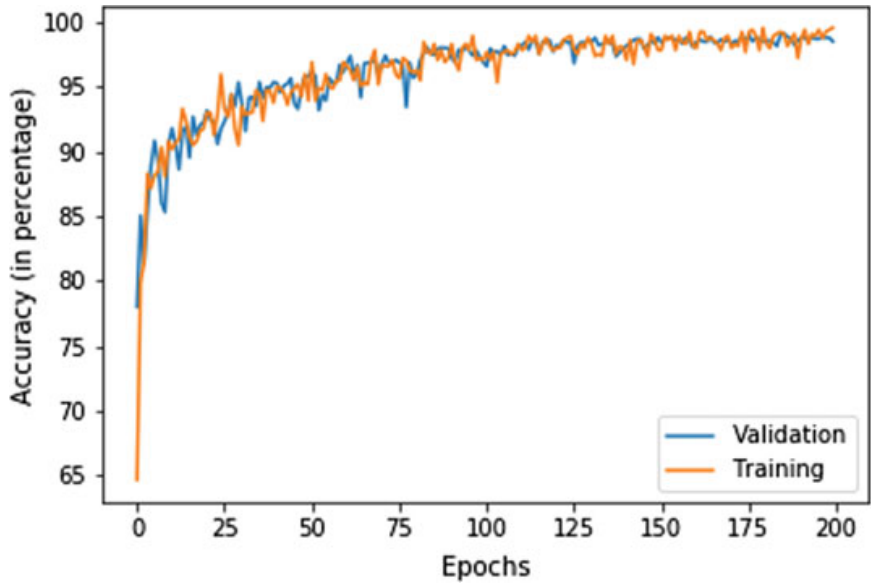


Fig. 8 Graphical representation of accuracy versus epochs

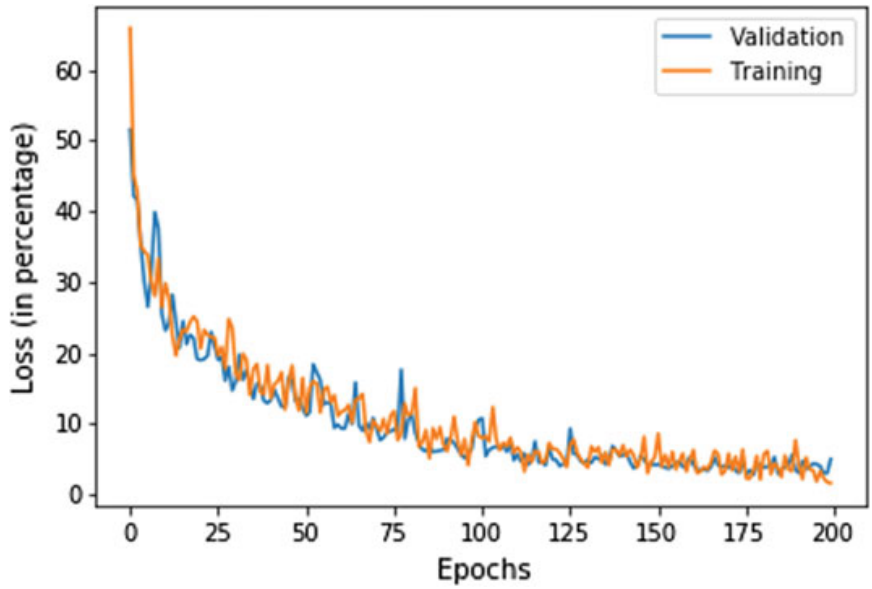


Fig. 9 Model loss curve

Elliptical 100%
Spiral 100%
Irregular 92.187%.

The testing accuracy of images of the various galaxy types is represented in the following Fig. 10 using bar chart representation.

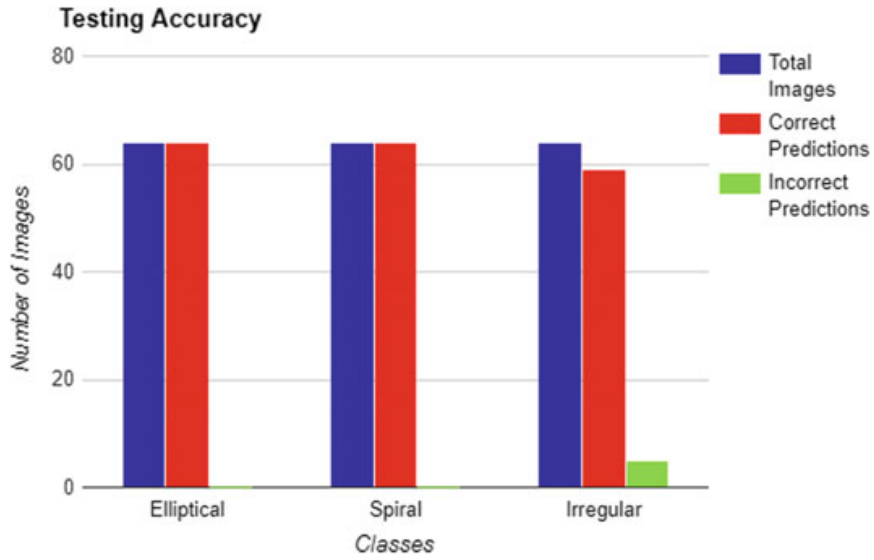


Fig. 10 The representation of testing accuracy using bar chart

5 Conclusion

In this research article, a galaxy classification algorithm has been proposed. The algorithm is based on deep convolution neural network with four hidden layers, max pooling layers after every convolution layer and one dense layer. The proposed research will be helpful for astronomical scientists and cosmologists. It will help to classify huge collection of galaxy images without manual effort of viewing each image individually. Research can be fine-tuned for further classification of galaxies into their subclasses as explained in the previous sections. The testing time was reduced to few seconds by saving the CNN weights file and thus it will be working on real time scenarios also.

References

1. Odewahn, S.C., Stockwell, E.B., Pennington, R.L., Humphreys, R.M., Zumach, W.A.: Automated star/galaxy discrimination with neural networks. In: Digitised Optical Sky Surveys, pp. 215–224. Springer (1992)
2. De Vaucouleurs, G.: Classification and morphology of external galaxies. In: Astrophysik iv: Sternsysteme/Astrophysics iv: Stellar Systems, pp. 275–310. Springer (1959)
3. Hupp, E., Roy, S., Watzke, M.: NASA finds direct proof of dark matter. Press Release (2006)
4. Basu, A., Mandal, A., Das, A.: Blueprint of the graphical galaxy of solar system segment: an algorithmic approach using CSS3.0
5. Liller, M.H.: The distribution of intensity in elliptical galaxies of the virgo cluster ii. Astrophys. J. **146**, 28 (1966)
6. Nieto, J.-L., Capaccioli, M., Held, E.V.: More isotropic oblate rotators in elliptical galaxies. Astron. Astrophys. **195**, L1–L4 (1988)
7. Mihalas, D., Binney, J.: Galactic astronomy* wh freeman co. San Francisco (1968)

8. Benjamin, R.A., Churchwell, E., Babler, B.L., Indebetouw, R., Meade, M.R., Whitney, B.A., Watson, C., Wolfire, M.G., Wolff, M.J., Ignace, R., et al.: First glimpse results on the stellar structure of the galaxy. Astrophys. J. Lett. **630**(2), L149 (2005)
9. Grebel, E.K.: The evolutionary history of local group irregular galaxies. In: Origin and Evolution of the Elements, p. 234 (2004)
10. Kormendy, J., Bender, R.: A revised parallel-sequence morphological classification of galaxies: structure and formation of s0 and spheroidal galaxies. Astrophys. J. Supp. Ser. **198**(1), 2 (2011)
11. Buta, R.J., Sheth, K., Athanassoula, E., Bosma, A., Knapen, J.H., Laurikainen, E., Salo, H., Elmegreen, D., Ho, L.C., Zaritsky, D., et al.: A classical morphological analysis of galaxies in the spitzer survey of stellar structure in galaxies (S4G). Astrophys. J. Supp. Ser. **217**(2), 32 (2015)
12. Shamir, L.: Automatic morphological classification of galaxy images. Mon. Not. R. Astron. Soc. **399**(3), 1367–1372 (2009)
13. Selim, I.M., Keshk, A.E., El Shourbugy, B.M.: Galaxy image classification using non-negative matrix factorization. Int. J. Comput. Appl. **137**(5) (2016)
14. Astronomía, M., Cordero Garayar, J.P., Campusano Brown, L.E., Blanc Mendiberri, G., De Propris, R., Muñoz Vidal, R.: The dry merger rate and merger relic fraction in the coma cluster core
15. Selim, I.M., Abd El Aziz, M.: Automated morphological classification of galaxies based on projection gradient nonnegative matrix factorization algorithm. Exp. Astron. **43**(2), 131–144 (2017)
16. Kim, E.J., Brunner, R.J.: Star-galaxy classification using deep convolutional neural networks. Mon. Not. R. Astron. Soc. 2672 (2016)
17. https://www.kaggle.com/c/galaxy-star-separation/data
18. Van Dyk, S.D., Peng, C.Y., Barth, A.J., Filippenko, A.V.: The environments of supernovae in post-refurbishment hubble space telescope images. Astron. J. **118**(5), 2331 (1999)
19. Haykin, S.: Network, neural: a comprehensive foundation. Neural Netw. **2**(2004), 41 (2004)
20. Radford, A., Metz, L., Chintala, S.: Unsupervised representation learning with deep convolutional generative adversarial networks (2015). arXiv:1511.06434
21. De La Calleja, J., Fuentes, O.: Machine learning and image analysis for morphological galaxy classification. Mon. Not. R. Astron. Soc. **349**(1), 87–93 (2004)
22. Angelica Marin, M., Enrique Sucar, L., Gonzalez, J.A., Diaz, R.: A hierarchical model for morphological galaxy classification. In: FLAIRS Conference (2013)
23. Khalifa, N.E.M., Taha, M.H.N., Ella Hassanien, A., Selim, I.M.: Deep galaxy: classification of galaxies based on deep convolutional neural networks (2017). arXiv:1709.02245

Advances in Network Technologies

Energy-Based Improved MPR Selection in OLSR Routing Protocol

Rachna Jain and Indu Kashyap

Abstract Wireless Ad hoc networks are consisting of wireless nodes that communicate over wireless medium without any centralized controller, fixed infrastructure, base station, or access point. The networks should be established in a distributed and decentralized way. Performance of mobile Ad hoc network depends on the routing scheme chosen. Extensive research has been taken place in recent years to suggest many proactive and reactive protocols to make them energy efficient. In this work, table-driven routing protocol, i.e., optimized link-state routing protocol (OLSR) is tried to make more energy efficient which also helps in prolonging the network lifetime. OLSR is a proactive routing protocol in Mobile Ad hoc Networks (MANETS) which is driven by hop-by-hop routing. The conventional OLSR is hybrid multipath routing, in which link-state information is forwarded only by Multi-Point Relays (MPRs) selected among one-hop and two-hop neighbor sets of host. In this work, a novel mechanism is introduced to select MPR among nodes neighbor set to make it more energy efficient by considering willingness of node. Proposed energy-aware MPR selection in MDOLSR is compared with conventional OLSR. Extensive simulations were performed using NS-2 simulator, and simulation results show improved network parameters such as higher throughput, more Packet Delivery Ratio (PDR) and lesser end-to-end delay as simulation time progresses.

Keywords Multi-point relay (MPR) · Optimized link-state routing (OLSR) · Packet delivery Ratio (PDR) · Residual energy · Throughput

R. Jain (✉) · I. Kashyap
Manav Rachna International Institute of Research and Studies, Faridabad, Haryana, India
e-mail: Rachna_19802000@yahoo.com

I. Kashyap
e-mail: Indu.fet@mriu.edu.in

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_41

1 Introduction

Mobile Ad hoc Networks (MANETS) are general-purpose multi-hopping Ad hoc networks [1]. In these networks, mobile nodes can freely roam within the communication range. These nodes have limited battery capacity [2]. Due to the mobility of nodes topology of the network can change at any time. Routing of data, conservation of energy and bandwidth, and security of information have become major challenges in these battery operated nodes. Despite all these challenges, ad hoc networks have become popular especially in military applications, search and rescue operations, and vehicular communications. Based on network topology routing protocols are broadly categorized as proactive, reactive, hybrid as well as cluster based. In proactive routing protocols, each node maintains a routing table to carry forward the data, whereas in reactive routing route is searched on demand. Hybrid routing protocol combines the advantages of both. Figure 1 shows a scenario of mobile Ad hoc networks.

Although many protocols have been suggested for energy-efficient routing in MANETS but selecting a single path results in quick depletion of energy among nodes. Multipath routing ensures better utilization of the network through load balancing. Routing protocols depending on network topology can be divided into proactive, reactive, and hybrid routing protocol. In proactive routing protocol, every node maintains a routing table. Destination Sequence Distance Vector (DSDV) and Optimized Link-State Routing (OLSR) [3] come under this category. In case of reactive protocols, route is searched on demand. Ad hoc Distance Vector Routing (AODV) and Dynamic Source Routing (DSR) come under this category. Hybrid routing combines the advantages of both proactive and reactive routing.

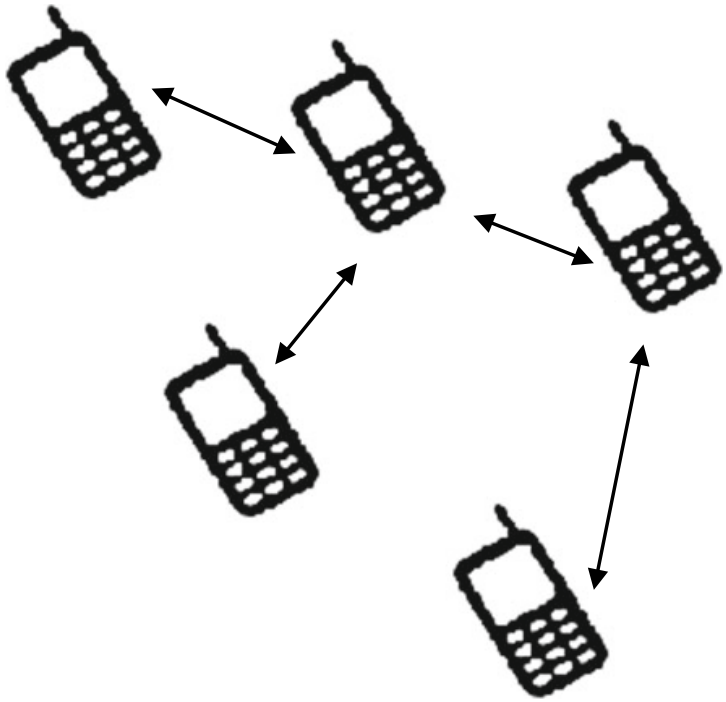


Fig. 1 Mobile Ad hoc networks

Zone Routing protocol (ZRP) comes under this category. This paper is organized as follows: Sect. 2 discusses related work, Sect. 3 discusses OLSR routing protocol in detail, Sect. 4 discusses the concept of energy willingness in detail. Section 5 tells about performance parameters studied and Sect. 6 shows the simulation scenario and results obtained. In last, Sect. 7 discusses the conclusion and future work.

2 Related Work

A detailed literature survey of energy-efficient MANET routing protocols especially focusing on OLSR routing protocol is discussed in this section.

In [4] authors have proposed a modified OLSRM protocol which is based on conventional OLSR. Energy-aware metric used in this protocol is the number of nodes alive against varying simulation time of nodes. Authors have framed an analytical model to compute the correct behavior of the network.

In [5] authors have proposed EEOLSR energy-efficient routing protocol in MANETS which extends network lifetime without losing other Quality of Service (QoS) parameters. Multi-point Relay selection is done on the basis of node willingness to become MPR. This process of MPR selection is discussed in detail in Sect. 3.

In [6] a different approach is used to make energy-efficient OLSR with Autoregressive Integrated Moving Average timeseries model (ARIMA). In this method per interval energy consumption is computed. Composite energy cost model is developed with the help of the calculation of residual energy of the node and consumed transmission power of the node. Sakthivel et al. [7] gave a different metric for MPR selection using energy accuracy metric. Since MPR node is selected from one-hop neighbor set of the host which covers its two-hop neighbors also. Modified MPR is selected from an ordered set based on the value of highest residual energy along with considering timestamp values.

Jabbar et al. [8] gave a new metric, i.e., Multi-Criteria Node Rank metric (MCNR) metric comprising residual battery energy along with node speed to propose a new energy and mobility aware (EMA-OLSR) scheme for selection of MPR in OLSR. Jabbar et al. [9] have proposed Multipath Battery Aware OLSR (MBA-OLSR) which is based on OLSRv2 and its multipath links between source and destination. In this work, the author has considered remaining battery power of nodes to calculate the initial cost of links between sender and destination. Simulations are done using EXata simulator and results show better network lifetime.

Natarajan and Rajendran [10] proposed AOLSR which is hybrid advanced OLSR by modifying Dijakstra's algorithm to compute multiple paths in both dense and sparse networks from source to destination. Whereas Dhanalakshmi [11] proposed Energy Conserving Advanced Optimized Link-State Routing (ECAO) model by calculating energy costs of all nodes and compared results with OLSR and

AOLSR. The author has also implemented modified Dijakstra's algorithm to compute multipath and also checked for link failure for further analysis.

Moussaoui [12] devised stability of nodes (SND) and fidelity of nodes (FND) parameters to elect Multi-Point Relays (MPR) in case of OLSR. The author has proved via simulation that his selected method provides better Quality of Service (QOS) parameters than the traditional method in OLSR for MPR selection which considers Expected Transmission Time metric (ETX). The limitation of using ETX is overestimating link delivery ratio when packet size is too large.

Rango [13] gave a unique routing protocol which considered both link-stability and energy-aware parameters by proposing Link-stability and Energy-aware Routing protocols (LAER). The author has considered link stability along with minimum drain rate energy consumption. Same author Rango [14] has extended Dynamic Source Routing (DSR) protocol packet format to make it energy efficient. Modified DSR header packet included cost function both in route request and route reply packet. Both DSR and OLSR routing protocols have been thoroughly studied from an energetic point of view. The path according to minimum cost function is organized from best to worst path. Many node disjoint paths have been found, out of which most energy-efficient path is chosen. Since node disjoint provides higher fault tolerance. In non-disjoint routes link or node failure causes many routes to fail. Whereas in node or link disjoint routes link failure will cause only a single route to fail.

Another approach to improve routing efficiency is to use multipath in comparison with a single path. More than one path is used to increase routing efficiency. Multiple paths can be linked disjoint or node disjoint. In link disjoint paths, there is no common link, whereas in node disjoint path there is no common node. In link, disjoint paths goal is to reduce delay and increase efficiency. Multipath AODV or Multipath OLSR is proposed to prove it.

3 Optimized Link-State Routing Protocol

Routing algorithm for Ad hoc networks can be classified as routers obtain routing information along with the type of information used to calculate routing path. According to the way routers get information, protocols are classified as proactive, reactive, or hybrid as shown in Fig. 2. According to the type of information proactive protocols are further subdivided as distance-vector-based or link-state-based. Distance-vector protocol uses distance information to obtain path up to destination, whereas link-state-based protocols use topological information to obtain path up to the destination. Link-state algorithms are less inclined to routing loop formation as changes in network topology are known by nodes as compared to distance vector routing protocols. But link-state protocols are more expensive for implementation since they require more CPU power.

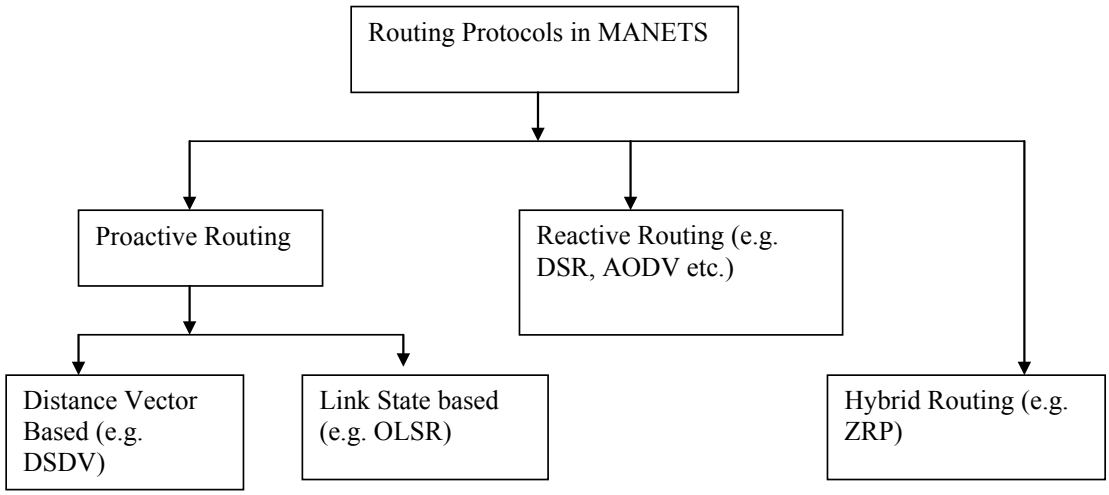


Fig. 2 Classification of routing protocols

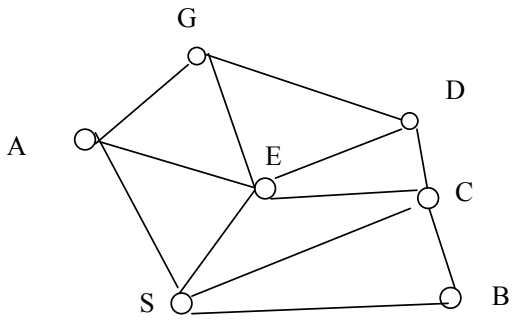
Optimized link-state routing protocol (OLSR) is a proactive, table-driven routing protocol which works on the concept of link sensing. Being a proactive routing protocol, routes are always available. But, if there is any change in network topology due to the mobility of nodes then the information should be flooded through the entire network. To control this flooding, Multi-Point Relays (MPR) are selected from one-hop neighbor set and two-hop neighbor set from the source node. Only selected MPR nodes are now responsible for the flooding of information, hence conserving the bandwidth by avoiding unnecessary flooding. Selecting the minimum number of one-hop neighbors which covers all two-hop neighbors is the goal of MPR selection. There are two types of control messages in OLSR.

(i) HELLO Messages
(ii) Topological Control Messages (TC Messages).

HELLO messages are periodically broadcasted to the neighbors that are only one-hop away. Hello messages obtain information about local links and its neighbors. Since links can be unidirectional or bidirectional, the host should know about all its neighbors. Links are subdivided as symmetric (bidirectional), asymmetric (unidirectional), or heard.

HELLO messages are periodically broadcasted after HELLO refresh time period and received by all one-hop neighbors. Each node attaches a list of its own neighbors to obtain the information about two-hop neighbors as shown by Fig. 3. Once the node has information about its one-hop and two-hop neighbors, it can select Multi-Point Relays (MPRs) which covers all two-hop neighbors as explained by Fig. 4.

Each node periodically broadcasts its HELLO messages to obtain the information about its neighbors and their link status. HELLO message contains a list of



Source Node	1-hop Neighbors	2-hop neighbors	Multipoint Relay
S	A,E,C,B	G,D	E

Fig. 3 Network example for MPR selection

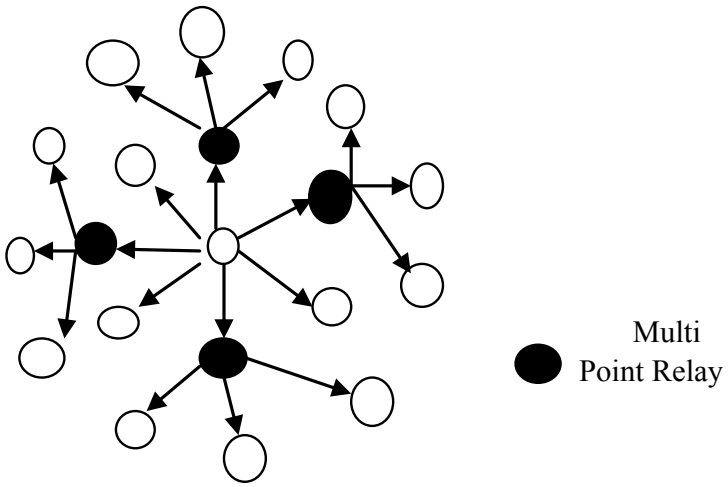


Fig. 4 Diffusion of broadcast message using MPR

addresses of neighbors to which valid bidirectional (symmetric) link exists. It also has a list of addresses of neighbors which have heard HELLO messages but the link is not bidirectional (asymmetric). Format of HELLO message is shown in Fig. 5.

Message Type	Vtime	Message Size	
Originator Address			
Time to Live	Hop Count	Message Sequence Number	
Reserved		Htime	Willingness
Link Code	Reserved	Link Message Size	
Neighbor Interface Address			
Neighbor Interface Address			

Fig. 5 HELLO message format in OLSR

Destination Address	Destination's MPR	MPR selector sequence number	Holding time

Fig. 6 TC message format in OLSR

HELLO message helps in link sensing as link code has the information about link type (Symmetric or Asymmetric) and neighbor type. Each node has knowledge about its two-hop neighbors so it helps in neighbor detection also. On the basis of this information, each node selects its MPR's which is responsible for the forwarding of control packets. On reception of HELLO message, each node constructs its MPR selector table. Multi-point relays are declared in the transmitted HELLO messages. If there is any change in the neighborhood up to two-hop neighbors, then multi-point relay set is recalculated.

Each node maintains a topology table on the basis of Topology Control (TC) messages. Routing tables are calculated on the basis of topology table. Format of TC message is shown in Fig. 6.

If there is entry to the same destination node with higher sequence number then the TC message is ignored. New topology entry is recorded if there is any entry with lower sequence number to the same destination node. Holding time of entry is refreshed in case entry is same as in TC message. New entry is recorded in case there is no corresponding entry.

Each node maintains a routing table to all known destinations in the network on the basis of topology table and neighbor table. Routing table is recalculated after a change in topology table. Destination address, next-hop address, and distance to every node are stored in the routing table.

4 Concept of EA: Willingness in OLSR

4.1 Energy Consumption Model

All the nodes in wireless networks are equipped with IEEE 802.11 g Network Interface Card. Energy required to transfer a packet p from any node is

$$E = i * v * tp \tag{2}$$

where i = Current (in Ampere), v = Voltage (in Volts), tp = Time taken to transmit the packet p (in Seconds)

Transmission time tp is given by

$$tp = (ph/6.10^6 + pd/54.10^6) \tag{3}$$

where ph = Packet Header in bits, pd = Size of payload

Assuming that energy consumption caused by packet overhearing is same as energy consumed in actually receiving a packet.

$$E(p, na) = Etx(p, na) + Erx(p, nb) + (N - 1)\{Eo(p, ni)\} \tag{4}$$

where $Etx(p, na)$ = Amount of energy spent in transmitting packet from node na, $Erx(p, nb)$ = Amount of energy spent in receiving a packet by node nb, $Eo(p, ni)$ = Amount of energy spent by overhearing nodes, N = Average number of neighboring nodes affected by transmission, $E(p, na)$ = Total energy spent in transmitting packet from node na to node nb.

From Eq. 4 it can be concluded that if the network is denser then overhearing causes much energy consumption. Modification is introduced in OLSR to include node willingness to be chosen as Multi-point Relay (MPR). In OLSR routing protocol, each node declares its willingness to be selected as MPR by default. In the proposed work, willingness is decided on the basis of remaining battery capacity of the node along with its lifetime which is dependent on drain rate.

4.2 Passive Hearing by Neighboring Nodes

As simulation time progresses, we can see from simulation results that energy of nodes that are not actively participating in routing is also becoming low. This is due to overhearing by neighboring nodes. Overhearing can be avoided by turning off device/node when unicast message is exchanged among neighboring nodes. It can be obtained by sending signaling messages (RTS/CTS) at MAC layer.

5 Performance Metrics

Performance parameters considered in this work are throughput, end-to-end delay, Packet Delivery Ratio (PDR), and energy required for transmission.

Throughput: It is the measure of rate of successful data transfer over the network. It is measured in bits per second (bps).

End-to-end delay: It is the total time taken in data transfer from source node to destination node. Latency in route discovery, delay in retransmissions at MAC, propagation delay as well as transfer time all contributes to end-to-end delay.

Packet Delivery Ratio (PDR): It is defined as a ratio of packet received by the destination node to the packets sent by source node. Higher the PDR, better is efficiency of routing protocol.

Consumed Energy: It is the sum of total energy spent by all the nodes.

Residual Energy: It is the initial energy minus consumed energy of the node.

Energy Cost per packet: It is the ratio of total consumed energy over number of successfully received packets at the destination.

6 Simulation Results

In this work, MDOLSR is built over UM-OLSR [19]. Simulation parameters are listed in Table 1.

Figure 7 shows that HELLO packet transmission is taking place while the simulation is in progress.

In this work, different performance parameters of modified OLSR, (MDOLSR) such as throughput of the network, packet delivery ratio, energy cost per packet are

Table 1 Network parameters

In this work following initial parameters are taken.	
Channel type	Wireless Channel
Radio propagation model	Two Ray Ground
Mobility Model	Random Way Point
MAC type	IEEE 802.11 a
Interface queue type	Drop Tail
Antenna model	Omni Antenna
Agent	UDP (User Data Protocol)
Application	CBR (Constant Bit Rate)
MANET topology	Dynamic
Speed of nodes	1 m/s…10 m/s
Max packet in ifq	50
Number of mobile nodes	20
Simulation area (*X*)	870
Simulation area (*Y*)	870
Simulation time	30 s….400 s
Pause time of nodes	50 s ….250 s
Initial energy	100 J per node

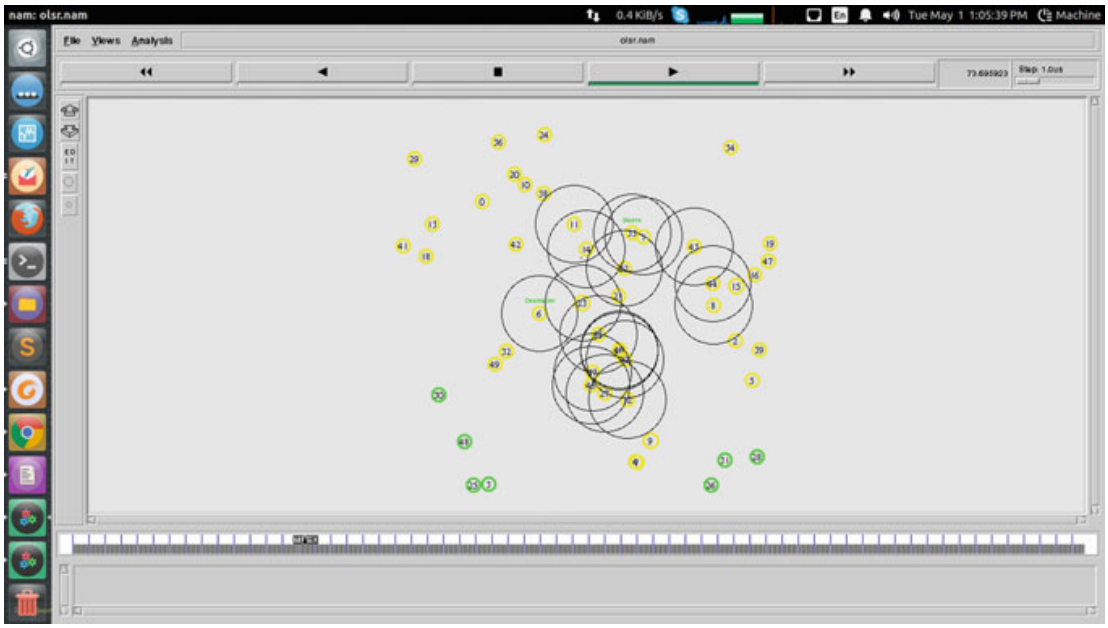


Fig. 7 Network scenario when the nodes have medium energy level

studied according to varying network conditions such as different values of pause times of nodes, varying node speed, and different values of simulation times (Table 2).

Simulation results obtained from modified OLSR are compared with conventional OLSR as shown in Fig. 8 which shows better results for PDR in case of MDOLSR, whereas Fig. 9 justifies the lesser time taken to search the route as the speed of node increases.

Now pause time of nodes is varied from 50 to 250 s and different performance parameters have been studied (Table 3).

The above results obtained are compared with conventional OLSR for same network parameters as depicted in the following graphs. Figure 10 shows there is a significant improvement in PDR of MDOLSR in comparison to traditional OLSR.

Now simulation time of nodes has been varied, and different performance parameters are captured. Table 4 clearly shows that as simulation time progresses residual energies of nodes deplete sharply.

Figure 11 shows the improvement in the throughput of the network of proposed MDOLSR in comparison to OLSR.

Table 2 Node's speed versus QoS parameters

Speed of node (m/s)	Total initial energy (J)	Consumed energy (J)	Residual energy (J)	Energy cost per packet (mWh)	Average energy (J)	Throughput (bytes/s)	PDR (%)	End-to-end delay (s)
1	2000	331.37	1668.62	9.36	16.57	54,464.2	50.38	1.24
4	2000	128.85	1871.15	1.97	6.44	93,571.6	93.08	0.006
6	2000	392.97	1607.03	8.34	19.65	73,470.6	67.05	0.908
7	2000	262.95	1737.04	4.22	13.15	92,629.7	88.56	0.013
8	2000	402.03	1597.97	8.49	20.10	68,018.2	67.36	0.887
9	2000	285.3	1714.7	29.36	14.27	11,263.8	13.83	3.35
10	2000	383.03	1616.96	6.37	19.15	98,112.5	85.61	0.0216

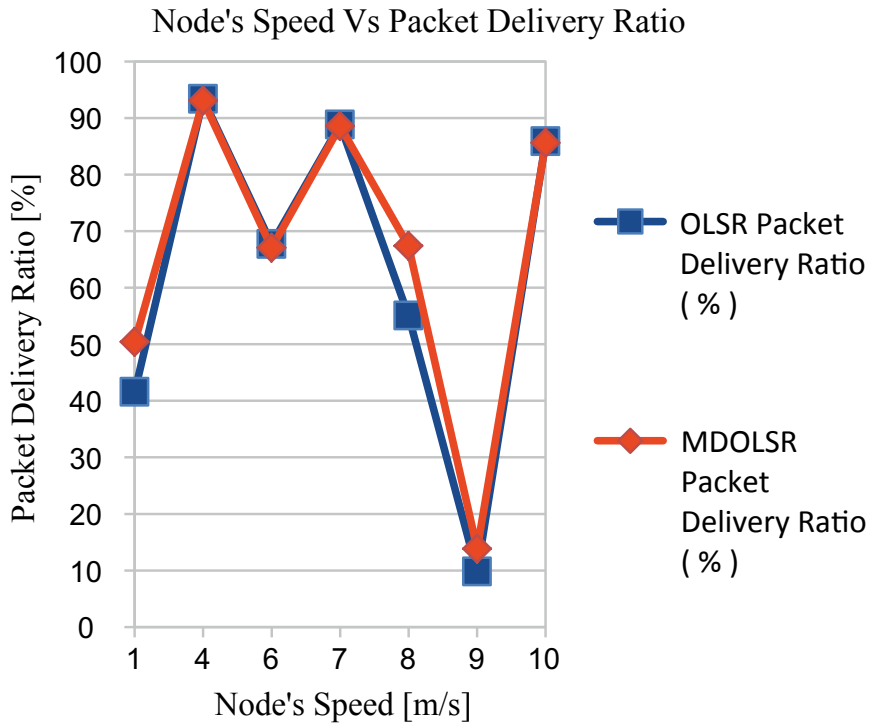


Fig. 8 Packet delivery ratio versus node's speed

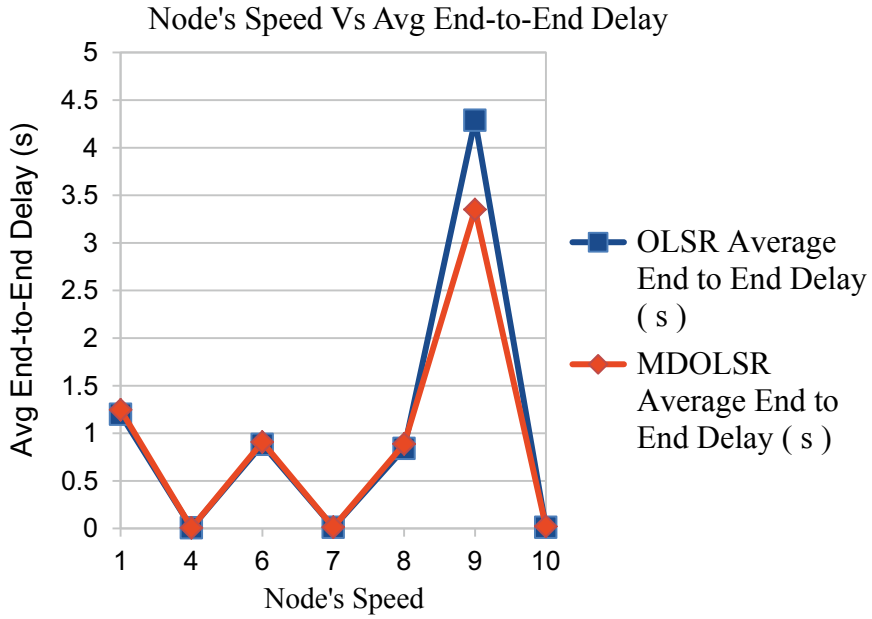


Fig. 9 End-to-end delay versus node's speed

Table 3 Pause time versus QoS parameters

Pause time of node (s)	Total Initial energy (J)	Consumed energy (J)	Residual energy (J)	Energy cost per packet (mWh)	Average energy (J)	Throughput (bytes/s)	PDR (%)	End-to-end delay (s)
50	2000	346.86	1653.13	9.7	17.34	52,355.5	50.89	1.24
100	2000	359.61	1640.38	7.54	17.98	71,637.2	67.81	0.88
150	2000	430.04	1569.96	12.14	21.50	53,490	50.43	1.23
250	2000	328.42	1671.58	5.38	16.42	91,143.7	86.88	0.014

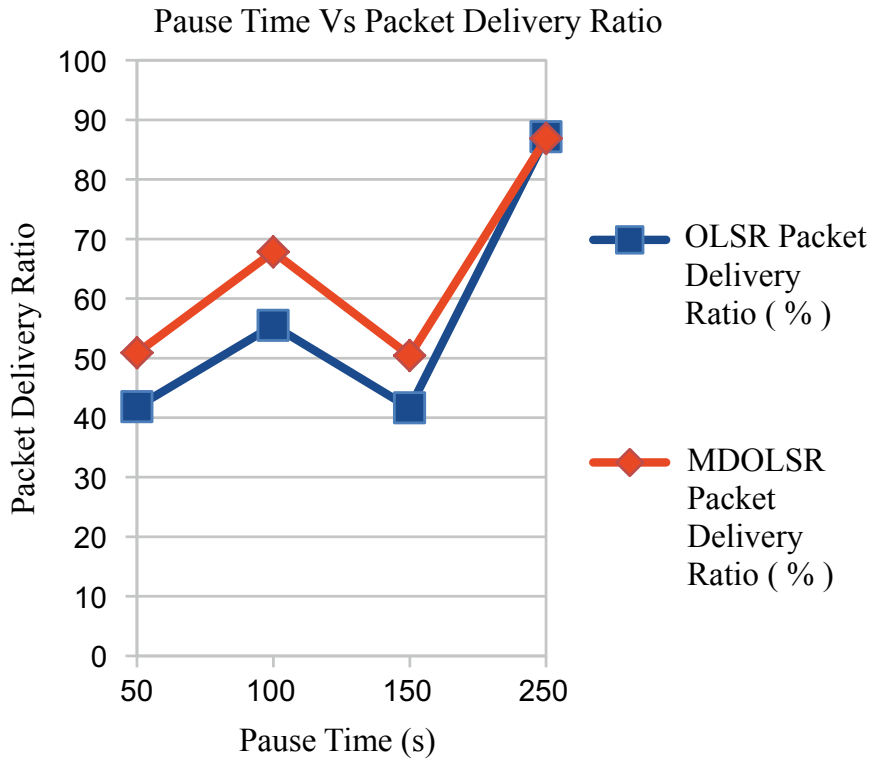


Fig. 10 Packet delivery ratio versus pause time

7 Conclusion and Future Work

Extensive simulations of OLSR routing protocol are chosen over other proactive protocols as it instantly knows about the status of the link and Quality of Service (QoS) information can be provided immediately. Another reason for considering OLSR is that reactive routing protocols result in large overhead as nodes are highly mobile but overhead in OLSR is independent of traffic as there is a fixed upper bound limit for overhead. This routing protocol is suitable for the large and dense network. Proposed methodology in modified MDOLSR results in better QoS parameters such as higher values of throughput and PDR whereas lesser values of energy cost per packet along with a smaller end-to-end delay. In future work energy-efficient OLSR can be further optimized using meta-heuristic techniques such as ant colony optimization, etc.

Table 4 Simulation time versus QoS parameters

Simulation time (s)	Total initial energy (J)	Consumed energy (J)	Residual energy (J)	Energy cost per packet (mWh)	Average energy (J)	Throughput (bytes/s)	PDR (%)	End-to-end delay (s)
30	2000	104.23	1895.77	1.59	5.21	98,170.5	93.08	0.0058
80	2000	689.14	1310.86	3.83	34.45	257,218	95.06	0.0119
130	2000	1609.8	390.19	6.45	80.49	349,408	90.24	0.0162
180	2000	1686.35	313.64	10.94	84.32	273,761	53.03	0.8454
200	2000	1601.37	398.63	11.33	80.06	197,894	55.10	1.6754
300	2000	1591.64	408.35	18.19	79.58	65,943.3	12.26	1.9375
400	2000	1685.6	314.4	6.74	84.28	280,896	73.32	0.0146

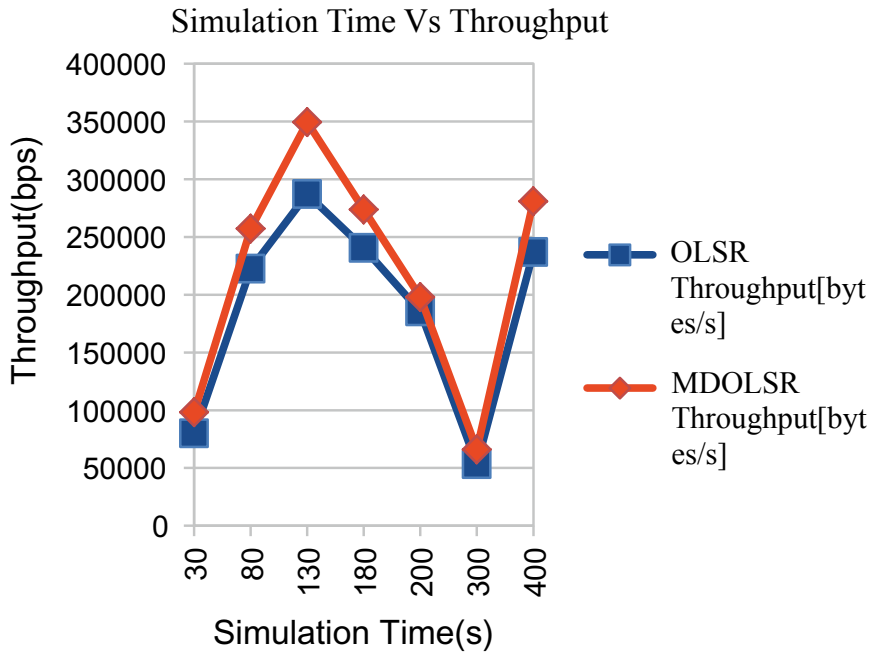


Fig. 11 Simulation time versus throughput

References

1. Chlamtac, I., Conti, M., Liu, J.J.N.: Mobile ad hoc networking: imperatives and challenges. Ad Hoc Netw. **1**(1), 13–64 (2003)
2. Conti, M., & Giordano, S.:. Multihop ad hoc networking: the evolutionary path. In: Mobile Ad Hoc Networking: Cutting Edge Directions, 2nd edn., pp. 1–33 (2013)
3. Clausen, T., Dearlove, C., Jacquet, P., Herberg, U.: The optimized link state routing protocol version 2. In: Internet Engineering Task Force (IETF), No. RFC 7181 (Standards Track) (2014)
4. Joshi, R.D., Rege, P.P.: Implementation and analytical modelling of modified optimised link state routing protocol for network lifetime improvement. IET Commun. **6**(10), 1270–1277 (2012)
5. De Rango, F., Fotino, M., Marano, S.: EE-OLSR: energy efficient OLSR routing protocol for mobile ad-hoc networks. In Military Communications Conference, 2008. MILCOM 2008. IEEE, pp. 1–7. IEEE (2008)
6. Guo, Z., Malakooti, S., Sheikh, S., Al-Najjar, C., Malakooti, B.: Multi-objective OLSR for proactive routing in MANET with delay, energy, and link lifetime predictions. Appl. Math. Model. **35**(3), 1413–1426 (2011)
7. Sakthivel, M., Palanisamy, V.G.: Enhancement of accuracy metrics for energy levels in MANETs. Comput. Electr. Eng. **48**, 100–108 (2015)
8. Jabbar, W.A., Ismail, M., Nordin, R.: Energy and mobility conscious multipath routing scheme for route stability and load balancing in MANETs. Simul. Model. Pract. Theory **77**, 245–271 (2017)
9. Jabbar, W.A., Ismail, M., Nordin, R.: Performance evaluation of MBA-OLSR routing protocol for MANETs. J. Comput. Netw. Commun. **2014**, 10 (2014)
10. Natarajan, D., Rajendran, A.P.: AOLSR: hybrid ad hoc routing protocol based on a modified Dijkstra's algorithm. EURASIP J. Wirel. Commun. Netw. **2014**(1), 90 (2014)

11. Dhanalakshmi, N., Alli, P.: Efficient energy conservation in MANET using energy conserving advanced optimised link state routing model. Int. J. Parallel Emergent Distrib. Syst. **31**(5), 469–480 (2016)
12. Moussaoui, A., Semchedine, F., Boukerram, A.: A link-state QoS routing protocol based on link stability for mobile ad hoc networks. J. Netw. Comput. Appl. **39**, 117–125 (2014)
13. De Rango, F., Guerriero, F., Fazio, P.: Link-stability and energy aware routing protocol in distributed wireless networks. IEEE Trans. Parallel Distrib. Syst. **23**(4), 713–726 (2012)
14. De Rango, F., Cano, J.C., Fotino, M., Calafate, C., Manzoni, P., Marano, S.: OLSR vs DSR: a comparative analysis of proactive and reactive mechanisms from an energetic point of view in wireless ad hoc networks. Comput. Commun. **31**(16), 3843–3854 (2008)

A Novel Approach for Better QoS in Cognitive Radio Ad Hoc Networks Using Cat Optimization

Lolita Singh and Nitul Dutta

Abstract Cognitive Radio is a Wi-Fi verbal exchange methodology that allows the user to engage except having a fixed preassigned radio spectrum. Cognitive Radio Networks (CRNs) are having the routing hassle that is one of the serious constraints. Ad hoc networks are non-centralized Wi-Fi networks that can be constructed and there is no need for any preexisting infrastructure for these networks. Here every point can work as a router. In this paper, the authors have explained the Cognitive Radio Networks (CRN) that are obtaining so a whole lot of recognition where the principal focus is on the dynamic undertaking of channels to wireless devices. In this paper, cognitive radio networks are primarily focused. Nowadays, almost all the networks rely on fixed allocated networks in an approved or unapproved frequency group. In this paper literature evaluates associated to CRN and an optimization algorithm to enhance the overall performance of TE under CRN has been discussed. Swarm intelligence technique is used in the paper. Swarm approach is clearly the combination of the decentralized attribute to gain excellent viable solutions. The motivation regularly creates from nature, more often than non natural outlines. One of the effective approachs known as Cat swarm has been used to acquire high price of accuracy and much low error rates which improves the lifespan of the network. The results are carried out by the use of CSO (Cat Swarm Optimization) algorithm and parameters like energy consumption, congestion, overhead consumption, and number of routing rules are used to analyze the overall performance of the algorithm.

Keywords Cognitive radio networks · CSO · Swarm approach · TE

L. Singh (✉)
Department of Computer Science Engineering, Marwadi University, Rajkot, India
e-mail: Lolita.singh101463@marwadiuniversity.ac.in

N. Dutta
Faculty of Technical Science, Jan Wyzykowski University,
ul. Skalnikow 6 B, 59-101 Polkowice, Poland
e-mail: n.dutta@ujw.pl2

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_42

1 Introduction

Cognitive Radio is developed by popular person Joe Mitola in 1999 [1]. They allow wireless communication to come into a new method. The CR science provides a flexible result for the trouble of spectrum shortage by means of potential of dynamic spectrum allocation for community communication. It permits the coexistence of CR successful devices with licensed band customers [2] and allows the later to use the licensed spectrum opportunistically. The CR units can correctly feel on hand idle spectrum, reconfigure parameters to get admission to the quickly unused spectrum, and produce unbearable interference to licensed users, which make them successfully network among themselves [1].

Routing in traffic engineering using cognitive radio networks is one of the major topics. The routing hassle is one of those recent research topics which shows up in Cognitive Radio Networks (CRNs). The CRN has confirmed itself as a network to improve spectrum strength via self-organization and dynamic reconfiguration and thus, the Cognitive Radio Ad Hoc Network (CRAHN) [3] is introduced as a promising conversation technology. But, CR users need to stop their conversation or lower their transmission power to keep away from disturbances to authorized licensed customers (called Primary Users (PUs)). Because PUs are the owners of the channel, and the CR units (also referred to as Secondary Users (SUs)) opportunistically use the channel when PUs have no statistics to transmit [4]. That is why the unused spectrum (called spectrum hole) needs to be used successfully to make CRN successful. Besides that, if transmissions of CUs do no longer purpose dangerous interferences with PU transmissions, then CUs can communicate among themselves. CU and SU actually mean the same. From overall discussion CU, SU, and CR are the terms used interchangeably for a cognitive node (Fig. 1).

Network investigation is succeeding in all areas nowadays. Rare greatest capable expertise like 4G [1], facts-centric communication, and software-defined interaction [4] are receiving substantial importance. However additional idea named green

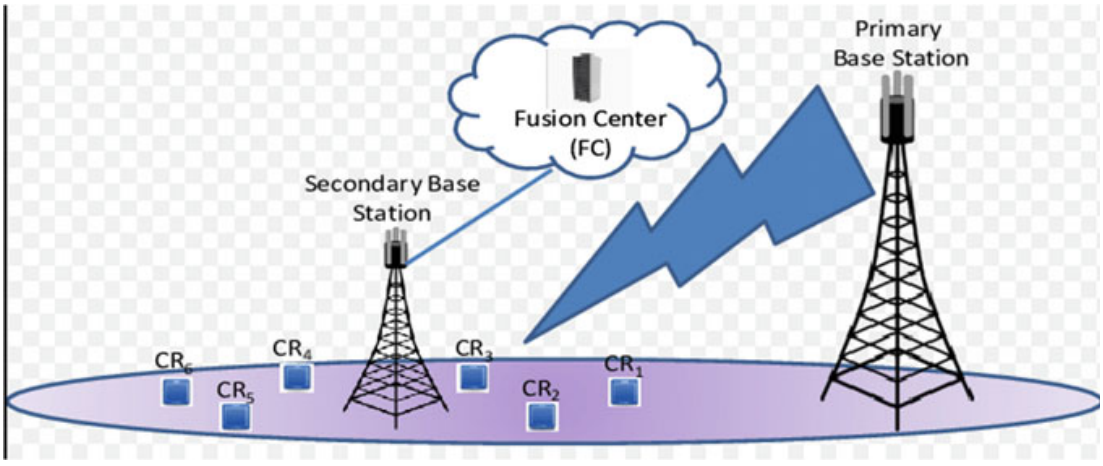


Fig. 1 CRN (cognitive radio network)

mobile computing is also attaining attention. In this, the procedures are intended to reduce power intake thus carbon emanation is minimized throughout the calculation. A review of such green computing constructed routing is established. Biology impressed algorithms are a unit extra step to green computing. Time serving packet forwarding and routing in Delay Tolerant Network (DTN) [9] is alternative space wherever several analysis work is established. These embrace transport network primarily based DTN and transport detector network. The routing in MANET with QoS improvement acquires an equivalent position in several analyses. Though, of these systems declared on top of depends on statically allotted channels either in authorized or unaccredited band. Such fixed channel job construct ends up in spectrum congestion and thus spectrum insufficiency in today's thickly inhabited situation. That's why, a replacement technology psychological feature i.e. Cognitive Radio Networks (CRN) attaining quality that works on vibrant task of channels to wireless devices.

2 Literature Review on CRN

Author	Description
Cacciapuoti et al. (2010)	Presented a routine procedure known as CRN On-demand Distance Vector or CAODV. This suggested technique designed for MANET is an adjunct of Ad Hoc On-demand Distance Vector routing protocol. Like AODV, this procedure also makes a path once there is a route for a package to send. Numerous manage packages of AODV, e.g., path appeal; neighbor discovery packets are moreover castoff in CAODV
Karim Habak et al. (2013)	Proposed a technique of routing known as LAUNCH. This procedure has certain features like effective usage of the frequent manage network, negligible path arrangement postponement. However, this procedure doesn't warranty the steadiness of the mounted path for the duration of conversation [6]
Sun et al. (2014)	Introduced a CRN-based totally vehicular network routing protocol. As mentioned in the paper, the protocol is appropriate for conditions like herbal disasters which cause extreme damage to communication infrastructure. Additionally, such herbal disasters are accompanied through sharp spikes in the utilization of commercially licensed spectrum. In such case, if pretentious victims try to transfer facts with excessive bandwidth requirement, wireless network, then the proposed protocol can suitably be used in CRN [11]
Ian F. Akyildiz et al. (2009)	Introduced CRAHNs current challenges and intrinsic features. The Software-Defined Networking (SDN) has been introduced and proposed. CRN deployments interference is managed by using the advantages of cloud computing and SDN in case of residual areas [5]
Anatolij Zubow et al. (2015)	A new CR SDN based structure has been introduced in which Spectrum Broker (SB), centralized controller primarily related cloud and it takes over spectrum task to CR Base Stations (CR-BS). The Wi-Fi information are combined to the SB beneath CR-BSs

(continued)

(continued)

Author	Description
	manipulate report, in the CRN community site visitors situation up to date records is done by way of SB controller with the aid of configuring acceptable policies in Open Flow-enabled CR-BSs [8]

3 CSO (Cat Swarm Optimization)

The Swarm Intelligence (SI) algorithms are optimization algorithms that were formulated for copying the intelligent behavior and nature of animals. In these technologies, a population of living things such as ants, bees, birds, and fish is communicating with each other and with their place via sharing information that results in the use of their place and things. Cat Swarm Optimization (CSO) algorithm is one of the more recent SI-based algorithms in which the nature of cats is simulated. The CSO algorithm is developed by Chu and Tsai (2007), and its things have been implemented for different optimization problems.

3.1 Natural Process of the Cat Swarm Optimization Algorithm

High alertness and curiosity about the environment is the nature of cats. They cannot spend most of their time resting; moving objects in their surroundings are their normal things. By this behavior, cats can cheat and seek their food. They spend very less time chasing prey to consume their energy as compared to the time dedicated to their resting. Chu and Tsai (2007) really got inspired by this behavior of cats and introduced CSO which is called as Cat Swarm Optimization. CSO has two modes: "seeking mode" it is the mode when cats are having rest and "tracing mode" it is the mode when cats are seeking their prey or food. CSO algorithm does the same thing. In CSO, a population of cats is first formulated and assigned in the M-dimensional solution space. Here each cat is representing a result. Then further two subgroups are made by the population. In the first group, the cats are rest mode and are very alert about their surroundings (i.e., seeking mode), while in the second group, the cats are trying to chase their preys (i.e., tracing mode). The hybridization of these two groups helps CSO to get the global solution in the M-dimensional solution space. The tracing subgroup should be small because the cats do not spend much time in the tracing model. The Mixture Ratio (MR) is defined for this which has a small value. New positions and fitness functions can be obtained after sorting the cats into these two modes. After this, the cat with an optimistic solution is

selected and saved in the memory. These steps have to be repeated until the stopping criteria are satisfied (Fig. 2).

The steps of CSO are given below:

Step 1: Initial population of cats has to be created and make them into the M-dimensional solution space (Xi,d) and provide each cat a velocity in the range of the maximum velocity value (ti,d). Then further two subgroups are made by the population. In the first group, the cats are in rest mode and are very alert about their surroundings (i.e., seeking mode), while in the second group, the cats are trying to chase their preys (i.e., tracing mode).

Step 2: Give every cat a flag to arrange them into the tracing mode process or seeking mode process. This depends on the value of MR. The tracing subgroup should be small because the cats do not spend much time in the tracing model. The Mixture Ratio (MR) is defined for this which has a small value.

Step 3: Calculate the fitness value of every cat and the cat with the best fitness function has to be selected. The position of the best cat tells the best solution so far. Xbest is the position of the best cat.

Step 4: The cats are applied to the seeking or tracing mode process and is based on their flags. New positions and fitness functions can be obtained after sorting the cats into these two modes. After this, the cat with an optimistic solution is selected and saved in the memory.

Step 5: If the termination process is accepted, end the process. Otherwise, repeat steps 2–5.

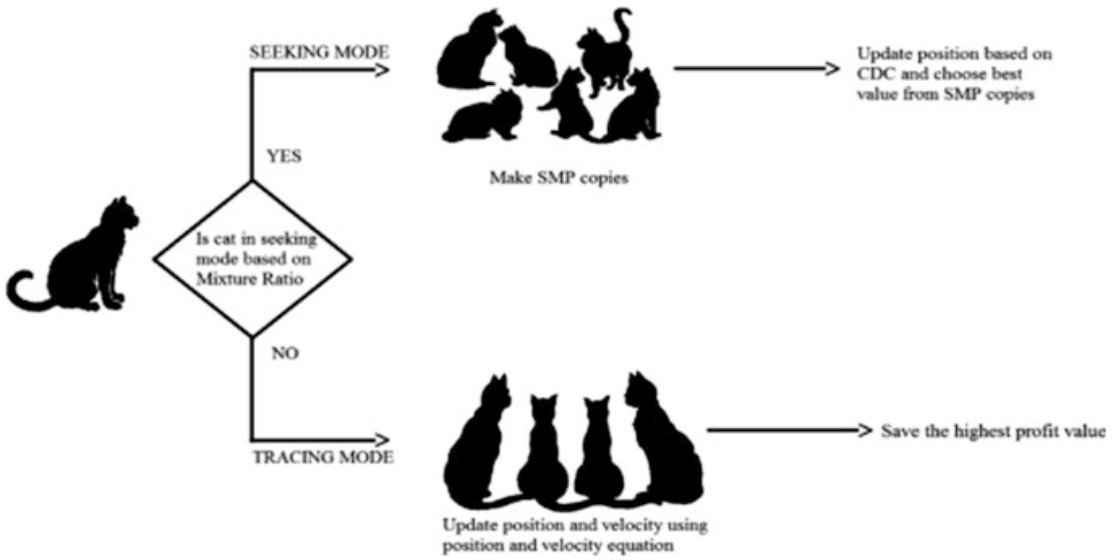


Fig. 2 Cat swarm optimization

3.2 CSO Algorithm

Algorithm for Cat swarm Optimization is given below:

Begin
Input parameters of the algorithm and the initial data
Initialize the cat population Xi (i = 1, 2. n), υ and w
While (the stop criterion is not satisfied or I < I max)
Calculate the fitness function values for all cats and sort them
Xg = cat with the best solution
For i = l: N
If w = l
Start seeking mode
Else
Start tracing mode
End if
End for i
End while
Post-processing the results and visualization
End

CSO algorithm normally uses the behavior of the cats and main focus is on two modes, i.e., "seeking mode" it is the mode when cats are having rest and "tracing mode" it is the mode when cats are seeking their prey or food.

4 Problem Formulations

Cognitive radio networks in routing in traffic engineering are one of the serious topics. The routing problem is one of those recent research topics which occur in Cognitive Radio Networks (CRNs). There is no need for preexisting infrastructure for Ad hoc networks (CRN). These are non-centralized wireless networks that can be easily formulated. An optimization algorithm is used to solve the problems in routing in traffic engineering using cognitive radio networks. Optimization algorithms are those that help in obtaining the best optimistic results. The optimization can be done using (CSO) Cat Swarm Optimization to optimize link utilization for the routing. Cat swarm is one of the effective approaches to get such optimize solutions to achieve a high rate of accuracy and fewer error rates which increase the lifespan of the network. The cat with an optimistic solution is selected and saved in the memory. The optimistic result would be the greatest and best position which is nothing but one of the cats to keep the best solution till it spreads the end of repetitions.

Objectives:

- To implement optimization algorithm called as CSO (Cat Swarm Optimization)
- To analyze the performance of the CSO for less overhead, congestion, and energy consumption to achieve a high quality of service.

The parameters are

- Energy consumption
- Congestion
- Overhead
- Variation of the routing rule.

Swarm intelligence approach has been used because it is a metaheuristic approach which is globally applied to any file to achieve high-end results according to the designed objective to achieve the best solution from the number of solutions.

5 Results and Discussions

The swarm intelligence approach is simply the combination of the decentralized characteristic to obtain the best optimistic results. The idea is utilized in an efficient method at computerized cognitive.

Swarm intelligence frameworks include regularly of a population of direct particles known as swarms interfacing with each other for the solution. The motivation frequently formulates from nature, mainly organic outlines. Cat Swarm optimization is used in this paper. The main aim of selecting this algorithm in the paper is to get optimistic solutions. In Cat Swarm Optimization, high alertness and curiosity about the environment is the nature of cats. They cannot spend most of their time resting; moving objects in their surroundings are their normal things. By this behavior, cats can cheat and seek their food. They spend very less time chasing prey to consume their energy as compared to the time dedicated to their resting. Chu and Tsai (2007) really got inspired by this behavior of cats and introduced CSO which is called as Cat Swarm Optimization. CSO has two modes: "seeking mode" it is the mode when cats are having rest and "tracing mode" it is the mode when cats are seeking their prey or food. CSO algorithm does the same thing. In CSO, a population of cats is first formulated and assigned in the M-dimensional solution space. Here each cat is representing a result. Then further two subgroups are made by the population. In the first group, the cats are rest mode and are very alert about their surroundings (i.e., seeking mode), while in the second group, the cats are trying to chase their preys (i.e., tracing mode). The hybridization of these two groups helps CSO to get the global solution in the M-dimensional solution space. The tracing subgroup should be small because the cats do not spend much time in the tracing model. The Mixture Ratio (MR) is defined for this which has a small value. New positions and fitness functions can be obtained after sorting the cats into these two modes. After this, the

cat with an optimistic solution is selected and saved in the memory. These steps have to be repeated until the stopping criteria are satisfied.

6 Result Explanation

The results are shown in the form of graphs and table.

Figure 3 shows the initialization of the network and shows that the cognitive radios are deployed and primary users are shown for the traffic communication and for the services and the common receiver is deployed to receive the packets and act as the destination.

Figure 4 shows the overhead consumption in which the overhead consumption is decreasing rapidly and is the desired output which is done using Cat swarm optimization approach and is able to achieve fewer collisions and intelligent transmissions (Fig. 5).

As we can see from the above figure that the energy consumption is decreasing in an efficient manner which is one of the main constraints in traffic engineering using the processing elements. The energy consumption is decreasing as the mobility increases which is also one of the desired outputs of the cognitive radios.

Figure 6 shows the congestion in the network which is one of the important parameters in the cognitive-based traffic engineering and shows that the congestion also increases as the mobility increases which increases the lifespan of the cognitive network (Fig. 7).

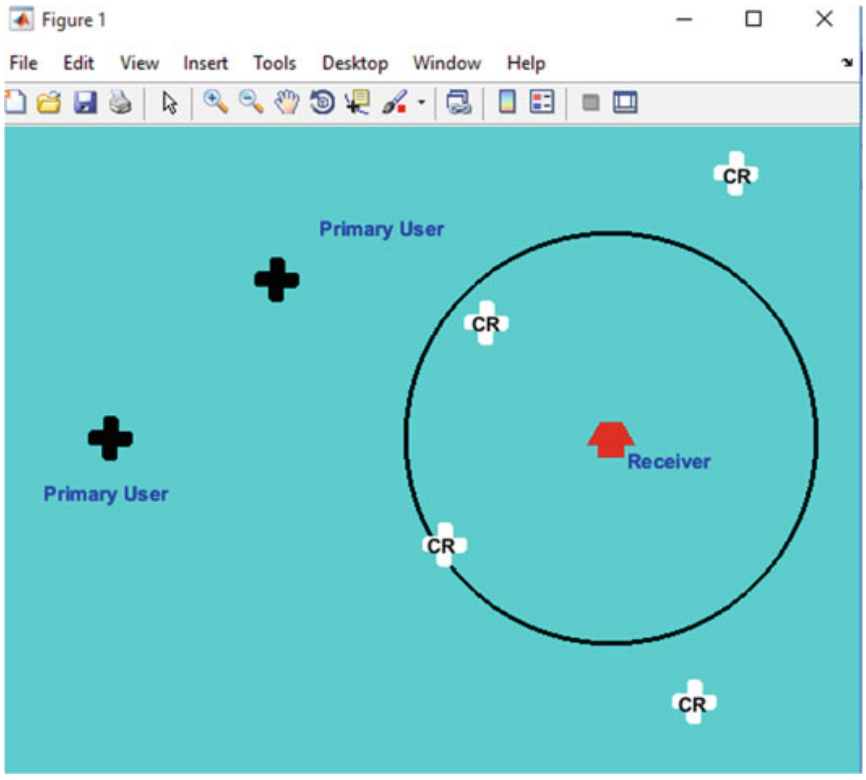


Fig. 3 Network initialization

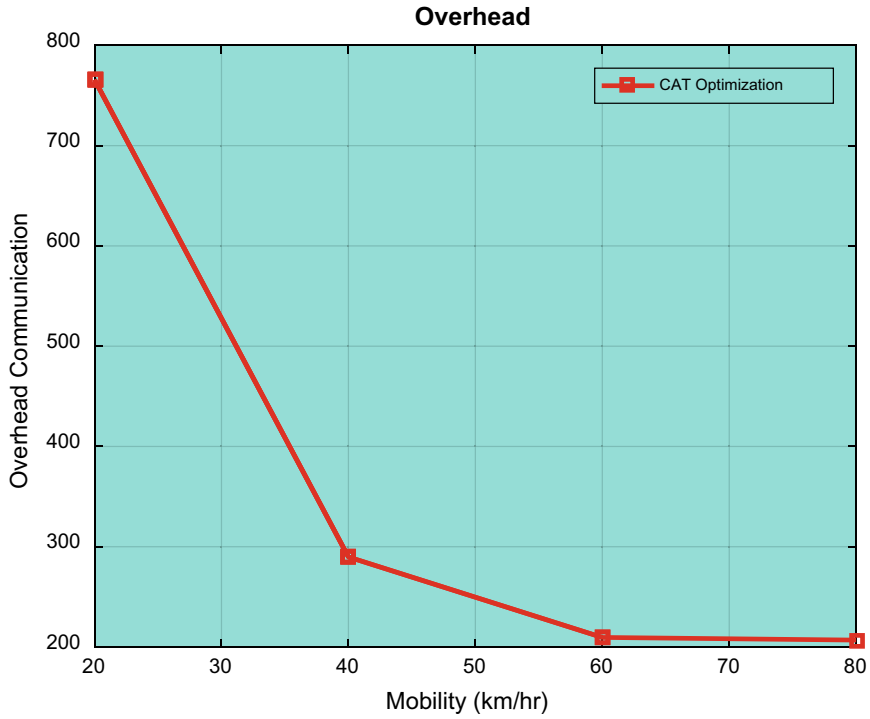


Fig. 4 Overhead consumption

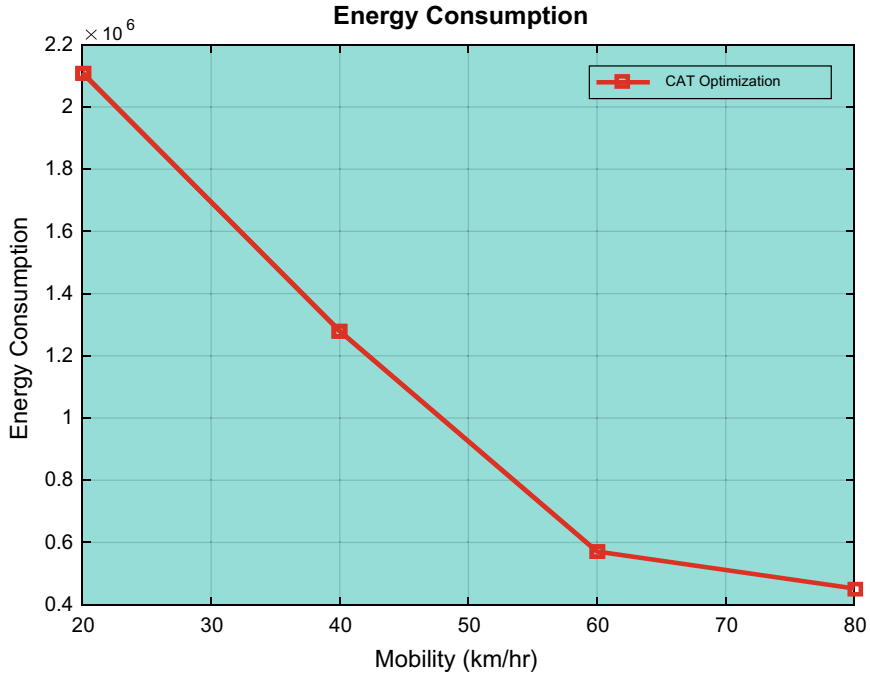


Fig. 5 Energy consumption

In the above figure, we can see the performance in terms of the routing rule which increase with an increase in the dynamic change in the topology using cognitive networks. Route rule is one of the significant parameters which divides

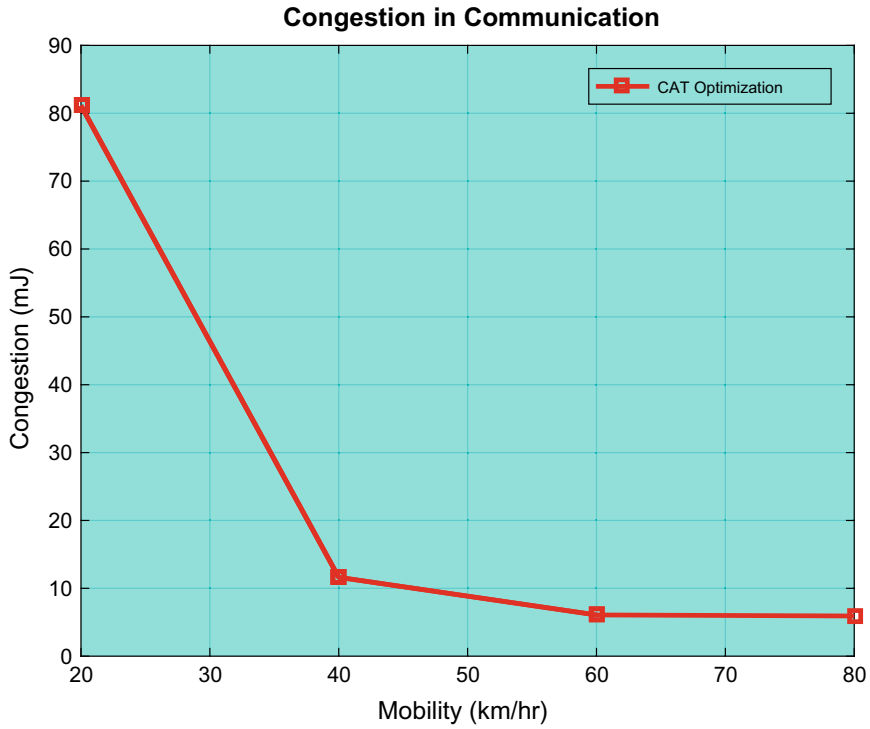


Fig. 6 Congestion

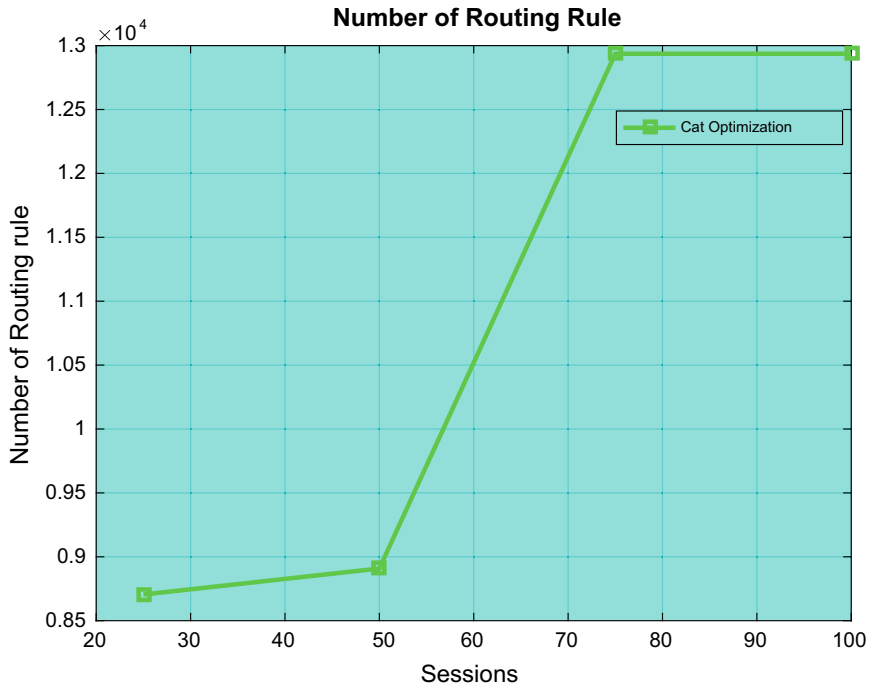


Fig. 7 Number of routing rule

the number of workflows into various threads and provides high hierarchy to the cognitive-based networks and is must be high.

Table 1 Performance analysis using CAT swarm optimization

Parameters	Proposed approach
Energy consumption	0.5×10^2 mJ
Congestion	5 mJ
Overhead consumption	200
Number of routing rules	1.3×10^4

Table 1 shows the performance analysis of CRN using Cat swarm optimization.

7 Conclusion and Future Scope

This paper explained the unique technique called Cognitive Radio Networks (CRN) that is becoming so much popular where the main aim is on active consignment of networks to wireless expedients. The routing in MANET with QoS improvement acquires an equivalent position in several analyses. Though, of these systems declared on top of depends on statically allotted channels either in authorized or unaccredited band. This static channel assignment proved spectrum congestion and thus spectrum shortage in a thickly occupied situation. Cognitive radio networks are introduced and a literature review related to CRN is discussed in the paper. An optimization algorithm to improve the performance of TE under CRN has been discussed. Swarm intelligence approach is used in the paper. Swarm approach is simply the combined part of the decentralized characteristic to obtain optimistic possible results. One of the optimistic approaches called Cat swarm has used to get a high rate of accuracy and fewer error rates which improves the lifespan of the network. The results are obtained using CSO (Cat Swarm Optimization) algorithm and parameters like energy consumption, congestion, overhead consumption, and number of routing rules are used to analyze the performance of the algorithm.

In future work, another Swarm optimization algorithm (BAT swarm) will be implemented to reduce congestion and overhead consumption. After getting results from BAT swarm hybrid model i.e. Cat Swarm +Bat Optimization will be achieved to get better results and comparison of this hybrid model will be done with the individual optimization techniques.

References

1. Cheng, G., Liu, W., Li, Y., Cheng, W.: Joint on-demand routing and spectrum assignment in cognitive radio networks. In: This full text paper was peer reviewed at the direction of IEEE Communications Society Subject Matter Experts for Publication in the ICC 2007 Proceedings, vol. 6501–6503 (2007)
2. Akyildiz, I.F., Lee, W.-Y., Chowdhury, K.R.: CRAHNs: Cognitive Radio Ad Hoc Networks. Elsevier Ad Hoc Networks, pp. 1–25 (2009)

3. Zubow, A., Döring, M., Chwalisz, M., Wolisz, A.: A SDN Approach to Spectrum Brokerage in Infrastructure-Based Cognitive Radio Networks. IEEE DySPAN 2015, pp. 213–221 (2015)
4. Alasadi, E., Al-Raweshidy, H.S.: SSED: servers under software-defined network architectures to eliminate discovery messages. IEEE/ACM Trans. Netw. **3**, 321–336 (2017)
5. Kaur, K., Garg, S., Aujla, G.S., Kumar, N., Rodrigues, J.J.P.C., Guizani, M.: Edge computing in the industrial internet of things environment: software-defined-networks-based edge-cloud interplay. IEEE Communications Magazine **5**, 44–51 (2018)
6. Tahaei, H., Salleh, R., Ab Razak, M.F., Ko, K., Anuar, N.B.: Cost effective network flow measurement for software defined networks: a distributed controller scenario. IEEE Access **12**, 501–514 (2017)
7. Hongli, Xu, Huang, He, Chen, Shigang, Zhao, Gongming, Huang, Liusheng: Achieving high scalability through hybrid switching in software-defined networking. IEEE/ACM Trans. Netw. **26**, 618–632 (2018)
8. Wang, Q., Zheng, H.: Route and Spectrum Selection in Dynamic Spectrum Networks. In: IEEE CCNC 2006 Proceedings, vol. 4, pp. 625–629 (2006)
9. Ali, A., Iqbal, M., Baig, A., Wang, X.: routing techniques in cognitive radio networks: a survey. Int. J. Wirel. Mob. Netw. **3**, 96–110 (2011)
10. Foerster, K.-T., Ludwig, A., Marcinkowski, J., Schmid, S.: Loop-free route updates for software-defined networks. IEEE/ACM Trans. Netw. **3**, 621–642 (2017)
11. Neama, G.N., Awad, M.K.: An energy efficient integral routing algorithm for software-defined networks. IEEE **5**, 401–406 (2017)
12. Asadollahi, S., Goswami, B., Raoufy, A.S.: Scalability of software defined network on floodlight controller using OFNet. In: 2017 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), vol. 23, pp. 557–561 (2017)
13. Karakus, M., Durresi, A.: Economic viability of QoS in software defined networks (SDNs). In: 2016 IEEE 30th International Conference on Advanced Information Networking and Applications, vol. 23, pp. 139–146 (2016)
14. Kuang, L., Yang, L.T., Wang, X., Wang, P., Zhao, Y.: A tensor-based big data model for QoS improvement in software defined networks. IEEE Netw. **21**, 30–35 (2016)
15. Körner, M., Stanik, A., Kao, O.: Applying QoS in Software Defined Networks by Using WS-Agreement. In: 2014 IEEE 6th International Conference on Cloud Computing Technology and Science, vol. 3, pp. 893–898 (2014)
16. Li, G., Wu, J., Li, J., Zhou, Z., Guo, L.: SLA-aware fine-grained QoS provisioning for multi-tenant software-defined networks. IEEE **7**, 121–134 (2017)
17. Ren, S., Dou, W., Wang, Y.: A deterministic network calculus enabled QoS routing on software defined network. In: 2017 9th IEEE International Conference on Communication Software and Networks, vol. 24, pp. 181–186 (2017)
18. Oluwaseun, A., Twala, B.: QoS functionality in software defined network. In: IEEE ICTC 2017, vol. 26, pp. 693–699 (2017)
19. Even, S., Itai, A., Shamir, A.: On the complexity of timetable and multicommodity flow problems. SIAM J. Comput. **4**, 691–703 (1976)
20. Al-Fares, M., Loukissas, A., Vahdat, A.: A scalable, commodity data center network architecture. In: ACM SIGCOMM Computer Communication Review, vol. 38, pp. 63–74 (2008)

(T-ToCODE): A Framework for Trendy Topic Detection and Community Detection for Information Diffusion in Social Network

Reena Pagare, Akhil Khare and Shankar Chaudhary

Abstract The increased use of social network generates a huge amount of data. Extracting useful information from this huge data available is the need of today. Study and analysis of this data generated provide insight into the behavior of the customers or users and thus will be beneficial to increase the sales of products or understand customers. To achieve the same, we propose a novel framework which will extract trendy topics, identify communities related to these trendy, topics, and also identify influential or seed nodes in communities. The framework intends to find the list of topics which are popular, second, find trend-driven communities, and from these trend-driven communities find nodes which act as seed nodes and thus dominate the spread of information in the community. Analysis of real-world data is done and results are compared with baseline approaches.

Keywords Community detection · Information diffusion · Topic detection · Trend topics · Social network

1 Introduction

Over the past few decades, we can see rapid change in technology and the way technology is used by the world. Social network has become an inseparable part of every person's life today. As the number of user increases day by day there lies great potential in analyzing the social network for the kind of relationships and the actors in the network. The study of social network is not only limited to the behavior of people, but it can be a relationship between people, organization, cells, genes, etc.

R. Pagare (✉) · S. Chaudhary
PAHER, Udaipur, Rajasthan, India
e-mail: reenawp5@gmail.com

A. Khare
MVSR COE, Hyderabad, India
e-mail: khare_cse@mvsrec.edu.in

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_43

Social network application expands over a vast horizon from psychology, sociology, marketing, communication, and political science.

Formally a social network can be defined as a network which is a collection of objects in which some pairs of these objects are connected by links. When the structure of the network is considered, it is more in terms of "who is connected to whom" and the other connection at the "level of behavior", one person can be friend of other, or he may be follower of the person, or he must have retweeted on a post or commented on a post, etc. This means that in addition to the language for discussing the structure of networks, there is also a need of a framework for reasoning the behavior and interaction in the network context.

Much research effort has been put into analyzing information diffusion, with most studies investigating which factors affect information diffusion, which information diffuses most quickly, and how information is disseminated. These questions are answered using information diffusion models and other methods, which play an important role in understanding the diffusion phenomenon. Information diffusion is about understanding who the important users are and which factors are influencing the information diffusion process. A good performance model is very important for understanding how to predict and influence information diffusion and has significant reference value to various applications, e.g., rumor controlling, behavior analysis, gaging public opinion, the study of psychological phenomena, and for resource allocation in public health care systems.

The social influence in a network can be identified through tracking communities and studying the behavior in the community. Social influence was studied in the context of influence maximization where the method was proposed to find the nodes in the network which will increase the spread of information in the community and hence in the network [1, 2]. Their work highlights the social diffusion model and marketing strategies for various applications [3, 4]. The work of [5] relates to the study of how one person influences the other person for buying of products. The author suggests a method to increase this influence to increase the spread of information in the network. Social influence is defined as the force exerted on a person due to her influencer and homophily is considered as the tendency that the other inactive node or person will follow the active node or influencer [6, 7, 8].

The work done in the above methods only exploit the topology of the network, and ignore other important properties, such as nodes' features and the way they process information. Most of the users in social network are in passive mode, based on this observation, researchers [9] suggested an approach similar to HITS algorithm which is a graph-based approach. The method assigns to every node in the network a value based on the influence and the rate at which the information is forwarded in the network. The users who work as a catalyst in spreading of information are called the seed nodes. These seed nodes are specific to a community. Therefore, Pal and Counts [10] develop a non-graph-based, topic sensitive method. The method uses structural as well as social attributes of the network to identify the most influential nodes. The set of identified users are clustered by using probabilistic technique and

therefore could find a list of nodes which are more in the network related to the topic. In paper [1], the author uses independent cascade and linear threshold model for influence maximization. The influence maximization problem asks for a parameter "k", to find a k-node set of maximum influence in the network. Depending on the basis of number of nodes activated at the end of information cascade, the influence of a given set of nodes is identified. The author uses a greedy hill-climbing strategy for the optimization problem and provides appropriate solutions.

A **Social network** can be defined as a graph in which vertices represent people or groups of people and edges represent some kind of social relationship or social interaction between vertices. These social graphs can further be defined as the ones in which direction to the links or edges may or may not be known as directed or undirected graphs like Twitter and Facebook, respectively [11].

A **community** in the social network is a set of nodes or users which can be potentially grouped and examined together such that the interactions or the ties between the communities are quite dense [1]. In terms of the graph, we can say it is a giant connected component that is very dense. The users in social network exhibit a behavior where users ignore the information available with themselves and are influenced by the behavior of the neighbor's. This decision taken is transferred in the form of information to all the connected users in the networks. This phenomenon is known as **information diffusion** in social networks.

Information diffusion is a process in the social networks in which the information is transferred from one user to the other, who are connected with each other via social network relationship. Information diffusion includes (a) Detection of popular topics (2) Community Detection (b) Identification of Influential spreaders (Fig. 1).

Even though the framework associated with social networks can change with time, communities stay fairly steady. The primary problem is how one can identify individual's communities which have higher influence inside a social network, and several techniques are suggested for this; some techniques use links between users and some use network attributes. Information diffusion is a hot topic within social networks investigation recently. Even though there has been numerous revolutionary research in this area, there are still some less explored areas.

The remaining paper is arranged as follows: second part is related to state of art; third part describes our framework; fourth section presents result and experimentation; last section presents conclusion and future work.

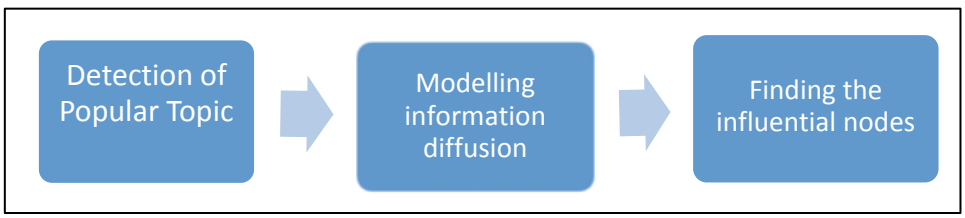


Fig. 1 Processes involved in information diffusion

2 Related Work

2.1 Information Diffusion

In paper [12], theauthor explores connections as well as designs developed by the aggregated interactions within Facebook webpages throughout catastrophe reactions. Evaluates social roles as well as crucial gamers utilizing social network evaluation. Research suggests measures to enhance information diffusion by considering social as well as network attributes. In Paper [13], theauthor suggests a differential equation model associated with information within thenetwork. The model is associated with temporal–spatial information diffusion. This model is set up prior to the topological framework associated with the network, semantic content material as well as interactive actions associated with users within thenetwork. Experimental calculations display the feasibility from the suggested model, conformity along with network topology with potential users associated with scalability with regard to big networks. The authors in paper [14] suggest parameters which evaluate topical importance as well as social. The work suggests that to study the information diffusion, it is important to study the behavior of user by taking into account different parameters like retweet count, followers information. In [15], through mixing users social as well as picture information, the model forecasts the actual diffusion route of the image much more precisely when compared with alternative baselines, which possibly encode just picture or even social functions, or even absence storage. Through mapping person users to person prototypes, the model generalizes for new users. The models are able to produce diffusion trees and shrubs, as well as display how the produced trees and shrubs carefully look like ground-truth trees and shrubs. Paper [16] detects communities based on not only structure properties of the network but also used ground-truth information. The author studied a collection of 180 real-world social, effort, and data networks exactly where nodes clearly condition their own community subscriptions. The suggested technique accomplishes 30% relative enhancement more than present nearby clustering techniques. In [17], the author proposes a CID model with regard to damaging information diffusion within competitors towards an optimistic information circulation. The model is an aggressive information diffusion model or even CID with regard to brief. This is the first model which talks about negative as well as positive information flow in the network and propagation of the same. The authors [18] proposes a grading opinion diffusion model. The opinion leaders play an important role in information diffusion in microblogs. The model is combined with communication power. The opinion leaders influence people who are unsteady concerning specific event. The work demonstrates how the brand new suggested models tend to be sensible as well as efficient for information spread within microblog information diffusion. The actual Information Systems (IS) community is constantly on the help to make inroads to improve using systems to aid catastrophe administration [19]. The actual community is constantly on the determine as well as lengthen the actual appropriate hypotheses, and also to create brand new paradigms

that may be delivered to keep about the ownership as well as diffusion associated with information techniques with regard to catastrophe administration. Research by authors in [20] handles the problem of advertising data diffusion, the degree to that data spreads, upon social media systems. This specific A–B–C framework (1) examines the textual options that come with Internet record articles making use of linguistic issue as well as phrase rely, (2) does apply the prior results in creating information suggestions, as well as (3) suppliers verified training supplies based on information tips to promote information diffusion among Internet record site visitors. The author suggests an evolutionary game theoretic framework [21] to model diffusion procedure within social networks. The framework is analyzed for degree and nonuniform degree network. Present social networks tend to be of very large scale producing huge data moves for each and every second. Exactly how data diffuses more than social networks offers has attracted a lot of interest. Author obtains diffusion characteristics within just complete networks, typical degree, without homogeneous degree networks, with identify connected with two distinctive networks. The authors in their work [22] depicted a person's behavior of online social networks opens new alternatives to systematically realize the data diffusion process after social networks. Random Recursive Tree (RRT) is used to help model with this advancement associated with Cascade trees. The use of stochastic sensitive look at forwards eliminated model to help demonstrate the particular potent individual behavior including producing, seeing, forwarding, and overlooking the information for the furnished social network.

The authors argued Twitter data is very loud [23] every twitter is brief, unstructured and along with casual vocabulary, challenging with regard to present subject modeling. About the additional hand, Twitter posts tend to be associated with additional data, for example, authorship, hashtags and the user-follower network. Taking advantage of this information, the author suggests the Twitter-Network (TN) subject model collectively model the written text and the social network inside a complete Bayesian nonparametric method. The TN subject model utilizes the hierarchical Poisson–Dirichlet Procedures (PDP) with regard to textual content modeling and the Gaussian procedure Random function model with regard to social network modeling. The author implies that the TN subject model considerably outperforms a number of current nonparametric models because of its versatility. Furthermore, the TN subject model allows extra educational inference, for example, authors' interests, hashtag evaluation, in addition to resulting in additional programs, for example, author suggestion, automated subject marking, and hashtag recommendation. Be aware the common inference framework may easily be reproduced to additional subject models along with inlaid PDP nodes. Another work [24] uses Louvain community detection algorithm and Girvan algorithm to find communities and centrality measure to find the nodes in the communities which help in the diffusion process. The author also uses the shortest path algorithm to find the communities formed using the seed nodes. The work considers the network attributes for community detection. The communities formed are not dense and are formed on the basis of only structure of the network. Does not consider the content in the network. As per study, in [25], twitter comprises a good obtainable system with

regard to learning and tinkering with the character of data dissemination. The author uses hashtag to find topics which disseminate most in the network. As per [26], it is vital to find out, monitor, review, as well as forecast well-liked subjects and occasions happening within the social networking within the space–time framework. Simultaneously, it's very helpful which a number of "what if" situations could be created to estimation the meme diffusion. The author has made use of location and time attribute to study the diffusion. The work [27] represents impact maximization issue in attempting to recognize some ``K'' nodes through that distribute associated with impact, illnesses, or even data is maximized. The author suggests a ``Greedy ad Community-Based'' formula.

2.2 Topic Detection

The paper [28] proposed LDA model which is the most popular and efficient model for topic modeling. But it does not consider the correlation among topics. It works only on content and is a probabilistic model that considers the distribution of words for identifying popular topics. An online LDA model is proposed [29], it considers the document in batches. Thus the memory requirement is reduced. It does not consider the social attributes of the network. Considers the distribution of word for topic definition. Takes only content into consideration while modeling. The authors [30] propose a model using wavelet analysis, detects event inferred using LDA gives a better description of the event and uses hashtag for topic inference. Topic popularity is found by using only content, the network attributes too can be considered to improve the accuracy of the model. In his work [8], the author proposes a model for unbounded topics. The hierarchical structure can help to identify popular topics. The model does not work for definite and bounded topics. A topic model based on time [31] is implemented, the model accuracy can be improved by consideration of other attributes of the network like retweet count and followers count. A parallel algorithm for topic modeling [32] gives a good speed up as compared to the sequential approach, but the resource requirement is huge. The model can be implemented on a small scale for less number of processor and it is important to study the behavior of the algorithm.

Most of the work done, find trendy topics irrespective of the social attributes and take content only into consideration, some works find communities and then find topics which are discussed mostly in these communities. In our proposed work, we make an attempt to identify trendy topics based communities that have played an important role in making the topic trendy. Given this real-time scenario, our objective is to (1) extract trendy topics (2) find communities related to these trending topics (3) identify the influential node in this topic-driven community.

To the best of our knowledge, our work is the first attempt toward presenting a framework which will identify trendy topics, find trend-driven communities, and then find the most influential nodes in the community. The topic-driven communities identified can be used to spread message or information related to the event or these set of users can be used for the marketing of products or to increase sales of

products. The work finds application in many areas related to identifying users for marketing, identifying users who spread wrong information in a closed network, or terrorist network. In order to detect communities related to trending topics, we focus on twitter, the widely used micro-blogging platform. The activity of twitter is depicted by tweets, retweets, comment, likes, and its structure of follower and followed unidirectional relationships.

To spread information in the network, it is needed that the users have information related to the topic. This information can be received by connecting with users having similar topic information. Thus forming of groups or communities of similar interest, users will be effective for spreading of information in the network. Accordingly, we present a method to identify these communities of topic-dependent users who created the post related to the topic Uc. Another important parameter of our work is followers users U_f information of these users Uc. These U_f users will be responsible for the spread of information in the network.

To summarize, given a trending topic, the objective is to detect community of users who (1) are related to the trending topic (2) are related to the same topic (3) are connected to each other (follow) (4) can be identified as U_c or U_f.

3 Proposed Work

3.1 Problem Formulation

Definition 1

Let $T = \{t_1, t_2, \ldots t_m.\}$ *be the set of tweets collected during a particular period.*

Each tweet $t_i = \{w_1, w_2, w_3, \ldots w_i \ldots w_n\}$ *where* w_i *represents the sets of word in the tweet* t_i.

Analyzing these tweets T for the word occurrences in each tweet and document as whole and finding which words w_j *are maximum talked about is topic detection.*

Where w_j *is subset of* w_i, $w_j \in w_i$
These w_j *words are the set of Trendy topics,* Tr.

The first goal is to identify this Tr set

Definition 2

Let $\mathrm{Tr} = \{\mathrm{Tr}_0, \mathrm{Tr}_1, \mathrm{Tr}_2 \ldots \mathrm{Tr}_9\}$ *be the set of trendy topics. Here we assume that we identify top 10 topics only. For each topic* Tr_i *belonging to* Tr *(a) find set of users U who have tweeted, retweeted about the topic* Tr_i *(b) Generate a graph G, of these users* u_i *from U where* u_i *belongs to U. The graph G formed represents a community* Ctr_i *which is related to topic* Tr_i. *Each vertex in community* Ctr_i *is related to user* u_i *who in turn is related to topic* Tr_i. *There exists an edge between* u_i *and* u_j, *where* $(u_i, u_j) \in U$ *iff* u_j *is in followers list of* u_i *or vice versa or* u_i *retweeted on* u_j *The users* $u_{z \in U,}$ *who are not a part of the connected graph are discarded.*

The second goal is to generate graph G (V, E) of users related to topics. This graph represents a set of users who will play an important role in identification of communities in next phase.

Definition 3
For the graph G generated for topic Tr_i *we run community detection algorithm to identify community. The community detection algorithm will identify communities associated with topic* tr_i*, which will be used to identify node with highest degree,* $u_{i \in U}$

A degree here means the node which will have the highest number of connections. The node having the highest connection is the one who will spread maximum information in the community.

To summarize, our contributions

(a) A model to find trend-driven communities and influential nodes.
(b) A unique framework which combines topic detection and topic-driven community detection.

3.2 Overview

We propose a novel framework where both topic modeling and community detection are collaborated to study the information diffusion in the network. The topic modeling module is the first module which takes as input tweets dataset and identifies topics which are trendy. This module can also be called as trendy topic detection. The second submodule is related to forming communities which are related to a particular topic. The communities formed are called as topic-driven communities. We name them so because the communities identified are related to the topic. After identification of communities, the nodes with highest degree centrality are identified since these are the nodes which influence the propagation of information in the network. Thus the framework will not only detect trendy topics but will also help to identify users and communities which play a substantial role in spreading of information (Fig. 2).

3.2.1 Data Collection

The Twitter dataset is used as input to the system. The twitter data is extracted using twitter API and stored in CSV form. The tweets collected have fields like retweet count, comment count, user who tweeted or retweeted about post, followers. We also store separate information about the users or owner (the user who first posted the tweet) of the tweet. We maintain separate information about the user—follower relationship.

The twitter data was collected from March 2018 to August 2018. Due to restriction by twitter, only 15 lacs of tweets were collected. The data collected included fields like user id, username, tweet, retweet count, mention count, followers of users, and followers count. Figure 3 shows the distribution of messages over different domains. The diagram below shows that floods, health, and politics topics were mostly covered in the tweets captured.

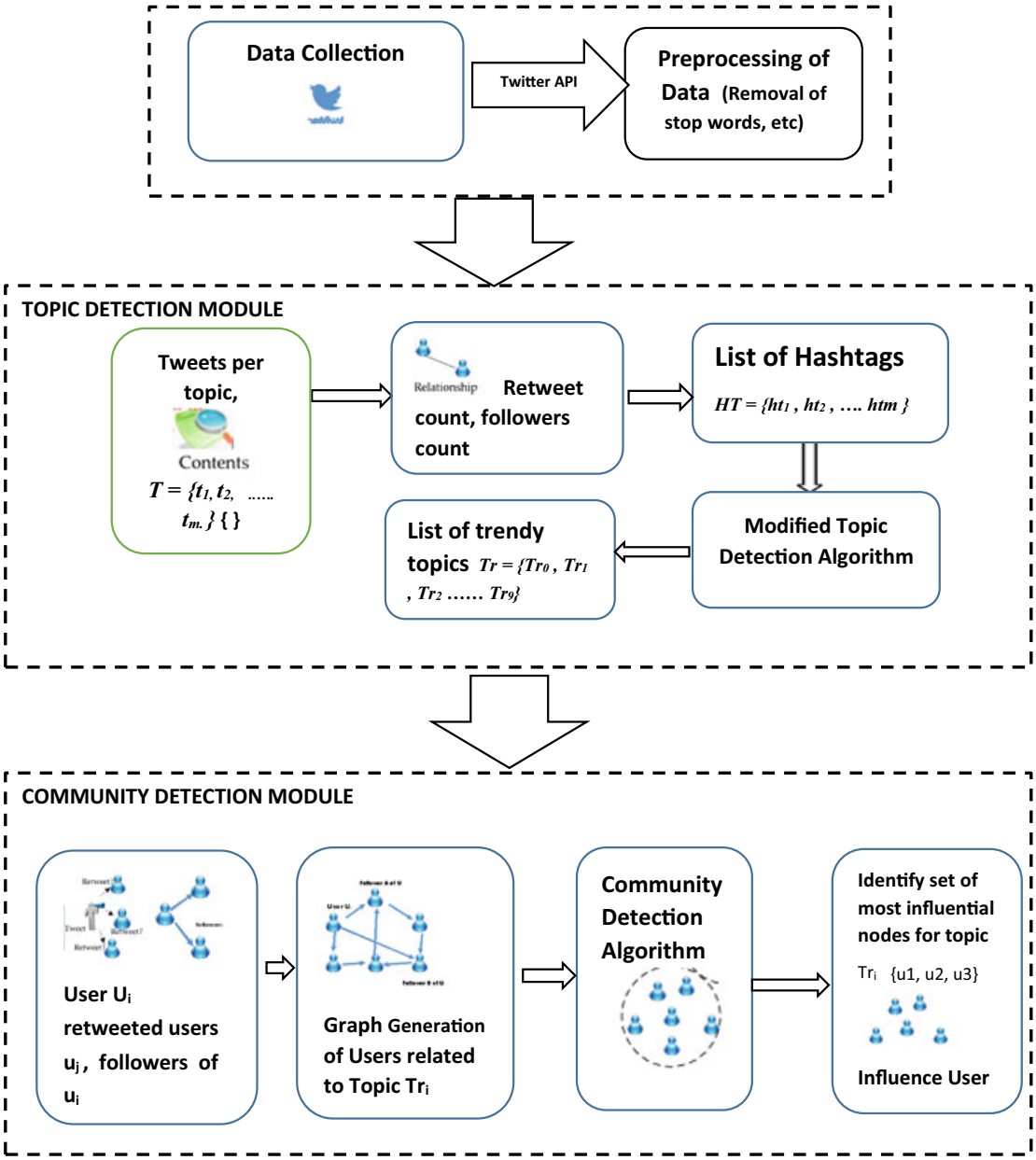


Fig. 2 Collaborative framework design

3.2.2 Data Preprocessing

In this step, the noise in the data is removed. The preprocessing deals with stop word removal, repeated data removal, filling of missing data by taking a mean of the values.

The next section provides details of our approach named T-ToCODE (**T**rendy **T**opic detection and trendy **CO**mmunity **DE**etection) which is able to detect hot topics and second the communities associated with every popular topic.

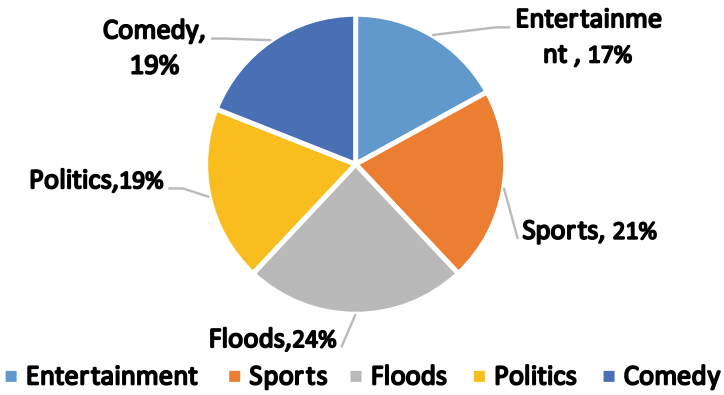


Fig. 3 Distribution of messages in the dataset over different domain

3.2.3 Trendy Topic Detection

Latent Dirichlet Allocation (LDA) is a classic model for topic modeling of documents. LDA detect underlying topics in text documents. LDA is based on the assumption that documents with similar topics will use similar groups of words. Documents are probability distributions over latent topics. Topics are probability distribution over words. Thus LDA is an unsupervised, probabilistic model. We propose a novel approach for finding trendy topics, where we consider the hashtag, comment count, retweet count, and likes for a particular post to identify if the topic is trendy or not. We compare our modified Multiparameter LDA (MP-LDA) with Twitter LDA and baseline LDA.

The proposed algorithm is mentioned below.

Generative Process for MP-LDA:

1. Initialize: Time distribution value of L_w, vector δM , δV.
   // word distribution for topic t
2. For each topic t = 1,2, …… T
   a. Draw $\emptyset\, t \sim dir(\beta)$
   End for
3. For each user u = 1,2, ….. U
   a. Draw $\theta u \sim \text{dir}(\alpha)$ *// topic distribution of user u*
   b. For each document/tweet s = 1,2,…….N_u
      i. Draw $Z_{u,s} \sim \text{Multi}(\theta u)$ *// Topic selection based on users topic distribution*
      ii. Draw $V_{n,s} \sim \text{dir}(\alpha)$
      End for
   c. For each word n = 1, 2,….. $N_{u,s}$ *// words are selected one by one based on value of P*
      i. compute P = Pw || Pd according to Lw
         if P = 0 then
            draw $Y_{u,s,n} \sim \text{Multi}(\vartheta g, u)$
         draw $W_{u,s,n} \sim \text{Multi}(\Phi z, Y_{u,s,n})$
      if P = 1 then
         draw $Y_{u,s,n} \sim \text{Multi}(\vartheta h, m)$
         draw $W_{u,s,n} \sim \text{Multi}(\Phi_{y,u,s}, \delta m, nu)$
      end if
   end for
end for

Let $\emptyset t$ be word distribution for topic t. P parameter is used to decide word belongs to a general topic or hot topic. The hashtag in the model is associated with a topic and word. δM is the hashtag associated with documents and δV is hashtag associated with word. The values that vectors can take is either 0 or occurrence count over the entire dataset. Latent variable learning is done by using Gibbs sampling. The MP-LDA is applied after preprocessing step to only those tweets which are above the threshold value of popularity count, comment count, and retweet count. All the tweets with parameter values below the threshold are not considered. Tweets with lesser retweet count or likes will not be popular and thus processing of these tweets for identifying popular topics is not required. The selection of threshold for retweet count, comment count and likes is based on the empirical analysis. It is seen from the statistics that most of the messages have less than 50 retweet and comment count. As the count reaches 1000 the distribution in number of messages is almost constant. Therefore we choose the threshold as 1000. After the application of MP-LDA to the dataset, we were able to find the topics which were most popular in the dataset. The topics like yoga for health, women's health, Asian games, US open championship and Kerala floods, Japan floods were found to be the trendiest topics.

3.2.4 Topic-Driven Community Detection

The steps involved in this submodule are given below.

Input: Topic Tr_i
STEP A: Identification of posts $P_t = \{P_1, P_2, P_3 \ldots P_l\}$ which are related to topic Tr_i
STEP B: Identification of user sets U_o who created the post, related to popular topic Tr_i
STEP C: Identification of user sets U_r who retweeted. The set identified in STEP B and STEP A together will be merged to form a set U_{rcl} where $U_{rcl} = U_r + U_o$
STEP D: Identification of followers of users U_f who created the post P_l. The followers play a crucial role in the propagation of information and thus follower's information will be useful.
STEP E: Graph creation based on user- follower data, user—U_{rcl} data
STEP F: Detection of communities related to a topic
Output: Community Ctr_i

STEP A:

Given a set of tweets $t = \{t_1, t_2, t_3, \ldots, t_m\}$ related to a topic. Identify tweets ti which are related to popular topic Tr_i. The data set of tweets will be searched and it will retry tweets related to these popular terms.

The output of this step will be a set of tweets $T_{Tr} = \{t_1, t_2, t_3, \ldots . t_j\}$ where $j < m$ and $t_j \in t$

STEP B:

From stepA all the tweets from set TTr are parsed to identify the tweets or post which are source post. The tweets are parsed to get the user id of the users who have

tweeted about a topic and is the origin of the post. This step will give output a set of users.

$U_o = \{U_{o1}, U_{o2}, ..U_k\}$ such that $U_{oi} \in U$ and U_{oi} is the owner of the post and $r < m$

STEP C:

Given a set of tweets T_{Tr}. Identify users who retweeted Ur on post in T_{Tr}, This third step gives a set of users who have retweeted a post.

At the end of step C, we get a set of users $U_{rcl} = \{U_o, U_r\}$ who may play an important role in the formation of communities

STEP D:

The belief that followers will play a substantial role in spreading message, the followers list is identified. Given a set of users Uo, find the follower of Uo and store it in the form (user id, user id follower[]). The set is identified as U_f

Also, a separate data related to User—U_{rcl} user data is stored. It contains those users who commented, retweeted a post by user.

STEP E:

Graph Generation:

Step D gives the vertices of the graph G. The set U_f (User - Follower) and the set U_m (user—U_{rcl}) will form the vertices of the graph.

Vertex set $\mathbf{V} = \mathbf{U_f} \cup \mathbf{U_m}$; union of set U_f and set U_m

The edge between a set of vertices can be of two types. The edge set is represented by E. User u1 and user u2 there exists link between u1, u2 if and only if

Type 1: User u2 retweet on user u1 post

$$R = \{\text{retweet}(u1, u2) = \text{TRUE}; u1, u2 \in U_m\}$$

Type 2: User u2 is follower of user u1

$$F = \{\text{follow}(u1, u2) = \text{TRUE}; u1, u2 \in U_f\}$$

The set of edges $\mathbf{E} = \mathbf{R} \cup \mathbf{F}$

On the basis of links and vertices a graph is formed. The belief that users who retweet or are followers of a particular user say u1 has more probability of spreading message which is posted by user u1 is used while formation of the graph G(V,E).

STEP F:

On the graph generated in step E, we apply Louvain community detection algorithm [33].

The decision of applying community detection algorithm was done because it can handle in a very short time a huge network. Also, the algorithm reveals hierarchical structure which can be useful in understanding the overall functioning of the network.

The Algorithms working is explained below

1. A greedy assignment of nodes to communities, favoring local optimizations of modularity.
2. New—coarse-grained network in terms of communities found in the first step.

The above two steps are repeated until no further modularity increasing reassignments of communities are possible.

3.2.5 Identification of Influential Nodes

Once the communities are identified finding nodes which are having a higher degree in the community is obviously the node which is going to be more responsible for spreading of information in the network. Average degree is used to find nodes which are most influential in the network.

The list of influential users is given by $U_{\text{In}} = \{U_1, U_2, \ldots U_x\}$ where $U_i \in$ to U

4 Experimentation

4.1 Topic Detection

4.1.1 Dataset

We use twitter dataset for experimentation. It consists of 132,009 posts/re-posts as well as 73,257 person nodes. Table 1 gives statistics of dataset. The data collection is as mentioned in Sect. 3.2.1

4.2 Analysis

Table 2 gives the list of hot topics and top words in each topic. When the result was compared to LDA and T-LDA it was observed that the top words in MP-LDA were more accurate as compared to T-LDA and LDA. For example, the first topic has words like Japan under topic FLOOD, but LDA and T-LDA did not return this word as a top word in the list. If we consider the result and compare with trending topic during the period, there are tweets related to floods in Japan during the same time as in India in Kerala.

Table 1 Statistics of dataset

Total tweet	Original tweets	Retweet	Users
1,32,009	9600	1,22,409	73,257

Table 2 Top topics with top words

Topic id	Top words
0	Flood rescue stressed relief Kerala donation Japan Gordon
1	Sports Asian Games, Indians US open, football Cleveland baseman, Serena
2	Health, women's health, Rohingya, yoga, mental health, medicine, organic, weight loss

Table 3 Coverage rate

	Top 10	Top 20	Top 30
MP-LDA	0.66	0.75	0.78
T-LDA	0.45	0.50	0.53
LDA	0.23	0.35	0.40

We also compared the coverage rate of T-ToCoDE with the baseline LDA method and Twitter LDA. It was observed that the result obtained by our method which considered social attributes gave better accuracy as compared to traditional methods.

$$\textbf{Coverage rate} = (\textbf{Extracted Hot Topics}/\textbf{Actual Hot Topics}) * \textbf{100}$$

The comparison was done for the top 10 topics, consecutively for 20, and then 30 topics. The proposed method showed consistent improvement in accuracy (Table 3; Fig. 4).

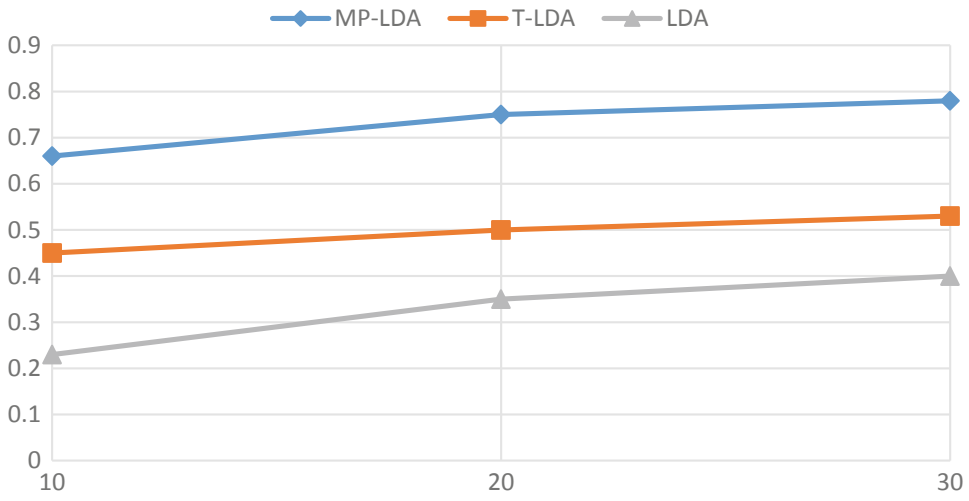


Fig. 4 Coverage rate versus no of topics for twitter data

4.3 Community Detection

4.3.1 Dataset

The trendy topic is given as input to the community detection module. The first topic which was trendy as per the topic detection was related to hashtag flood. First, we collected all tweets which were related to hashtag flood. This gave a list of 30,000 tweets related to floods. The captured tweets were then separated for the original tweet and retweet. It was found there were 2930 original tweets and remaining were retweets for flood dataset. From this set of original tweets identified, we considered only those users original tweets which had retweet count more than the threshold. After applying the threshold, we were left with 670 users who were original tweet users for flood dataset. We extracted the followers of these original tweet users. To this collected data then was Louvain Community detection method applied to find the communities related to Topic FLOOD.

The communities formed are evaluated on the basis of disconnected users, cohesiveness among users, and clustering coefficient. We compared our work in three different scenarios.

Scenario 1 (M1): The entire network is considered irrespective of tweet and retweet and an attempt was made to identify the communities. No threshold was applied for the selection of users. The graph was constructed on the basis of the followers and the followee network. Dataset considered for this method had all the 23,257 users included.

Scenario 2 (M2): The graph was constructed on the basis of retweet count and the retweet relationship and threshold considered was 1000. For the tweets collected for hashtag flood, out of 30,000 tweets, there were 2930 original tweet users, 7640 retweet users, and the remaining were duplicate users. There were a total of 10,570 users. When the threshold was applied to the original set of users, only 670 original users were considered for graph construction. The edge was considered between two users if user u1 retweets about the post by user u2 and vice versa.

Scenario 3 (T-ToCODE): The graph was constructed as per the proposed method in the framework. After the threshold applied, there were 670 users. For these 670 users, the followers list was considered and the graph was constructed on the basis of retweet as well as followers information. Due to the restriction by twitter, we could retrieve followers list for only 200 users and the total user were 41,235.

4.3.2 Metrics

Disconnected Users: This metrics will help to find users who are not part of the community

$$\text{ratio of disconnected users} = \frac{\text{users who are not connected to any other user}}{\text{total no of users}}$$

Cohesiveness among users: The second metric used is cohesion among users, this metrics allows to evaluate the quality of connection among users by using standard metrics such as modularity, the ratio between no. of communities and number of users, and density.

The ratio between the no. of community and no. of users allows to evaluate the ability of the approach to group the individual users into communities.

Lower the value, the better communities are formed, it signifies that communities are less in number with more number of users in each community.

The density is the ratio between the numbers of edges per node to the number of possible edges.

Clustering Coefficient: Clustering Coefficient quantifies the extent to which nodes in the graph tend to cluster together.

4.3.3 Analysis

Analysis of Disconnected Users

The result showed that 80% of users took part in the topic with our approach as compared to 76% with scenario 2 as compared to 72% of scenario 1. These results demonstrate that our graph construction approach includes more no of users.

Our approach considers the followers also thus increasing the number of users who can contribute toward spreading of information. In scenario 2, the graph is constructed using only retweet users thus leaving out a good percentage of users who may contribute toward information spread. Though in scenerio1 there are more numbers of users, due to consideration of the entire network a large number of users are disconnected users.

Analysis of Cohesiveness Among Users

Density is influenced by the fact that T-ToCODE framework share retweet, followers link and therefore has higher density compared to approach in scenario 2. The average number of communities found in Scenario 2 is 0.28 per user that exceeds the amount of community in T-ToCODE (Table 4).

Table 4 Cohesiveness among the users

Method	Modularity	Ratio of communities per node	Density
Scenario 1	0.420	0.321	0.010
Scenario 2	0.530	0.278	0.020
T-ToCODE	0.780	0.173	0.029

Table 5 Analysis of the structure of community (average)

Method	% of Post creators	% of followers	Clustering coefficient
Scenario 2	4.14	6.87	0.088
T-ToCODE	5.20	8.23	0.078

Analysis of Community Structure

The percentage of users who are the owners of post is more in scenario 3 (Proposed method) as compared to scenario 2 where only retweet relationship is considered. The increase in scenario 3 is obvious since we consider the retweet relationship along with follower's relationship. T-ToCODE creates communities on the basis of retweet relationship and followers relationship; therefore, the followers percentage is observed more as compared to scenario 2. Since the communities formed are related to a topic and the communities are dense thus the clustering coefficient is also improved in T-ToCODE method (Table 5).

After identification of communities, the nodes with the highest degree were found in the network. For finding these nodes, we used GEPHI visualization tool. It was observed that the nodes which had the highest degree on an average were the nodes which were the most influential nodes in the network. The entire community could be reached through the users who were having a higher degree.

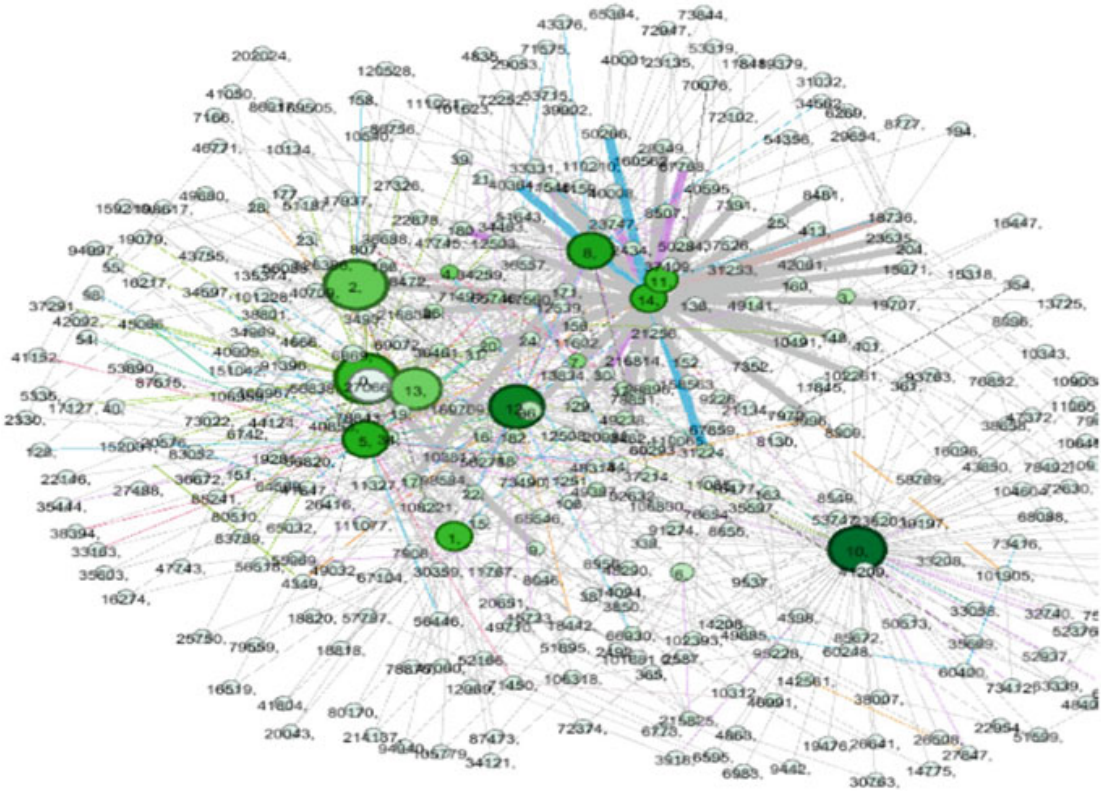


Fig. 5 Influential nodes for topic FLOOD: Scenario 3

The fig. 5 shows the influential nodes in the network ;nodes 8 , 11 and 13 are more dominant in the network. The edge thickness represents the ties among the nodes. From the graph, it can be observed that the nodes 8, 11, and 13 are the influential nodes in the community, along with a few other nodes in the network. There are communities identified for topic FLOOD, and there are nine influential nodes identified, the nodes which have the highest degree. The nodes at the center of each community are the node with the highest degree and thus they are the influential nodes in the network.

5 Conclusion

The novel framework proposed allows to identify hot topics and the communities which are influential in making those topics trendy. The analysis did prove that communities formed with relation to topics have more strength and lesser communities with more no of nodes are formed. The communities formed are denser and less overlapping. The topic detection model performs better as compared to the traditional model. The accuracy of topics identified is increased due to the inclusion of social attributes like retweet count, comment count and hashtag. Consideration of behavior of users in the network gives better results is proved.

5.1 Future Scope

The framework can be extended by considering the spatial and temporal attributes of the data. Thus adding these attributes can give a more personalized analysis, which can be used by an individual as well as businesses to target the customers.

Acknowledgements The authors declare that they have no conflict of interest. All procedures performed in studies involving human participants were in accordance with the ethical standards. Informed consent was obtained from all individual participants included in the study.

References

1. Kempe, D., Kleinberg, J., Tardos, É.: Maximizing the spread of influence through a social network. In: Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ‘03), pp. 137–146. ACM, New York, NY, USA (2003). http://dx.doi.org/10.1145/956750.956769
2. Bharathi, S., Kempe, D., Salek, M.: Competitive influence maximization in social networks. In: Deng, X., Graham, F.C. (eds.) Proceedings of the 3rd international conference on Internet and network economics (WINE'07), pp. 306–311. Springer, Berlin, Heidelberg (2007)

3. Bonchi, F.: Influence propagation in social networks: a data mining perspective. In: 2011 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology, vol. 2(1) pp. 2–2 (2011)
4. Domingos, P., Richardson, M.: Mining the network value of customers. In: Proceedings of the Seventh {ACM} {SIGKDD} International Conference on Knowledge Discovery and Data Mining, pp. 57–66 2001. https://doi.org/10.1145/502512.502525
5. Leskovec, J., Adamic, L.A., Huberman, B.A.: The dynamics of viral marketing. ACM Trans. Web (TWEB) **1**(1), 1–39 (2007)
6. Goyal, A., Bonchi, F., Lakshmanan, L.V.S.: Learning influence probabilities in social networks. In: Proceedings of the Third ACM International Conference on Web Search and Data Mining—WSDM '10, p. 241 (2010). Available at: http://portal.acm.org/citation.cfm?doid=1718487.1718518
7. Saito, K., Nakano, R., Kimura, M.: Prediction of information diffusion probability ies for independent cascade model. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 67–75 (2008)
8. Chen, W., Wang, C., Wang, Y.: Scalable influence maximization for prevalent viral marketing in large-scale social networks. In: Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '10), 1029–1038. ACM, New York, NY, USA (2010). https://doi.org/10.1145/1835804.1835934
9. Liu, L., et al.: Modelling of information diffusion on social networks with applications to WeChat. Physica A Stat. Mech. Appl. **496**, 318–329 (2018)
10. Pal, A., Counts, S.: Identifying topical authorities in microblogs. In: WSDM '11, pp. 45–54 (2011)
11. Easley, D., Kleinberg, J.: Networks, Crowds, and Markets: Reasoning About a Highly Connected World. Cambridge University Press, New York (2010)
12. Kim, Jooho, Hastak, Makarand: Social network analysis: characteristics of online social networks after a disaster. Int. J. Inf. Manag. **38**(1), 86–96 (2018)
13. Tu, H.T., Nguyen, K.P.: Differential information diffusion model in social network. In: Asian Conference on Intelligent Information and Database Systems. Springer, Cham (2018)
14. Shi, J., et al.: Determinants of users' information dissemination behavior on social networking sites: an elaboration likelihood model perspective. Internet Res. **28**(2), 393–418 (2018)
15. Hu, W., et al.: Who will share my image?: Predicting the content diffusion path in online social networks. In: Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. ACM (2018)
16. Yang, Jaewon, Leskovec, Jure: Defining and evaluating network communities based on ground-truth. Knowl. Inf. Syst. **42**(1), 181–213 (2015)
17. Tu, H.T., Nguyen, K.P.: Competitive information diffusion model in social network with negative information propagation. In: Asian Conference on Intelligent Information and Database Systems. Springer, Cham (2018)
18. Liu, X., Liu, C.: Information Diffusion and Opinion Leader Mathematical Modeling Based on Microblog. IEEE Access (2018)
19. Beydoun, G., et al.: Disaster management and information systems: insights to emerging challenges. Inf. Syst. Front. **20**, 1–4 (2018)
20. Liang, Y., Kee, K.F.: Developing and validating the ABC framework of information diffusion on social media. New Media Soc. **20**(1), 272–292 (2018)
21. Jiang, C., Chen, Y., Liu, K.J.R.: Evolutionary dynamics of information diffusion over social networks. IEEE Trans. Signal Process. **62**(17), 4573–4586 (2014)
22. Romero, Galuba, W., Asur, S., Huberman, B.: Influence and passivity in social media. In: ECML/PKDD '11, pp. 18–33 (2011)
23. Lim, K.W., Chen, C., Buntine, W.: Twitter-network topic model: a full Bayesian treatment for social network and text modeling (2016). arXiv preprint arXiv:1609.06791

24. Jain, S., Mohan, G., Sinha, A.: Network diffusion for information propagation in online social communities. In: 2017 Tenth International Conference on Contemporary Computing (IC3). IEEE (2017)
25. Stai, E., et al.: Temporal dynamics of information diffusion in twitter: modeling and experimentation. IEEE Trans. Comput. Soc. Syst. **5**(1), 256–264 (2018)
26. Ye, X., et al.: Open source social network simulator focusing on spatial meme diffusion. In: Human Dynamics Research in Smart and Connected Communities. pp. 203–222. Springer, Cham (2018)
27. Jalayer, M., Azheian, M., Kermani, M.A.M.A.: A hybrid algorithm based on community detection and multi attribute decision making for influence maximization.". Comput. Ind. Eng. **120**, 234–250 (2018)
28. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. J. Mach. Learn. Res. **3**, 993–1022 (2003)
29. AlSumait, L., Barbará, D., Domeniconi, C.: On-line LDA: adaptive topic models for mining text streams with applications to topic detection and tracking. In: 2008 Eighth IEEE International Conference on Data Mining, Pisa, pp. 3–12 (2008). https://doi.org/10.1109/icdm.2008.140
30. Cordeiro, M.: Twitter event detection: Combining wavelet analysis and topic inference summarization. In: Doctoral Symposium on Informatics Engineering, DSIE (2012)
31. Liu, G., Xu, X., Zhu, Y., Li, L.: An improved latent dirichlet allocation model for hot topic extraction. In: 2014 IEEE Fourth International Conference on Big Data and Cloud Computing, pp. 470–476. Sydney, NSW (2014). https://doi.org/10.1109/bdcloud.2014.55
32. Wu, C., Wu, B., Wang, B.: Event evolution model based on random walk model with hot topic extraction. In: Li, J., Li, X., Wang, S., Li, J., Sheng, Q. (eds.) Advanced Data Mining and Applications. ADMA 2016. Lecture Notes in Computer Science, vol. 10086. Springer, Cham (2016)
33. Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. J. Stat. Mech. Theory Exp. **10**, P10008 (2008)

ns-3 Implementation of Network Mobility Basic Support (NEMO-BS) Protocol for Intelligent Transportation Systems

Prasanta Mandal, Manoj Kumar Rana, Punyasha Chatterjee and Arpita Debnath

Abstract In an Intelligent Transportation System for a Smart City, seamless connectivity is essential for each user during mobility for efficient data communication. For a group of mobile users in a vehicle (bus/train/flight), due to high mobility, implementing a protocol in order to manage handoffs smoothly is a challenging task. Network Mobility Basic Support (NEMO-BS) protocol was proposed to comply with this requirement. It is an extended version of Mobile IPv6 (MIPv6). But, the MIPv6 implementation in ns-3, which is the most widely used open-source simulator, is still not extended so far, to support network mobility. In this work, we have implemented the functionality of the NEMO-BS protocol in ns-3.25 by modifying the existing MIPv6 module to enrich the ns-3 library.

Keywords Intelligent Transportation Systems · Mobile IPv6 · Network · Mobility · Smart City

1 Introduction

In the era of Smart City and Internet of Things (IoT), each and every user needs to be connected to the Internet seamlessly. Intelligent Transportation Systems is one of the main aspects of the Smart City. In case of group mobility, i.e., when a group of

P. Mandal · M. K. Rana · P. Chatterjee
School of Mobile Computing and Communication, Jadavpur University, Kolkata, India
e-mail: prasanta.mtechdmc@gmail.com

M. K. Rana
e-mail: manoj24.rana@gmail.com

P. Chatterjee
e-mail: punyasha.chatterjee@gmail.com

A. Debnath (✉)
The Calcutta Technical School, Kolkata, India
e-mail: arpita.debnath03@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_44

users is moving in a vehicle (bus/train/flight), due to the mobility of the vehicle [1], the wireless interface of it sometimes switches the on-road access points. This event is called the IP handoff. This results in severe service degradation of the vehicular users. To handle it, the Internet Engineering Task Force (IETF) standardizes the Network Mobility Basic Support (NEMO-BS) protocol by extending the base Mobile IPv6 (MIPv6) scheme [2]. The objective of this is to provide uninterrupted Internet connectivity to the vehicular mobile users by maintaining ongoing sessions during handoff. UMIP (Usagi Patched MIPv6 stack) is an open-source implementation of MIPv6 and NEMO-BS protocols for Linux and can be used for network simulator ns-3. But it is only executed on ns-3 Direct Code Execution with the Linux native stack and that is why if any changes may be proposed in NEMO, then the users must have to trace the Linux stack. But, it is quite complicated to the naïve users to do any changes in Linux stack.

To fix the above issue, we have implemented the NEMO-BS protocol [3–6], using the ns-3 stack, by modifying the existing MIPv6 module [7, 8] in ns-3, that may be added to the ns-3 library which may help the ns-3 users a lot.

The paper is arranged as: Sect. 2 describes the Literature Survey, Sect. 3 describes the working principle of the protocol, Sect. 4 gives the description of the Class Diagram Implementation, Simulation and Results are given in Sect. 5 and finally, Concluding remarks emerge in Sect. 6.

2 Literature Survey

In [1], they proposed a group-based network mobility controlling technique to diminish the signaling problem. Besides this, they reduce the handover latency as well. But they did not investigate the finest mode for vehicles to be clustered together or they did not employ the neighboring association among the vehicles when alliance them.

In [8], they introduced a NEMO support protocol for Intelligent Transportation Systems that support mobility management as well as handover. But they did not support the guaranteed seamless connectivity during a handover.

Hager et al. [9], illustrate MINT- a Mobile Internet Router having enough computational capability to execute all essential communication protocol operations while enabling connections for the nodes. Transparency of the communication software is provided by the MINT router, and there is no requirement of any modifications with the basic software of mobility support while connecting via such a router to the Internet.

In the Request for Comments [10] which was on IP mobility support, clearly specifies the fact of using a Mobile Router for the mobile network.

By considering all the limitations of the above-related study, we try to implement NEMO-BS protocol using ns-3 network simulator that ensures the seamless network connectivity as well as the IP handoff management.

3 Working Principle of NEMO-BS Protocol

The working principle of the protocol is depicted in Fig. 1. The network entities for the handoff operation includes a Mobile Network Node (MNN), Mobile Router (MR), Home Agent of the MR (MR-HA), and Correspondent Node (CN).

Step 1: When MNN enters into a vehicle, it connects with the MR as MR is the default Internet Service Provider for the MNN inside a vehicle and acquires a Care-Of-Address (CoA), configured from the MRs prefix.
Step 2: As MNNs home prefix differs from that prefix, it binds the CoA with its HA (MNN-HA) through a Binding Update (BU) process.
Step 3: When the MR changes access point, it performs the BU process with the MR-HA. The BU processes result in two tunnels between entities (MR and MR-HA, MNN and MNN-HA).

A packet from CN to MNN follows the route: CN → MNN – HA → MR – HA → MR → MNN. It is encapsulated first by the MNN-HA such that it could reach the MR's home. Reaching MR's home, the packet is automatically redirected toward the MR-HA as the corresponding HA contains MR's home links. MR-HA has the binding of the MR's advertised prefixes to the MR's current CoA. So, the MR-HA encapsulates the packet, setting MR's CoA as the destination. The tunnel headers are decapsulated in the reverse order, i.e., in the MR first and then, in the MNN.

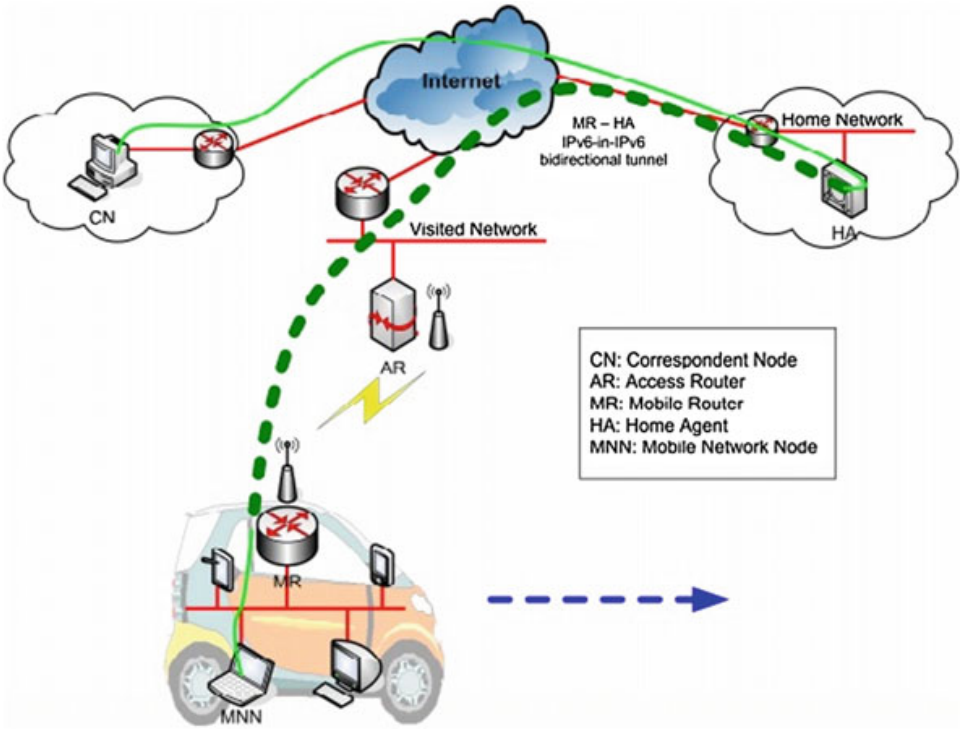


Fig. 1 NEMO-BS operation [4]

The working process of the NEMO-BS differs from the MIPv6 in the following two ways:

- The MR registers all its prefixes (which it advertises to the MNNs) during its BU process, instead of only the home address.
- Unlike MIPv6, the MR-HA must have to advertise all those MR's prefixes to its neighbors such that the MR-HA can receive all the packets, destined to the MNNs CoAs, configured from those prefixes.

Here, we will describe the Binding Update (BU) process and data packet processing in NEMO-BS in detail.

3.1 Binding Update (BU) Process of NEMO-BS

In MIPv6, when a mobile host configures a new address in its interface, it sends single BU to its HA. But in NEMO-BS instead of sending single BU, each MH in the mobile network sends BU to its corresponding HA through MR when they configure a new address in the mobile network. In the same way, when MR configures a new IPv6 address in a new subnet from a new access router advertised prefixes, it also sends a BU to its HA. When a node configures an address on an interface, immediately Duplicate Address Detection (DAD) method is started to verify the delicacy of that address. After the time-out session of DAD, it calls the SetState() which is defined in ns3 core.

3.2 Packet Processing of NEMO-BS

After the successful completion of Binding Update process, a bidirectional tunnel is established between MR and MR-HA and another tunnel is established between mobile network MN and MNHA (Fig. 2). That is why when a mobile host in the

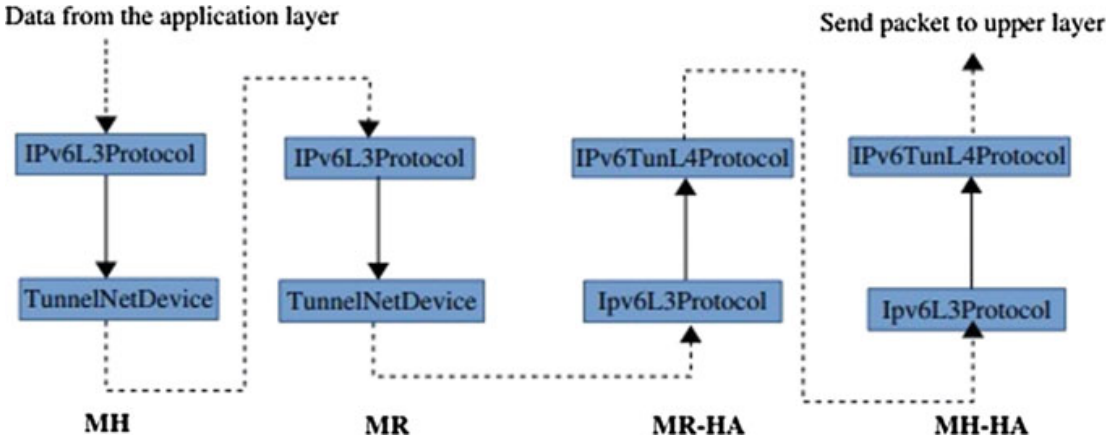


Fig. 2 Data packet processing from MH to MHHA

mobile network sends a data packet to a Correspondent Node (CN), the first level encapsulation is done by the tunnel between in MH and MHHA, and then is sent to MR. After that, second-level encapsulation is performed by MR because of the tunnel between MR and MR-HA. When the two-level encapsulated packet reaches the MR-HA, it decapsulates the outer encapsulation and forwards the remaining packet to the MHHA. When this packet reaches MHHA, it decapsulates the first level of encapsulation and sends it to the right destination that is the correspondent node. Besides this, the data packets are processed in reverse direction from the CN to MH.

4 Class Diagram Implementation

The existing classes of the implemented MIPv6 module [7] are modified to support the NEMO-BS functionalities as shown in the class diagram in Fig. 3. We divide our NEMO-BS classes into mainly four modules:

- Header
- Internet stack
- Net device
- Helper.

The important operations and functionalities of each module are detailed as follows:

4.1 Header

We have implemented the NEMO-BS functionality in ns3, by modifying some of the existing MIPv6 headers message formats in ns3 as defined in RFC3963. All the headers are inherited from the base class Header in ns3 system. In Fig. 3, we have depicted the main data members of NEMO-BS headers and relationship among them as defined in RFC3963. For NEMO-BS in ns3, we have added m flagR data field in Binding Update (BU) and Binding Acknowledgement (BA) header to denote the mobile node's status. If this flag value is set in BU, then HA assumes that MIPv6 mobile node acts as a NEMO-BS MR otherwise it considers it as MIPv6 Mobile Host (MH). A new mobility option header named as Mobile Network Prefix Option Header is added with existing headers shown in Fig. 3, which is also inherited from MIPv6OptionHeader class. It carries the one or more prefix information (128 bit or more) from MR to its HA within BU message and these prefixes are advertised by MR-HA to receive the packet correctly for the mobile host in the mobile network.

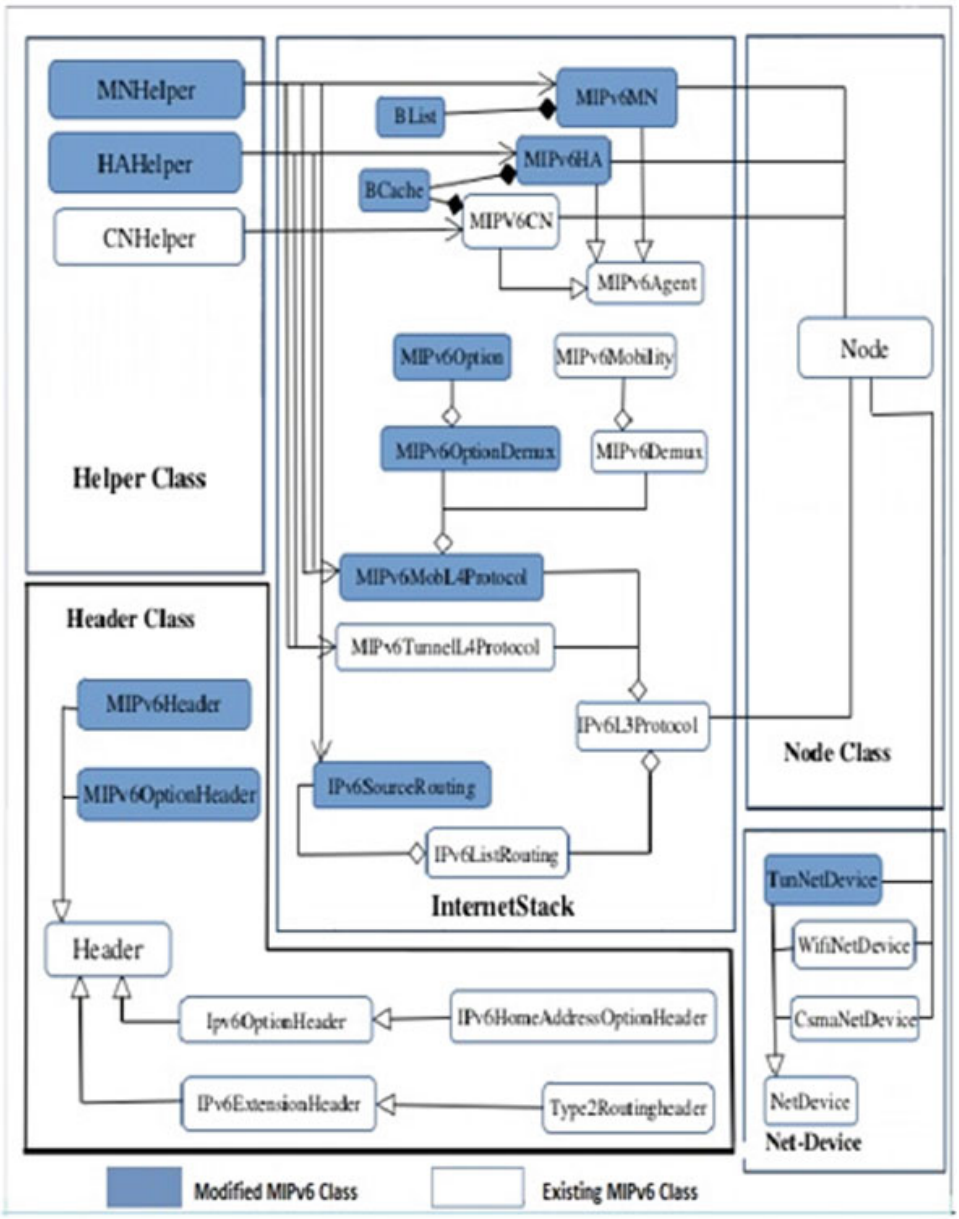


Fig. 3 NEMO-BS classes and their functional relationships

4.2 InternetStack

NEMO-BS implementation contains the two types of Demux classes named as MIPv6Demux and MIPv6OptionDemux. Details of the functionality of Demux classes are described in paper [7]. For NEMO-BS, a new mobility option is added to the MIPv6Option class. The MIPv6MobL4Protocol deals with mobility messages necessary for communication, whereas the MIPv6TunL4Protocol deals with data packets in the time of communication. When from lower layer, a packet is recognized by IPv6L3 Protocol and then sends it to the MIPv6-layer. The summarized data packet or a mobility message contains the MH type value in its header

which is dissimilar from normal data packet. The MH Type field value is checked by IPv6L3Protocol class, and depending on this value, appropriate upper layer L4 protocol class is called. If the packet is received by Receive() function of MIPv6TunL4Protocol class, then its Receive() function checks the summarized packet tunnel line which may use to de-capsulate the packet. It drops the packet if there is no matching tunnel interface found. The MIPv6TunL4Protocol class may be added, removed, or modified a tunnel.

4.3 NetDevice

A tunnel is made for virtual MAC layer by the TunNetDevice class. The TunNetDevice class reimplements the base class Send() function. After the reception of a packet from the upper layer, the Send() function accomplishes IPv6inIPv6 encapsulation by making a new IPv6 header. The main functionality of the MIPv6TunL4Protocol class is TunNetDevice class.

4.4 Helper

This is in the top of the above all implemented classes. When a user wants to use the NEMO-BS protocol, they don't have to bother about the complex internal structure rather they only call the Install() function of MIPv6Helper class to install the functionality on a particular node.

5 Simulation and Results

The simulation framework is shown in Fig. 4. We have used ns-3.25 for simulation. The MR has two interfaces.

- It connects with the Access Router (AR) through Wi-MAX interface and
- With the MNNs through Wi-Fi interface.

Simulation setup: The simulation snapshot is shown in Fig. 5. The Mobile Host (MH) and Correspondent Node (CN) perform as a host and remaining all others as a router. A middle router R1 connects the AR1 and AR2 connects with CSMA interface to MR-HA, another middle router R2 connects the CN and MHHA through CSMA interface to R1. The MR has two interfaces, it connects with access router through Wi-Fi interface, and connects with the mobile host in the mobile network through Wi-Max interface. IEEE 802.11 radio is used over here in AR1 and AR2. Beacons are generated in every 100 ms. Here, the data rate is set at

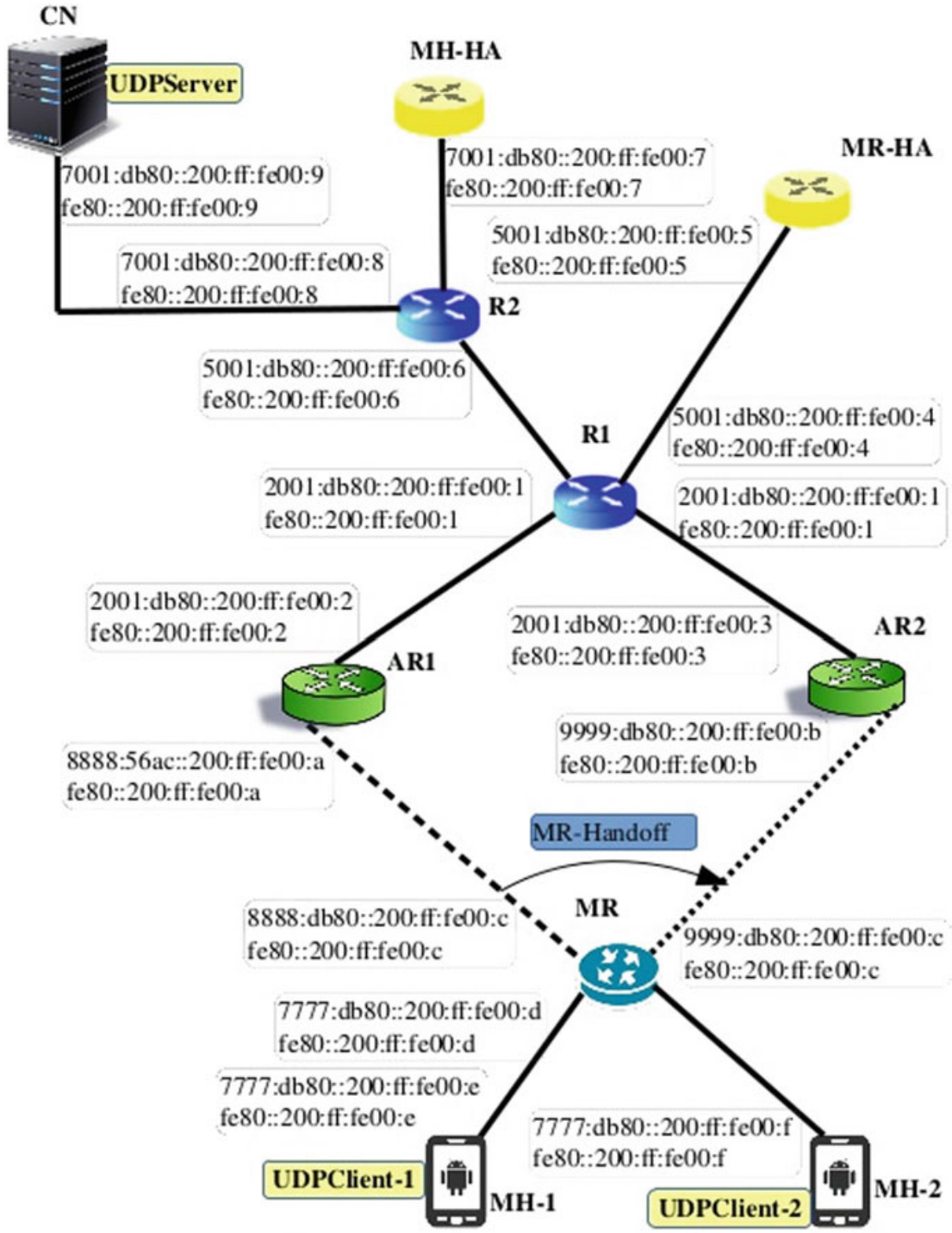


Fig. 4 Simulation Framework

2 Mbps and the link delay at 1.0 ms. On mobile hosts (MH1 and MH2), a UDP echo-server application is installed. On the CN, an echo-client application is running with 1024 packet size and maximum 100,000 packets. Here, the MR home network prefix is 5001:db80:/64, and MR's HoA of MR is expressed from this.

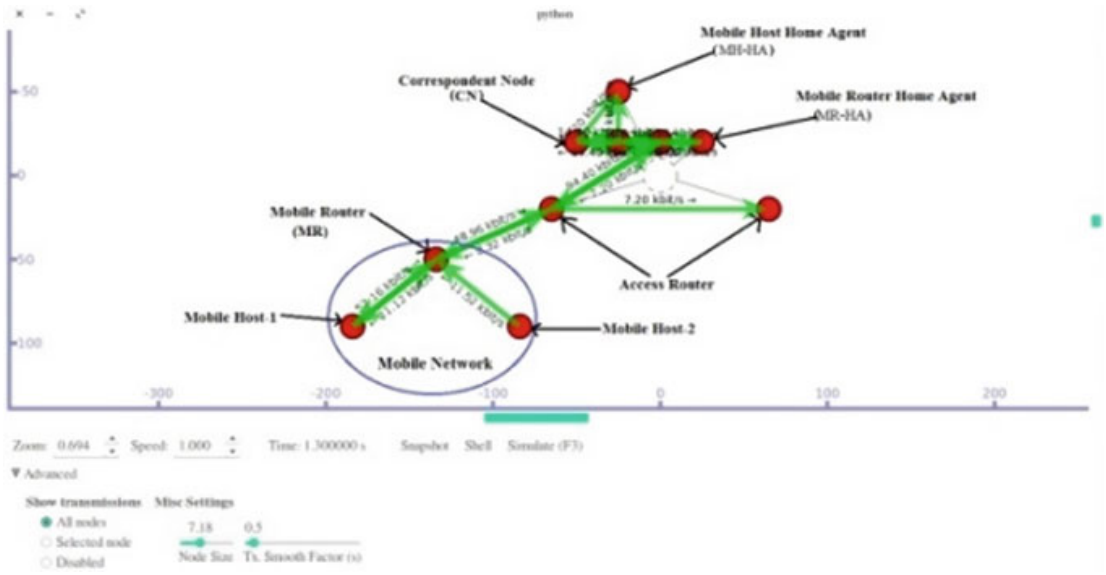


Fig. 5 Simulation snapshot of implemented NEMO-BS in NS3

The constant velocity model (20 m/s) is used for all the MR and moves from left to right fashion.

The simulation parameters are given below:

- Both the AR1 and the AR2 generate beacons in every 100 ms.
- All CSMA links use 2 Mbps data rate and 1.0 ms link delay.
- A UDP echo-server is fixed on the MNNs (MNN is called as MH) (MNN-1 and MNN-2) (say at port 9).
- The MR with all the MNNs in the mobile network uses the constant velocity mobility and travels from the left to right with 20 m/s velocity.

From the packet capture (PCAP) files, we can calculate the delay experienced by MNN1 in the mobile network due to the handoff of MR from AR1 to AR2. At t = 7.240 s, the MNN1 successfully receives the preceding packet from the CN. After that MR changes the point of attachment and after successful completion of the handoff of MR, MNN1 again successfully receives the next packet at t = 12.245 s. So, the delay experienced by MNN1 = (12.245–7.240) s = 5.005 s.

6 Conclusion

NEMO-BS protocol provides seamless Internet connectivity by providing smooth and fast hand-off for a group of users moving from one place to another. In this paper, we have implemented this protocol as an extension of MIPv6 in ns3 environment, so that it can be appended as a new library file in ns-3 in future, which will help the naive users. The above implementation complies with the standard, defined by the IETF. But, a few future works are also required to make it fully usable by the

research community. As the open-source simulator like ns-3 presently mandates validation for any new implemented module, our implementation must be validated against some of the strong existing implementations. Above all, integrating NEMO into LTE would be very interesting as recent research trends mostly rely on LTE communication.

References

1. Kim, M.S., Lee, S.K.: Enhanced network mobility management for vehicular networks. IEEE Trans. Intell. Transp. Syst. **17**(5), 1329–1340 (2016)
2. Perera, E., Sivaraman, V., Seneviratne, A.: Survey on network mobility support. SIGMOBILE Mob. Comput. Commun. Rev. **8**(2), 7–19 (2004)
3. Network Mobility Basic Support Protocol. Retrieved July 25, 2018 from https://github.com/prasanta2018/NetworkMobility
4. Bernardos, C.J., De La Oliva, A., Caldern, M., von Hugo, D., Kahle, H.: Nemo: network mobility. Bringing ubiquity to the internet access. Demonstration at IEEE INFOCOM (2006)
5. Khan, M.Q., Andresen, S.H., Khan, K.N.: Pros and cons of route optimization schemes for network mobility (NEMO) and their effects on handovers. In: Proceedings of the 8th International Conference on Frontiers of Information Technology (FIT '10). ACM, New York, NY, USA, Article 24
6. Bauer, C., Ayaz, S., Ehammer, M., Grupl, T., Arnal, F.: Infrastructure-based route optimization for NEMO based on combined local and global mobility. In: Mobility Conference, p. 61 (2008)
7. Rana, M.K., Rana, B., Mandal, S., Saha, D.: Implementation and performance evaluation of a mobile IPv6(MIPv6) simulation model for ns-3. Simul. Model. Pract. Theory **72**, 122 (2017)
8. Lee, J.H., Ernst, T., Chilamkurti, N.: Performance analysis of PMIPv6Based network mobility for intelligent transportation systems. IEEE Trans. Veh. Technol. **61**(1), 74–85 (2012)
9. Hager, R., Klemets, A., Maguire, G.Q., Reichert, F., Smith, M.T.: MINT-AMobile internet router. In: 1st International Symposium on Global Data Networking, Cairo, Egypt, Dec. 13–15, 1993 (1993)
10. Perkins, C.: IP Mobility Support, IETF. RFC October 1996 (2002)

Modified DFA Minimization with Artificial Bee Colony Optimization in Vehicular Routing Problem with Time Windows

G. Niranjani and K. Umamaheswari

Abstract A NP-hard problem, vehicular routing is a combinatorial optimization problem. Vehicular routing problem with time windows indicates vehicular routing with specified start and end time. There will be "n" number of vehicles starting from the depot to cater to the needs of "m" customers. In this paper, Gehring and Homberger benchmark problems are considered wherein the size of customers is taken to be 1000. Artificial Bee Colony Optimization algorithm is executed on these 60 datasets and the number of vehicles along with total distance covered is recorded. The modified version of Deterministic Finite Automata is applied along with the Artificial Bee Colony Optimization and the results produce 25.55% efficient routes and 15.42% efficient distance compared to simple Artificial Bee Colony Optimization algorithm.

Keywords Artificial Bee Colony Optimization · Minimization · Number of routes · Total distance

1 Introduction

Vehicular routing problem with time windows is a NP-hard problem. It is a combinatorial optimization mechanism. Vehicular routing problem is similar to travelling salesman problem. In the travelling salesman problem, a sales person has to start from one city, travel to all the other cities only once and then return to the original city with the minimum cost possible. In VRPTW, the number of customers and the number of vehicles are specified. These vehicles start from the depot that is also specified in the problem, go to different customers that are distributed across

G. Niranjani (✉)
Department of CSE, PSG Institute of Technology and Applied Research, Coimbatore, India
e-mail: niragopal@gmail.com

K. Umamaheswari
Department of IT, PSG College of Technology, Coimbatore, India
e-mail: umakpg@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_45

and then return to the depot. The customers are specified as x and y coordinates of the point. When a customer is reached, the next customer is selected which has the minimum distance from the current point among all the points that are not yet visited. This paper considers the 60 datasets pertaining to Gehring and Homberger benchmark problems that service 1000 customers.

In the real world, these vehicle routing problems are being used in distribution and transportation logistics [1] and also specific applications such as solid waste, beverage, food, diary and newspaper industries [2].

The dataset, in general, contains the following details: vehicle count and the capacity of each vehicle. For each vehicle, the following details are specified: (1) Customer number: for depot, the customer number is zero. All the other vehicles have natural numbers sequentially. (2) x-coordinate of the customer. (3) y-coordinate of the customer. (4) Demand: this denotes the priority. (5) Start time: the customer needs to be serviced after the start time. (6) End time: the customer needs to be serviced before the end time. (7) Service time: the time taken to perform service for this particular customer. The Gehring and Homberger benchmark problems have six sets of problems: C1 type, C2 type, R1 type, R2 type, RC1 type and RC2 type. Each type has 10 problems. The types vary in terms of vehicle count and the capacity of each vehicle.

1.1 Reason for Considering the Gehring and Homberger's Problem for the Experiments

1. Solomon's benchmark problems have been considered in many algorithms associated with vehicular routing problem with time windows. In these problems, the maximum number of customers in a single instance is 100. Gehring and Homberger benchmark problems are a variation of Solomon's benchmark problems with more number of customers for each instance.
2. The number of customers represented in each of the instances is 1000. To demonstrate that the algorithm can handle a huge amount of customers comparative to the other algorithms such as branch and price, Tabu search, firefly algorithms and even some variations of ABC [Tournament selection, Vector Evaluated Technique and Roulette Wheel Selection technique], these problems are selected.

2 Artificial Bee Colony Optimization

Karaboga [3] started the Artificial Bee Colony algorithms after observing the behaviour of honey bees and applied them to real-time problems. He along with Karaboga and Basturk [4] further extended the usage of the algorithm for solving

various optimization problems. Karaboga and Basturk [4] and Singh [5] propose a solution that shows the better performance of bee colony algorithms in comparison with particle swam optimization and genetic algorithms, particle swarm optimization [6, 7], Tabu search and ACO algorithms, respectively.

2.1 Working of Honey Bees

Artificial Bee Colony Optimization algorithms are based on the working of real-world honey bees. These algorithms come under swarm intelligence—wherein problems are solved based on the collective behaviour of social insect colonies and other animals. In this paper, the honey bees are considered. There are three types of honey bees: (1) Scout bees: these are used to search for new food sources randomly. (2) Onlookers: the onlooker bees calculate the nectar amounts available in the food sources. (3) Employed bees: these determine the nectar amount and the probability value with which the food source can be reached from the bee hive.

2.2 Essential Components of Forage Selection

(1) Food sources: a particular food source is selected based on its proximity to the hive, richness/concentration and the ease of extracting energy. (2) Employed forager: it carries information about the food source, how much distance it is from the hive and the direction of the food source from the hive, profitability and it shares information with a certain probability. (3) Unemployed forager: the scout bees and the onlooker bees come under this category.

2.3 Modes of Behaviour

There are two important modes of behaviour of the honey bees with respect to the food sources. (1) Recruitment of the food source, (2) Leaving behind the food source. The honey bees exchange information about the food sources among themselves by means of the “waggle dance”.

2.4 Basic Self-Organization Properties

(1) Positive feedback: when the nectar in the food source is high, the number of onlooker bees is high. (2) Negative feedback: when the nectar in the food source is poor, there are no bees nearer. (3) Fluctuations: probability value of the nectar in the

food source in comparison with others [8]. (4) Multiple Interactions: Communication between different types of bees with regard to the food source. Many tasks are performed at the same time leading to the division of labour [3].

2.5 Algorithm

The general Artificial Bee Colony Optimization algorithm proposed by Dervis Karaboga [4] is as follows:

Send the scout bees onto the initial food sources

REPEAT

Send the employed bees and find the nectar amounts in the food sources
The onlooker bees decide the selection of food source based on its probability value
They go to the respective food source and determine the nectar amount
The food source selected by the bees are exploited
Again, send the scout bees to search for new food sources
Remember the food sources exploited so far.

UNTIL (constraints are satisfied).

2.6 Parameter Setting for the ABC Algorithm

1. All the vehicles start from the depot.
2. It is assumed that the vehicles are not affected by traffic conditions.
3. All the customers will be serviced only after the start time but before their end time.
4. A particular vehicle should be available to service a customer within the service time of that particular customer.

3 VRPTW

The study of vehicular routing started in 1959 when G. B. Dantzig and J. H. Ramser [9] proposed the truck dispatching problem applicable to the delivery of gasoline to service stations. It was followed by Clarke [10] to provide an iterative procedure to select the optimum or near-optimum route in a better way. A lot of algorithms [11] have been proposed and some implemented to this problem which provide both exact and approximate solutions.

The vehicular routing problem with time windows has been solved by many algorithms. Some of the well known algorithms are presented in Table 1.

3.1 Objectives

The following are the objectives of the VRPTW:

1. Minimization of Total routes:
   The total number of routes refers to the total number of vehicles needed to solve a particular problem. The number of vehicles varies depending on the placement of customers. If the customers are nearer, a single vehicle is enough to service all those customers, provided that particular vehicle is available within the start time and the end time of that customer. The algorithm is designed such that the total number of routes is minimized [12].
2. Minimization of the total distance travelled:
   Distance of a single vehicle is measured as the sum of the distance between the depot and the first customer that the vehicle is servicing and the distance between the first and the second customer and then the second and the third and so on till the last customer. The distance between the last customer and the depot is added along with these values. The total distance of a particular customer is calculated as the summation of distances of individual vehicles. This algorithm takes care of the fact that the total distance is reduced.
3. Minimization of the total time taken:
   Time taken for a single vehicle is calculated as the summation of the start time of the first customer and its service time and then subsequent service time of the corresponding customers. Total time is calculated as the summation of the time taken by individual vehicles which is to be reduced by the proposed algorithms.

   It has to be taken care that the capacity constraints are not violated [13].

3.2 Procedure

There are different variations for solving VRPTW using Artificial Bee Colony Optimization algorithm. Some are presented in Table 2.

The vehicle count, the capacity of vehicles and the depot details—customer number: 0, x-coordinate and y-coordinate of the depot, demand, start time, end time and unload time—are got from the input dataset along with the details of the 1000 customers. The number of routes is initialized to 0.

The following procedures are executed until all the customers are serviced:

1. MemorizeBestSource (): This function is used to select the customer with the least start time and also that has not been serviced yet.

Table 1 Existing solutions for VRPTW

Author and year	Title	Strategy	Parameter	Dataset	Methodology	Benefits	Limitations
Bulhões and Uchoa [16]	A branch-and-price algorithm for the Minimum Latency Problem	Minimum Latency Problem	Minimize the sum of waiting time of customers	40 TSP Library instances with up to 200 vertices	In branch-and-price method, the arcs are considered instead of specific nodes	Roberti and Mingozzi (2014) did not solve nine instances. This method solved them. Also, branch-and-price could solve larger instances of size 159, 175, 180 and 195	When the vertices are greater than 150, it is difficult to solve
Michael Schneider, September 2015 [17]	The vehicle-routing problem with time windows and driver-specific times	Tabu search	Minimize travelling distance and working duration	Solomon's 100 customer instances	Relocation, Exchange of data and 2-Opt methods	10 runs are executed and the optimal ones are considered	Meta-heuristic method can be improved by means of splitting methods for solutions of higher quality
Huang et al. [18]	Multi-agent ACO for VRP with soft time windows and road condition	A mathematical model of VRP with soft time windows [penalty for lateness] and road factor	Transport cost, fuel consumption and customer satisfaction achieved by the arrival time of vehicles and the expecting time of customers	40 customer problem	Adaptive pheromone trail, state transition rule and pheromone updating, 3-opt solution	Multiagent ant colony optimization is considered better than ACO and AGA	For larger instances, yet to be verified

(continued)

Table 1 (continued)

Author and year	Title	Strategy	Parameter	Dataset	Methodology	Benefits	Limitations
Matthopoulos and Sofianopoulou [19]	A firefly Algorithm for the Heterogeneous Fixed Fleet VRP	HFVRP is a variation of VRPTW wherein the customers have different capacities and costs for goods distribution	Relative brightness, Degrees of attraction, Distance between fireflies	Augerat et al. (1995)	Relative brightness, Degrees of attraction, Distance between fireflies, Position update and Parameters values selection	Produces the best known solution for small datasets—up to 39 nodes	The next node is determined by means of a random walk of the firefly and hence, it is not much effective

Table 2 Existing solutions for ABC in VRPTW

Author and year	Title	Strategy	Parameter	Dataset	Methodology	Benefits	Limitations
Shi et al. [20]	A modified ABC algorithm for vehicle routing problems with time windows	Tournament selection	Minimize the total distance	Solomon's R102 problem	Fitness points are calculated between pairs of points and at the end, the one with the larger fitness value is selected	Travelled distance is better than the ant colony algorithm, iterated local search and genetic algorithms	Only one instance of 100 customer problem was considered
Nahum et al. [21]	Multi-Objective VRPTW: a Vector-Evaluated ABC Approach	Vector-Evaluated Technique	Count of vehicles and the total cost	Solomon's 25, 50 and 100 customer problem	(1) Minimize the number of vehicles, (2) The total path length is optimized based on the results calculated beforehand	Optimal solution can be obtained by running the algorithm again and again	The end result is based on previous runs. So, it takes lots of time for computation
Bhagade and Puranik [12]	Artificial Bee Colony (ABC) Algorithm for Vehicle Routing Optimization Problem	Effective path planning	Tour length and bee travel time	Results are simulated	Control parameters are set which help in determining the next node using the nearest neighbour method	Optimal distance is obtained by greedy selection strategy	Very few parameters are considered
Alzaqebah et al. [22]	Modified Artificial Bee Colony for the vehicle routing problems with time windows	Roulette wheel selection	Convergence speed increased	Solomon's 100 customer problem	The bees select the best node based on the random selection of nodes for the count of maximum trials	67% of the results are better than simple- ABC	The results need to be increased

2. SendEmployedBees (): This function is used to execute the work of the employed bees. The customer selected by the MemorizeBestSource () is to be processed. This function selects the route that this customer is to be placed in. A group of routes is selected based on the following conditions:

   - The distance from the depot should be greater than the distance from the routes latest point to this customer.
   - The time taken by the route until the latest point should be less than the customer's end time.

3. CalculateProbabilities (): This function selects the route with the minimum distance among the set of routes selected in the SendEmployedBees () function.
4. SendOnlookerBees(): If no such routes are selected by the CalculateProbabilities () function, a new route is created with the following conditions:

   - The vehicle count of the new route is set to 1.
   - The customer is marked as serviced.
   - A new route is created with the present customer as the first vehicle.
   - Route time is equal to the sum of the starting time of a particular customer and the time taken to unload.
   - Distance of the route is calculated as the distance from the depot to the present customer.
   - The number of routes is incremented by 1.

When a sample route is selected to fit in this customer, the following procedures are carried out:

- The customer is marked as serviced.
- The present customer is added with the selected route as the last one.
- The number of vehicles processed by the current route is incremented by 1.
- If the start time of the present customer is before the route time, the route time is calculated as the sum of route time and the unload time for this particular customer. Else, the route time is calculated as the sum of the starting time of the present customer with its unload time.
- Distance of the route is calculated as the sum of the previous distance and the distance between the previous point and the present customer.
- The present customer is made as that particular route's last point.

All the distances are measured using the Euclidean distance—The distance between two points (x_1, y_1) and (x_2, y_2) is

$$\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

5. SendScoutBees (): The scout bees manage the customers that can be selected in the next iteration.
   The route number, the number of vehicles, specific distance, number of customers serviced and time taken is displayed for each route. In the end, the total

number of routes, total distance—the summation of the distance of individual routes and the total number of serviced customers—are displayed. The number of vehicles and the total distance travelled is recorded to be compared.

4 Minimization of Deterministic Finite Automata

Deterministic Finite Automata refers to a machine that has a countable number of states and each input state has exactly one output state on a given input. The deterministic finite automata is indicated as $(Q, \sum, \delta, q0, F)$ wherein Q refers to the set of states, $\sum$ refers to the input alphabet, δ refers to the transition equation which contains the input state, input alphabet and the output state, $q0$ refers to the initial state and F refers to the set of final states [14].

For minimizing a DFA, initially, the final states are considered as one component and all the other states are considered as another component. Each component is checked whether they all give the states belonging to the same component for each alphabet. If so, they are retained in the same component. Else, the component is split and the same procedure is followed again until there is no change. In the end, the states belonging to the same component are considered as one state and the Deterministic Finite Automata is reconstructed [15].

4.1 Algorithm for Modified DFA on ABC

On applying modified DFA minimization, the ABC algorithm changes as follows:

Send the scout bees onto the initial food sources
Group the food sources based on the distance between them
REPEAT

Send the employed bees and find the nectar amounts in the food sources
The onlooker bees decide the selection of food source based on their probability value
They go to the respective food source and determine the nectar amount
The food source selected by bees are exploited.

For each food source that is available in the group of the current food source

If the food source is currently available and is not yet exploited, it is exploited by the bees

Again, send the scout bees to search for new food sources
Remember the food sources exploited so far

UNTIL (constraints are satisfied).

4.2 Parameter Setting for Modified DFA on ABC

In addition to the above-mentioned parameters, a few are added to the modified DFA with the ABC algorithm:

1. The customers in a single group will be serviced by a single vehicle only when the start time and the end time coincide with the free time of the vehicle. Otherwise, different vehicles will be assigned to vehicles in a single group.

5 Modified DFA Minimization with Artificial Bee Colony Optimization in Vehicular Routing Problem with Time Windows

5.1 Changes to Artificial Bee Colony Optimization in VRPTW

In normal DFA minimization, the states that are grouped as a single component are considered as a single state. Here, it is modified such that the customers are grouped such that they belong to the same route if the capacity does not exceed. If the capacity exceeds the specified limit, the group can be split into different customers belonging to different routes.

The customers are considered in pairs having a minimum distance between them and they are grouped based on the following conditions:

- When both the customers are not in any groups, and they have a minimum distance between themselves compared to others; those two customers are combined together to form a new group.
- When any one customer of the pair is grouped and the other one is not paired, the customer not paired belongs to the group of the paired customer.
- When both the customers are paired, then two conditions are checked: (1) When they both belong to the same group, they are left as such. (2) When they both belong to different groups, any one of the customer groups is discarded and the customers belonging to the discarded group is added to the group of the other customer.

The functions of the Artificial Bee Colony Optimization for vehicular routing problem are executed as such except the functions:

- SendOnlookerBees (): In this function, a small change is done. When a customer belonging to a group is selected to be attached to a route, the customers belonging to the same group are given a preference to be selected in the same route, provided the constraint pertaining to the capacity of the vehicle is not violated.

- SendScoutBees (): This function is used to service the customers belonging to the group of the last serviced customer whose start time is less than the route time of the present route.

The route number, the number of vehicles, specific distance, number of customers serviced and time taken are again displayed for each route. In the end, the total number of routes, total distance—the summation of the distance of individual routes, the total number of serviced customers—are displayed. The number of routes and the total distance travelled is recorded to be compared with the previous one.

6 Results and Discussion

6.1 Comparison of Routes

The routes obtained from the Artificial Bee Colony algorithm is compared with the routes obtained in the Minimized DFA combined with ABC algorithm for the VRPTW. The comparison is done for each problem type separately.

C1_Type: Ten problems with count of vehicles: 200 and capacity: 250. There are 1000 customers in these problems each with service time of 90. The modified DFA with ABC algorithm gave better results for 90% of problems than the ABC algorithm as indicated in Fig. 1. The number of routes using minimized DFA is found to be 7.10% better than ordinary Artificial Bee Colony Optimization algorithm.

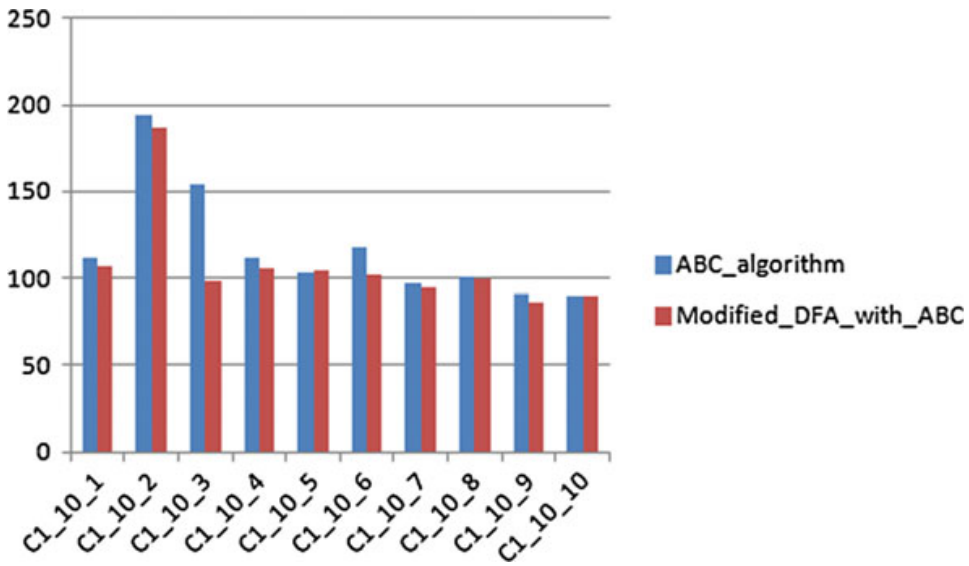


Fig. 1 Comparison of the number of routes for the ABC algorithm with the Modified DFA with ABC algorithm for C1 type problem

C2_Type: Ten problems with count of vehicles: 700 and capacity: 250. There are 1000 customers in these problems each with service time of 90. The modified DFA with ABC algorithm gave better results for 80% problems than the ABC algorithm as indicated in Fig. 2. The efficiency when minimized DFA is applied improves by 16.8% in terms of the number of routes.

R1_Type: Ten problems with count of vehicles: 200 and capacity: 250. There are 1000 customers in these problems each with service time of 10. The modified DFA with ABC algorithm gave better results for 90% problems than the ABC algorithm as indicated in Fig. 3. The number of routes is found to be 24.38% more efficient when minimized DFA is applied.

R2_Type: Ten problems with count of vehicles: 1000 and capacity: 250. There are 1000 customers in these problems each with service time of 10. The modified DFA with ABC algorithm gave better results for 100% problems than the ABC algorithm as indicated in Fig. 4. When minimized DFA is applied over Artificial Bee Colony Optimization, the results obtained are found to be 40.59% more efficient in terms of the number of routes.

RC1_Type: Ten problems with count of vehicles: 200 and capacity: 250. There are 1000 customers in these problems each with service time of 10. The starting time of customers in this data-set is far less compared to the previous problems. The modified DFA with ABC algorithm gave better results for 90% problems than the ABC algorithm as indicated in Fig. 5. These instances are found to be 10.43% more efficient when minimized DFA is applied for calculating the number of routes.

RC2_Type: Ten problems with count of vehicles: 1000 and capacity: 250. There are 1000 customers in these problems each with service time of 10. The starting time of customers in this dataset is far less compared to the R1, R2, C1 and C2 type problems. The modified DFA with ABC algorithm gave better results for 100%

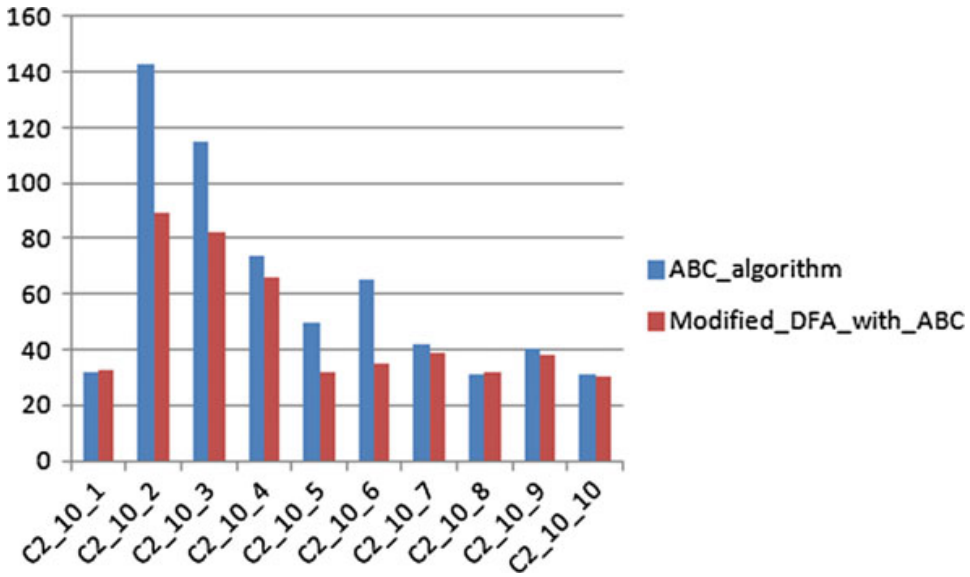


Fig. 2 Comparison of the number of routes for the ABC algorithm with the Modified DFA with ABC algorithm for C2 type problem

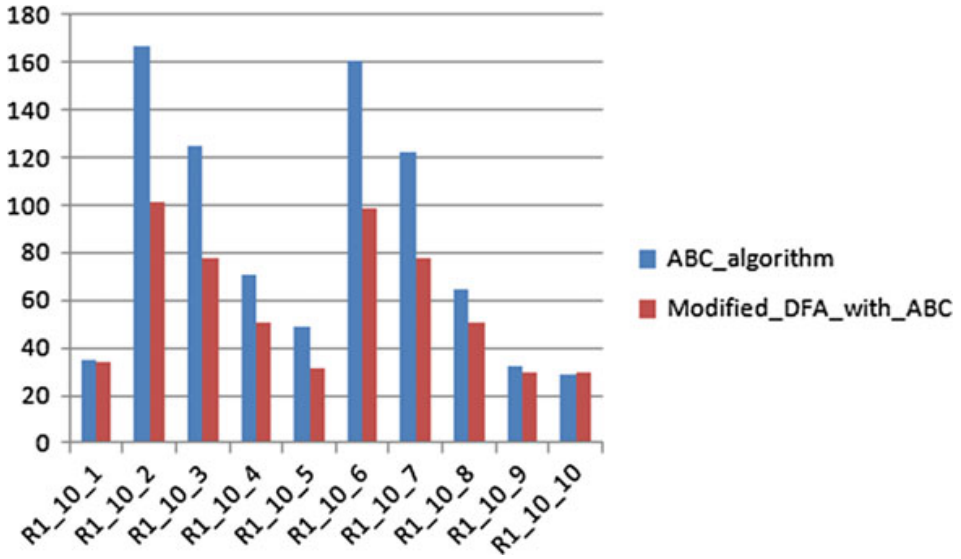


Fig. 3 Comparison of the number of routes for the ABC algorithm with the Modified DFA with ABC algorithm for R1 type problem

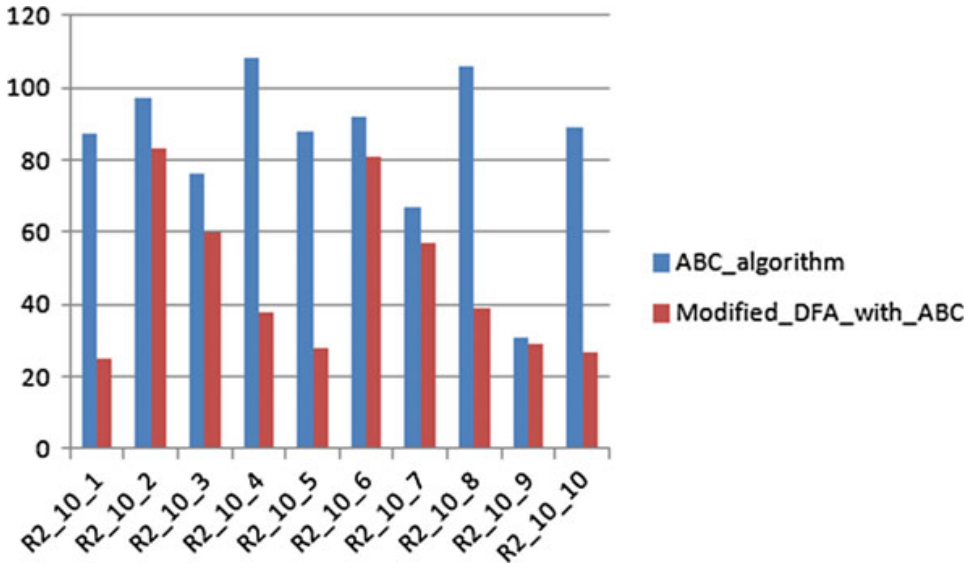


Fig. 4 Comparison of the number of routes for the ABC algorithm with the Modified DFA with ABC algorithm for R2 type problem

problems than the ABC algorithm as indicated in Fig. 6. 53.97% more efficiency is achieved for the RC2 type problems when minimized DFA is applied compared to Artificial Bee Colony Optimization algorithms for finding the number of routes.

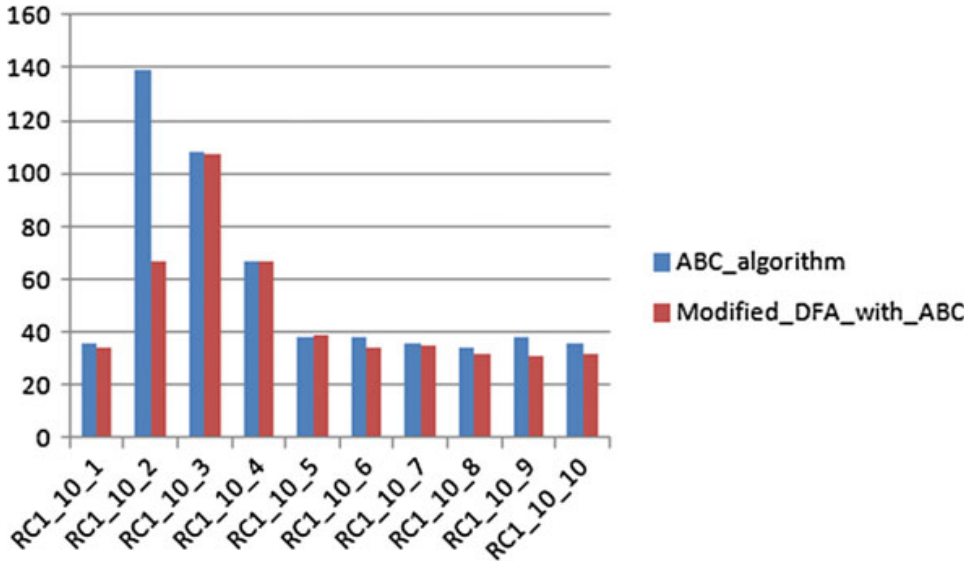


Fig. 5 Comparison of the number of routes for the ABC algorithm with the Modified DFA with ABC algorithm for RC1 type problem

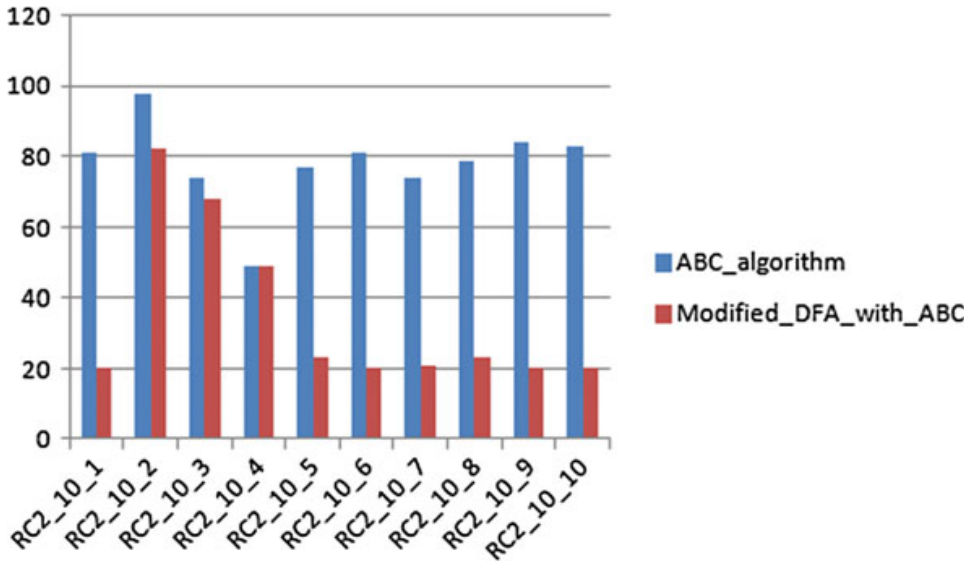


Fig. 6 Comparison of the number of routes for the ABC algorithm with the Modified DFA with ABC algorithm for RC2 type problem

6.2 Comparison of Distances

C1_Type: Ten problems with count of vehicles: 200 and capacity: 250. There are 1000 customers in these problems each with service time of 90. The modified DFA with ABC algorithm gave better results for 100% problems than the ABC algorithm as indicated in Fig. 7. The instances are found to be 7.29% more efficient when minimized DFA is applied on top of Artificial Bee Colony Optimization algorithms.

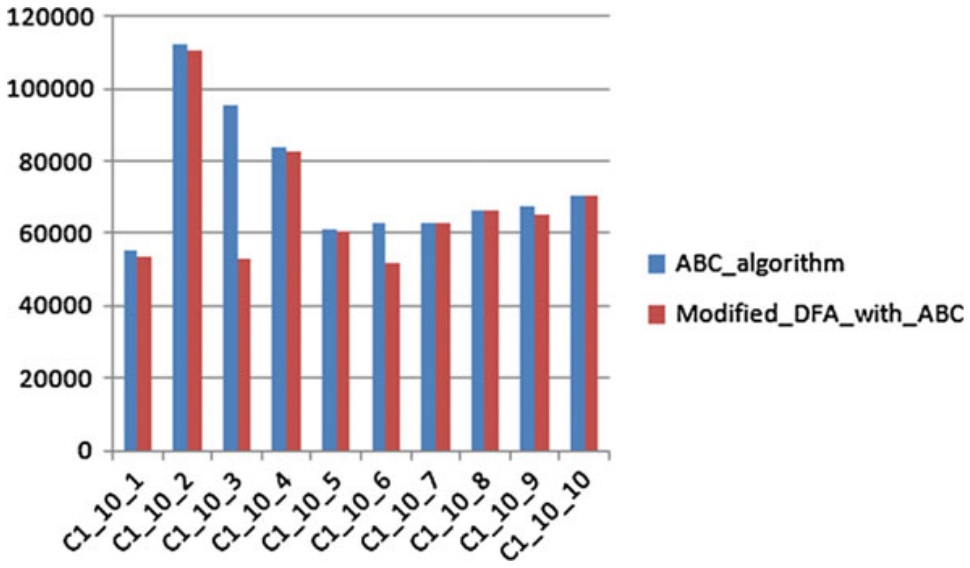


Fig. 7 Comparison of the total distance for the ABC algorithm with the Modified DFA with ABC algorithm for C1 type problem

C2_Type: Ten problems with count of vehicles: 700 and capacity: 250. There are 1000 customers in these problems each with service time of 90. The modified DFA with ABC algorithm gave better results for 90% problems than the ABC algorithm as indicated in Fig. 8. The efficiency when minimized DFA is applied improves by 18.02% in terms of the total distance.

R1_Type: Ten problems with count of vehicles: 200 and capacity: 250. There are 1000 customers in these problems each with service time of 10. The modified DFA with ABC algorithm gave better results for 80% problems than the ABC

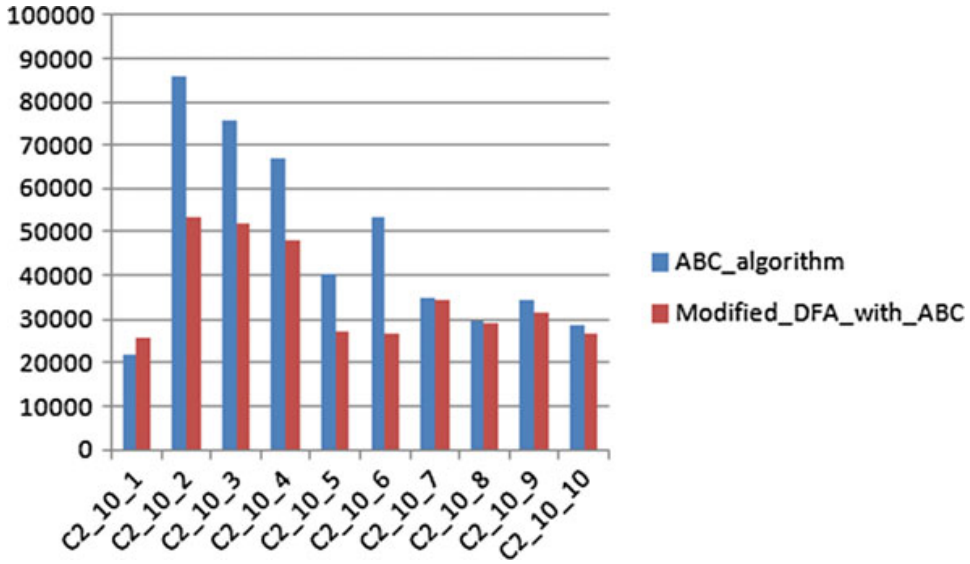


Fig. 8 Comparison of the total distance for the ABC algorithm with the Modified DFA with ABC algorithm for C2 type problem

algorithm as indicated in Fig. 9. The total distance is found to be 8.44% more efficient when minimized DFA is applied.

R2_Type: Ten problems with count of vehicles: 1000 and capacity: 250. There are 1000 customers in these problems each with service time of 10. The modified DFA with ABC algorithm gave better results for 100% problems than the ABC algorithm as indicated in Fig. 10. When minimized DFA is applied over Artificial Bee Colony Optimization, the results obtained are found to be 31.72% more efficient in terms of the total distance.

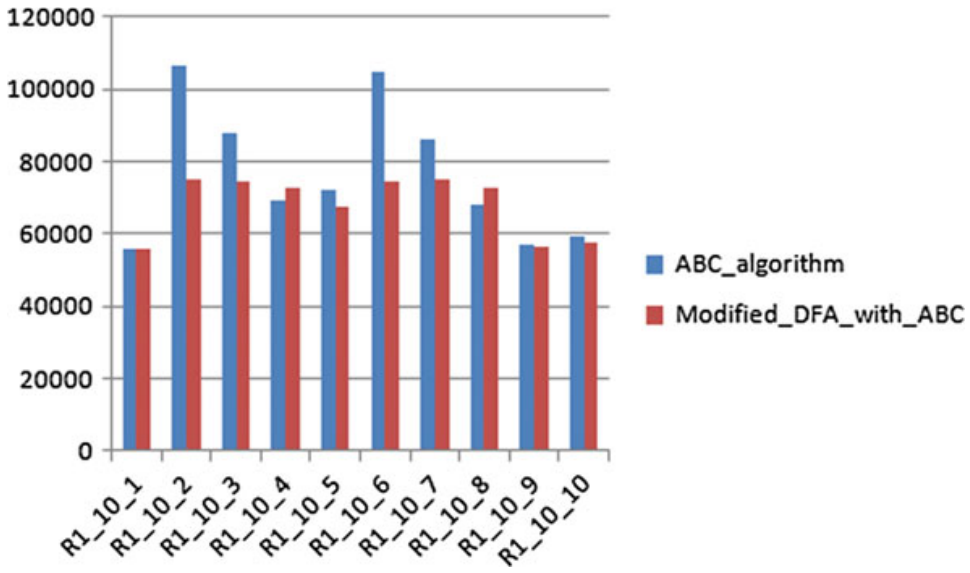


Fig. 9 Comparison of the total distance for the ABC algorithm with the Modified DFA with ABC algorithm for R1 type problem

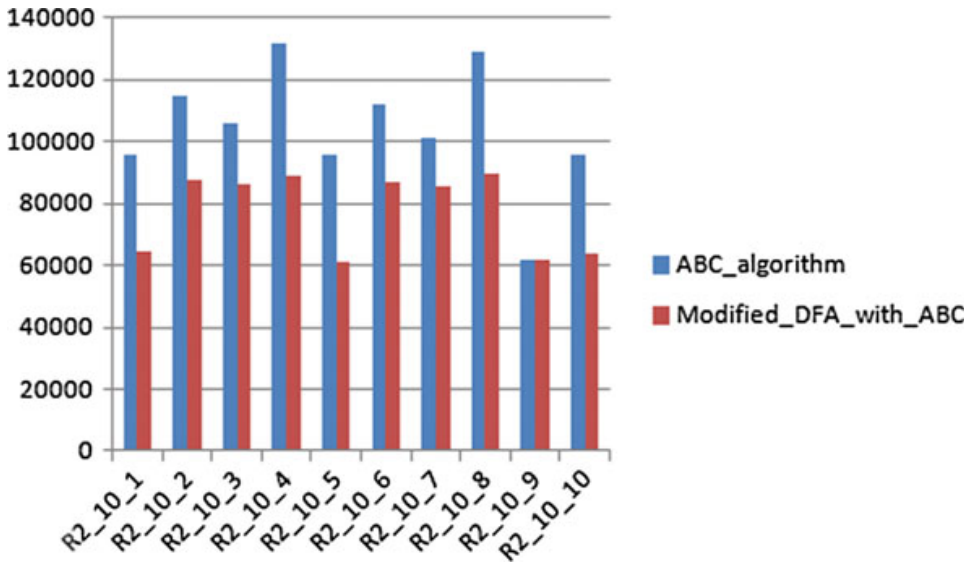


Fig. 10 Comparison of the total distance for the ABC algorithm with the Modified DFA with ABC algorithm for R2 type problem

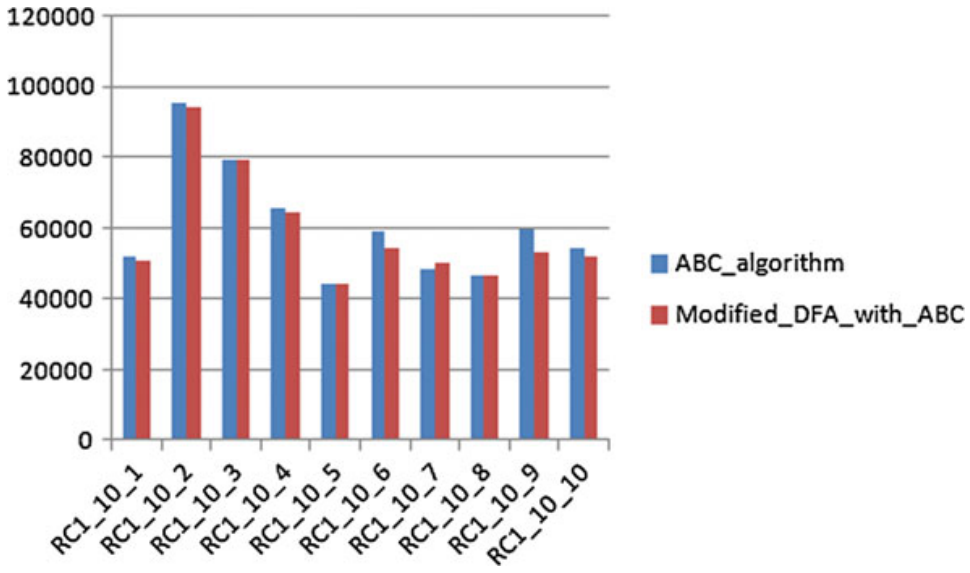


Fig. 11 Comparison of the total distance for the ABC algorithm with the Modified DFA with ABC algorithm for RC1 type problem

RC1_Type: Ten problems with count of vehicles: 200 and capacity: 250. There are 1000 customers in these problems each with service time of 10. The starting time of customers in this dataset is far less compared to the previous problems. The modified DFA with ABC algorithm gave better results for 90% problems than the ABC algorithm as indicated in Fig. 11. These instances are found to be 2.47% more efficient when minimized DFA is applied for calculating the total distance.

RC2_Type: Ten problems with count of vehicles: 1000 and capacity: 250. There are 1000 customers in these problems each with service time of 10. The starting

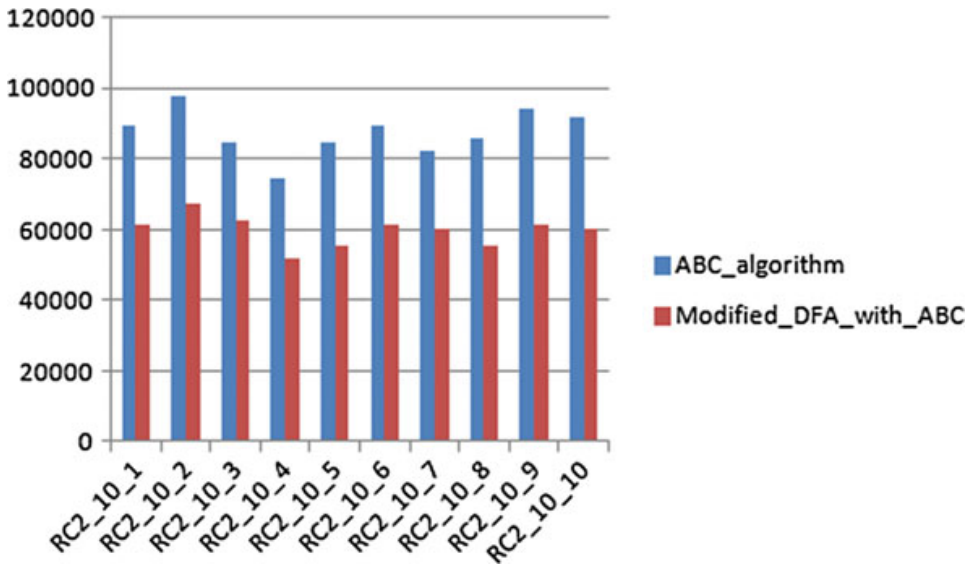


Fig. 12 Comparison of the total distance for the ABC algorithm with the Modified DFA with ABC algorithm for RC2 type problem

time of customers in this dataset is far less compared to the R1, R2, C1 and C2 type problems. The modified DFA with ABC algorithm gave better results for 100% problems than the ABC algorithm as indicated in Fig. 12. 53.97% more efficiency is achieved for the RC2 type problems when minimized DFA is applied compared to Artificial Bee Colony Optimization algorithms for finding the number of routes.

7 Conclusion and Future Work

The Modified DFA with Artificial Bee Colony algorithm has a higher performance when compared to the Artificial Bee Colony Optimization algorithm when applied to vehicular routing problem with time windows. Overall performance: 91.67% problems got better results, i.e. 25.55% more efficient in terms of the number of routes and 93.33% of the problems got better results, i.e. 15.42% more efficient for the total distance travelled by all the vehicles as a whole for the vehicular routing problem with time windows. The total time taken can also be added as a parameter along with these values. Even other algorithms such as genetic algorithms, particle swarm optimization and ant colony optimization can be optimized by applying DFA minimization along with those concepts for the VRPTW.

References

1. Toth, P., Vigo, D.: The vehicle routing problem. In: Toth, P., Vigo, D. (eds.) SIAM Monographs on Discrete Mathematics and Applications. SIAM, Philadelphia (2001)
2. Golden, B.L., Assad, A.A., Wasil, E.A.: Routing vehicles in the real world: applications in the solid waste, beverage, food, dairy, and newspaper industries. In: Toth, P., Vigo, D. (eds.) The Vehicle Routing Problem, pp. 245–286. SIAM, Philadelphia (2002)
3. Karaboga, D.: An idea based honey bee swarm for numerical optimization. Technical Report-TR06, October (2005)
4. Karaboga, D., Basturk, B.: A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. J. Global Optim. **39**(3), 459–471 (2007)
5. Singh, A.: An artificial bee colony algorithm for the leaf-constrained minimum spanning tree problem. Appl. Soft Comput. **9**(2), 625–631 (2009)
6. Umamaheswari, K., Sarathambekai, S.: Task scheduling in distributed systems using discrete particle swarm optimization. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **4**(2):510–522 (2014/2)
7. Sarathambekai, S., Umamaheswari, K.: Task scheduling in distributed systems using heap intelligent discrete particle swarm optimization. J. Comput. Intell. **33**(4):737–770 (2017/11)
8. Bonabeau, E., Dorigo, M., Theraulaz, G.: Swarm Intelligence: From Natural to Artificial Systems. Oxford University Press, New York (1999)
9. Dantzig, G.B., Ramser, J.H.: The truck dispatching problem. Manag. Sci. **6**(1), 80–91 (1959)
10. Clarke, G., Wright, J.W.: Scheduling of vehicles from a central depot to a number of delivery points. J. Oper. Res. Soc. Am. **12**(4), 568–581 (1964)
11. Desrochers, M., Desrosiers, J., Solomon, M.: A new optimization algorithm for the vehicle routing problem with time windows. J. Oper. Res. Soc. Am. **40**(2), 342–354 (1992). https://doi.org/10.1287/opre.40.2.342

12. Bhagade, A.S., Puranik, P.V.: Artificial bee colony (abc) algorithm for vehicle routing optimization problem. Int. J. Soft Comput. Eng. (IJSCE) **2**(2), 2231–2307 (2012)
13. https://www.sintef.no/projectweb/top/vrptw/solomon-benchmark/
14. Watson, B.W.: A taxonomy of finite automata minimization algorithms. Computing Science Report 93/44 Eindhoven, January 1995 (1995)
15. Perrin, D.: CHAPTER 1—Finite Automata Formal Models and Semantics, Handbook of Theoretical Computer Science, pp. 1–57. Elsevier, Amsterdam (1990)
16. Bulhões, S.R., Uchoa, E.: A branch-and-price algorithm for the minimum latency problem. Comput. Oper. Res. **93**, 66–78 (2017)
17. Schneider, M.: The vehicle-routing problem with time windows and driver-specific times. Eur. J. Oper. Res. **250**(1), 101–119 (2016). https://doi.org/10.1016/j.ejor.2015.09.015
18. Huang, G., Cai, Y., Cai, H: Multi-agent ant colony optimization for vehicle routing problem with soft time windows and road conditions. In: MATEC Web Conference, Volume 173, 2018 International Conference on Smart Materials, Intelligent Manufacturing and Automation [SMIMA] (2018). https://doi.org/10.1051/matecconf/201817302020
19. Matthopoulos, P.-P., Sofianopoulou, S.: A firefly algorithm for the heterogeneous fixed fleet VRP. Int. J. Ind. Syst. Eng. **33**(1) (2018)
20. Shi, Y.-J., Meng, F.-W., Shen, G.-J.: A modified artificial bee colony algorithm for vehicle routing problems with time windows. Inf. Technol. J. **11**(10), 1490–1495 (2012)
21. Nahum, O.E., Hadas, Y., Spiegel, U.: Multi-objective vehicle routing problems with time windows: a vector evaluated artificial bee colony approach. Int. J. Comput. Inf. Technol. **3**(1), 41–47 (2014)
22. Alzaqebah, M., Abdullah, S., Jawarneh, S.: Modified artificial bee colony for the vehicle routing problems with time windows. SpringerPlus **5**(1), 1298 (2016)

Coverage-Aware Recharge Scheduling Scheme for Wireless Charging Vehicles in the Wireless Rechargeable Sensor Networks

Govind P. Gupta and Vrajesh Kumar Chawra

Abstract Recent advancement in the wireless power transfer technology has motivated the development of a wireless rechargeable sensor network (*WRSN*). In WRSNs, the formation of an optimal recharging schedule for each wireless charger vehicle is a well known NP-complete problem. To determine the optimal recharging schedule for each wireless charger vehicle, this paper presents a coverage-aware recharge scheduling scheme (*CRS*) where *ACO*-based metaheuristic algorithm is employed. In order to provide fast recharging in WRSN, the proposed scheme employs multiple wireless charger vehicles to perform the charging task. Performance analysis of the proposed scheme confirms its superiority in terms of charging latency.

Keywords Recharging scheduling · Wireless rechargeable sensor networks · ACO-based metaheuristic

1 Introduction

In *WSRNs*, a set of rechargeable sensor nodes are deployed over an area of interest where each sensor node carry a wireless rechargeable battery and one or more wireless charging vehicle (*WCV*) are deployed for efficient recharging task [1, 2]. Recharging of the sensor nodes before their battery power vanishes by using multiple *WCVs* is a very critical task. Hence, optimal scheduling of the WCV is a fundamental research issue in WSRNs [1–5].

In WRSNs, the formation of an optimal recharging schedule for each mobile recharger is a well known NP-complete problem [2–7]. There are many solutions

G. P. Gupta (✉) · V. K. Chawra
Department of Information Technology, National Institute of Technology, Raipur, C.G. 492010, India
e-mail: gpgupta.it@nitrr.ac.in

V. K. Chawra
e-mail: vkchawra.phd2017.it@nitrr.ac.in

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_46

proposed in the literature for designing the recharge schedule for *WCV*. Most of the solutions consider only single *WCV* for completing the recharging task. However, these techniques suffer from charging latency problem [3–6].

In order to enhance charging efficiency and reducing the charging latency, this paper proposes a coverage-aware recharge scheduling (*CRS*) scheme. In the proposed scheme, *ACO*-based metaheuristic algorithm is used for deriving the optimal recharge schedule for each *WCV*. In *CRS*, multiple *WCVs* are used for delay efficient recharging of the deployed sensor nodes. The proposed scheme contains mainly two phases: task assignment process for the *WCV* and derivation of the optimal recharge schedule for each *WCV*. To ensure the coverage of *WCVs*, proposed scheme divides the monitoring area into a set of logical grids and a subset of grids are allocated to each *WCV* for power transfer task. In order to schedule the traversal path of each *WCV*, an *ACO*-based metaheuristic algorithm is used for deriving the recharge schedule for each *WCVs*.

The remaining parts of the paper are structured as follows. A brief overview of the related work on recharge scheduling methods for the *WCV* is described in Sect. 2. Section 3 presents the network model and various assumptions used in this work. Section 4 discussed the proposed coverage-aware recharge scheduling scheme for wireless charging vehicles. In Sect. 5, simulation result analysis and a detailed performance comparison of the proposed scheme with the existing scheme are discussed. Finally in Sect. 6, the paper is concluded.

2 Related Work

In the literature, several research works have been proposed for deriving the recharge schedule for wireless charging vehicles. In [3], the authors have described a mobile charging method where the first clustering algorithm is employed to cluster the nodes based on their energy consumption rate. After the formation of the cluster, a set of nested *TSP* tours are formed for traversing the nodes that require energy. In [4], authors have presented two different charging algorithms, keeping an objective to minimize charging time as well as travel distance.

In [5], Wang et al. have discussed a node deployment scheme where a data gathering vehicle and multiple charging vehicles are employed to balance the energy resources and maximize the network lifetime of the network. In this scheme, only a single mobile charger is used for recharging task. The authors in [6] have presented a multimode wireless charging method where the mobile charger can charge multiple nodes at a time. In this scheme, the network area is partitioned into hexagonal cells. In this scheme, it is assumed that charging point is always at the center of the cell.

In [7], the authors have considered charging problem as two subproblems such as tour formation and assignment problem. In this scheme, a greedy technique-based heuristic algorithm is proposed for solving these problems. In [8], Jia et al. have discussed an integrated solution for charging and routing problem in

WRSNs. In this scheme, a heuristic algorithm is employed to derive optimal charging tour using a predetermined routing tree. This scheme employed a *GA*-based metaheuristic optimizing technique for optimizing charging and the routing tasks. This scheme considered only a single mobile vehicle for charging and data gathering task. Thus it suffers from charging latency problem.

The authors in [9] have proposed a collaborative scheme for solving the recharging problem in *WRSNs*. In this scheme, the network area is divided into a set of concentric circles and authors have tried to optimize the number of charging vehicles required for completing the charging task. In the literature, there are many solutions proposed by different researchers to solve the recharge schedule problem in *WRSNs*. Most of the scheme only utilized a single mobile charger vehicle for completing the recharging task. Thus, suffers from charging latency problem. In this paper, multiple *WCVs* are employed for reducing the charging latency and enhancing total charging efficiency. In addition, the proposed work also considers the coverage issue of recharging so that all required sensor nodes are recharged by the dispatched *WCVs*.

3 Network Model and Assumptions

These research works consider a wireless rechargeable sensor network where network area is logically divided into a set of equal size grids. Each grid can contain a set of sensor nodes and one of the them is selected as cluster head node which collects the sensed data from its cluster members and handovers to mobile sink during its data collection process [11–13]. In this work, we assume that all sensor nodes are homogeneous and haveequal sensing and communication range [13–16]. Figure 1 illustrates the network model used in this paper.

4 Coverage-Aware Recharge Scheduling Scheme for Wireless Charging Vehicles

This section presents a detail explanation of the working of the proposed coverage-aware recharge scheduling scheme (*CRS*). In the proposed scheme, ACO-based metaheuristic algorithm is used for deriving the optimal recharge schedule for each wireless charging vehicle (*WCV*). In CRS, multiple *WCVs* are used for delay efficient recharging of the deployed sensor nodes. The proposed scheme contains mainly two phases: task assignment process for the *WCV* and derivation of the optimal recharge schedule for each *WCV*.

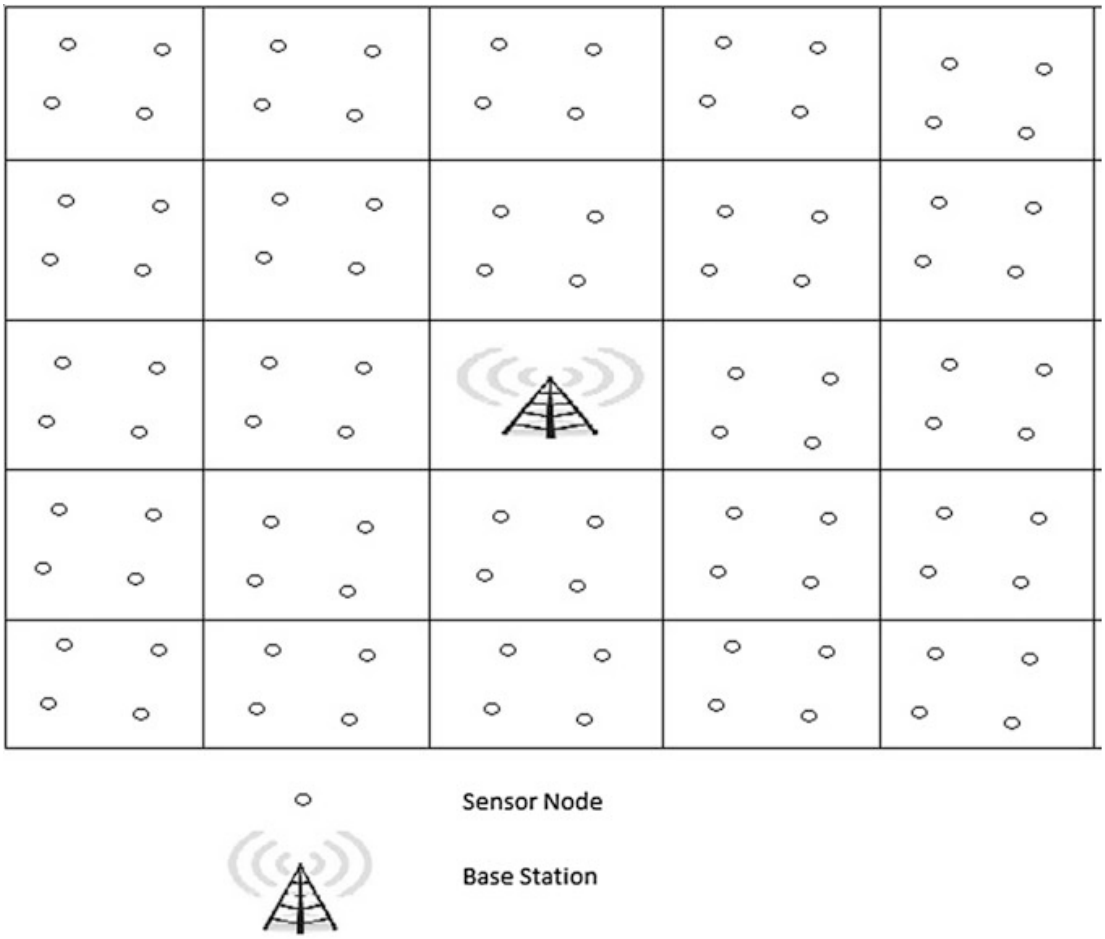


Fig. 1 Network model used in the proposed scheme

4.1 Task Assignment Process for the WCV

In order to assign the recharging task to each involved *WCV*, available area of interest is logically divided into equal size grids and a subset of grids is assigned to each *WCV* for traversing and recharging the sensor nodes located in that grid. In the experiments, the proposed scheme has taken 25 × 25 dimensions for each grid. In order to ensure coverage of recharging task by the *WCs*, the monitoring area is divided into logical grids. For scheduling the traversal path for each *WCV,* an *ACO*-based metaheuristic algorithm is used which is described in Sect. 4.2.

4.2 *Derivation of the optimal recharge schedule for each* WCV

For the derivation of the optimal recharge schedule for each deployed *WCV,* the proposed scheme employed an *ACO*-based metaheuristic algorithm. The working of the *ACO*-based optimal recharge scheduling algorithm is described as follows:

(i) **Initialization** In this phase, the proposed scheme selects a set of location points where *WCV* stay for a fixed charging time to wirelessly transmit power to recharge all sensor nodes located in the grid. Since each *WCV* stays at a particular point of each grid, a population set is randomly generated by selecting a random location point from each grid. Let a *WCV* needs to visit a set of *M* grid so a set of M stay points are randomly selected and need to devise a optimal visiting schedule for traversing theses stay point such that total traveling cost of *WCV* will be minimum. Each *WCV* is dispatched from the base station to visit each grid and recharge all sensor nodes located in each grid. Base station (BS) and a set of stay points (S_i) form a connected graph *G* where source node *BS* connects all destination points S_i. Each link between (BS, S_i) is initialized with a variable called pheromone trail ($\tau_{ij} = \tau_0$). The value of the pheromone trail can be read and updated by ants.

(ii) **Selection of visiting schedule for optimal recharging** In this phase, probability of each link (i, j) of the graph is calculated by using the following formula as given in [17, 18]. Here, N_i is set of neighboring stay points of the points i.

$$p_{ij} = \begin{cases} \frac{1}{d_{ij}} \times \frac{\tau_{ij}}{\sum_{j \in N_i} \tau_{ij}} & \text{if } j \in N_i \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

After calculation of probability (p_{ij}) of each link, a sequence of stay points is selected based on their computed probability value. This sequence of stay points considers as a recharge schedule which starts from the base station and ends at the base station. After getting the recharge schedule, calculate its cost (C_s) by calculating total traveling distance by *WCV*.

(iii) **Pheromone Update** In this phase, after calculation of cost of the derived schedule, the value of pheromone trail (τ_{ij}) for each link is updated using the following formula which is given in [17, 18].

$$\tau_{ij}(t+1) = \rho \times \tau_{ij}(t) + \frac{q}{C_s(t)} \tag{2}$$

Here ρ is the evaporation rate whose value is between 0 and 1. Q is a constant and $C_s(t)$ is the cost of the derived schedule in t iteration. After updating the value of pheromone trail, Step (ii) and (iii) are repeated until we get an optimal cost schedule for *WCV*.

5 Result Analysis

This section presents the result analysis of the proposed scheme and compares its performance with the single *WCV-based* scheme. Implementation of the proposed scheme and the existing scheme are done using *MATLAB R2014*. In the experiments, this paper considers four *WCVs* for recharging tasks. Figure 2 illustrates traversing of each *WCV* for recharging the sensor nodes of each grid.

Figure 3 illustrates the result analysis of the proposed scheme (*CRS*) and compares its performance with the single *WCV*-based scheme in terms of charging latency by varying the speed of the *WCV*. It can be depicted from Fig. 3 that charging latency for the proposed scheme is significantly much lower than the existing scheme. This is due to the use of multiple WCV and derivation of optimal recharging schedule in the proposed scheme.

Figure 4 illustrates the result analysis of the proposed scheme (*CRS*) and its comparison with the single *WCV*-based scheme in terms of charging latency by varying the grid size from (10 × 10) to (60 × 60) within the network. It can be observed from Fig. 4 that charging latency for the proposed scheme is significantly much lower than the existing scheme. This is due to the use of multiple *WCV* and derivation of optimal recharging schedule in the proposed scheme.

Figure 5 illustrates the result analysis of the proposed scheme (*CRS*) and its comparison with the single *WCV*-based scheme in terms of charging latency by varying the number of stay points from 9 to 400 within the network. It can be

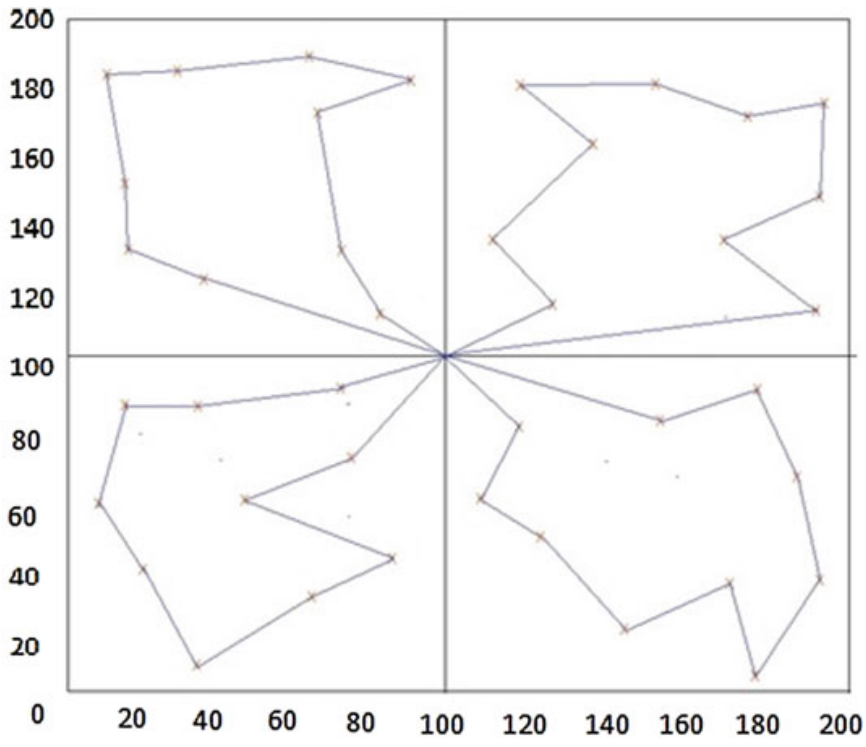


Fig. 2 Illustration of the traversal path of each *WCV* for recharging tasks

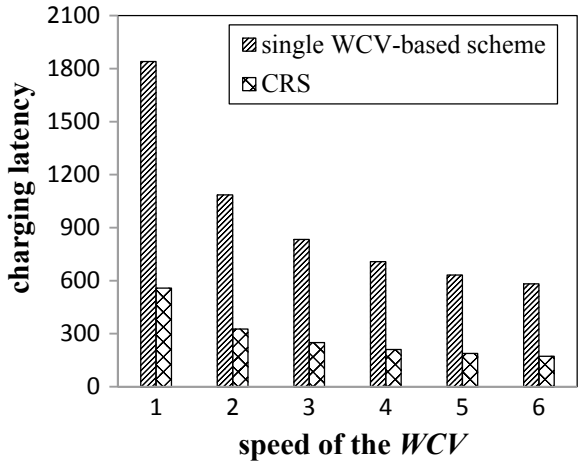


Fig. 3 Speed versus charging latency

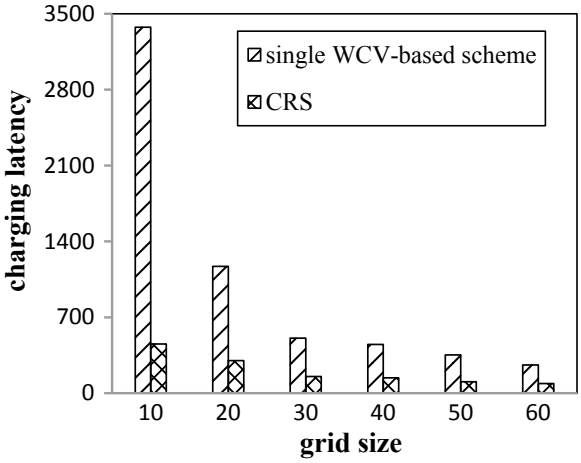


Fig. 4 Grid size versus charging latency

observed from Fig. 5 that as the number of stay points increases, the charging latency also increases. Charging latency for the proposed scheme is significantly much lower than the existing scheme. This is due to the use of multiple *WCV* and derivation of optimal recharging schedule in the proposed scheme.

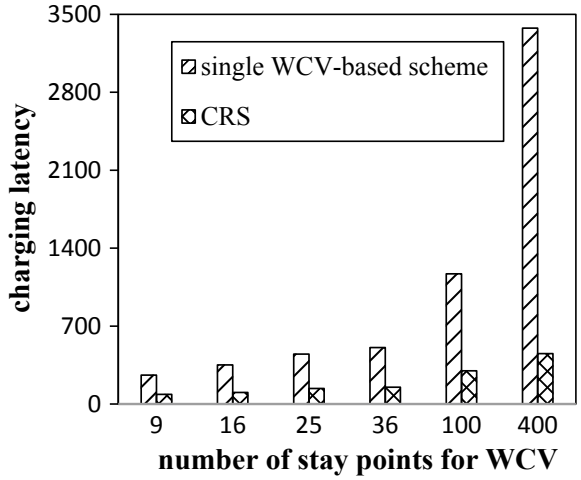


Fig. 5 Number of stay points versus charging latency

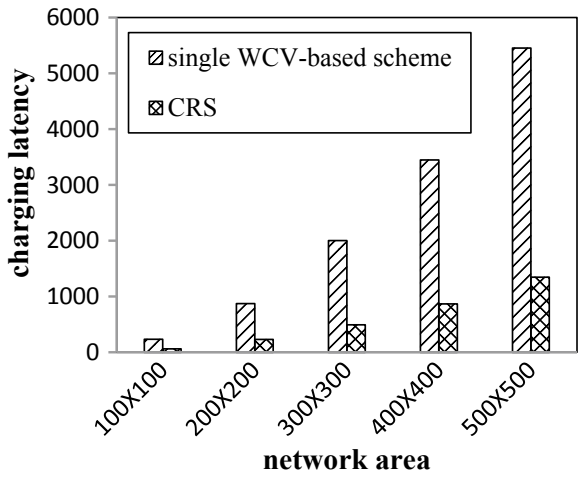


Fig. 6 Network area versus charging latency

Figure 6 illustrates the result analysis of the proposed scheme (*CRS*) and its comparison with the single *WCV*-based scheme in terms of charging latency by varying the size of the network area (100 × 100) to (500 × 500). It can be observed from Fig. 6 that as the size of the network area increases from (100 × 100) to (500 × 500), charging latency also increases. This is due to the fact that the total number of grids visited by a *WCV* also increases. Charging latency for the proposed scheme is significantly much lower than the existing scheme. This is due to the use of multiple WCV and derivation of optimal recharging schedule in the proposed scheme.

6 Conclusion

This paper proposed a coverage-aware recharge scheduling scheme for wireless charging vehicle where ACO-based metaheuristic algorithm is employed to determine the optimal schedule for each wireless charging vehicle. In order to reduce charging latency, the proposed scheme employs multiple charging vehicles that are dispatched by the base station concurrently. Simulation results and its analysis confirm that the proposed scheme performs significantly better than the existing scheme in terms of charging latency.

References

1. He, S., Chen, J., Jiang, F., Yau, D.K.Y., Xing, G., Sun, Y.: Energy provisioning in wireless rechargeable sensor networks. IEEE Trans. Mob. Comput. **12**, 1931–1942 (2013)
2. Liao, J.-H., Hong, Jiang, J.-R.: An adaptive algorithm for charger deployment optimization in wireless rechargeable sensor networks. In: ICS (2014)

3. Fu, L.K., He, L., Cheng, P., Gu, Y., Pan, J.P., Chen, J.M.: ESync: energy synchronized mobile charging in rechargeable wireless sensor networks. IEEE Trans. Veh. Technol. **65**(9), 7415–7431 (2016)
4. Lin, C., Wu, G.W., Obaidat, M.S., Yu, C.W.: Clustering and splitting charging algorithms for large scaled wireless rechargeable sensor networks. J. Syst. Softw. **113**(March), 381–394 (2016)
5. Wang, C., Li, J., Ye, F., Yang, Y.: A mobile data gathering framework for wireless rechargeable sensor networks with vehicle movement costs and capacity constraints. IEEE Trans. Mob. Comput. **65**(8), 2411–2427 (2016)
6. Xie, L., et al.: Multi-node wireless energy charging in sensor networks. IEEE/ACM Trans. Netw. **23**(2), 437–450 (2015)
7. Hu, C., Wang, Y.: Minimizing the number of mobile chargers to keep large-scale WRSNs working perpetually. Int. J. Distrib. Sens. Netw. **2015**, 1–16 (2015)
8. Jia, J., Chan, J., Deng, Y., Wang, X., Aghvami, A.H.: Joint power charging and routing in wireless rechargeable sensor networks. Sensors **17**(10), 2290 (2017)
9. Chen, Z., Chen, X., Zhang, D., Zeng, F.: Collaborative mobile charging policy for perpetual operation in large-scale wireless rechargeable sensor networks. Neurocomputing **270**(2017), 137–144 (2017)
10. Gupta, G.P.: Efficient coverage and connectivity aware data gathering protocol for wireless sensor networks. In: Proceedings of the 3rd IEEE International Conference on Recent Advances in Information Technology (RAIT-2016), pp. 50–55 (2016)
11. Angelopoulos, C.M., Nikoletseas, S., Raptis, T.P.: Wireless energy transfer in sensor networks with adaptive, limited knowledge protocols. Comput. Netw. **70**, 113–141 (2014)
12. Gupta, G.P., Jha, S.: Integrated clustering and routing protocol for wireless sensor networks using Cuckoo and Harmony Search based metaheuristic techniques. Eng. Appl. Artif. Intell. **68**, 101–109 (2018)
13. Gupta, G.P.: Improved Cuckoo Search-based clustering protocol for wireless sensor networks. Procedia Comput. Sci. **125**, 234–240 (2018)
14. Gupta, G.P., Misra, M., Garg, K.: Towards scalable and load-balanced mobile agents-based data aggregation for wireless sensor networks. Comput. Electr. Eng. **64**, 262–276 (2017)
15. Kumari, S., Gupta, G.P.: Differential evolution-based sensor allocation for target tracking application in sensor-cloud. In: Bhateja, V., Coello Coello, C., Satapathy, S., Pattnaik, P. (eds.) Intelligent Engineering Informatics. Advances in Intelligent Systems and Computing, vol. 695. Springer, Singapore (2018)
16. Gupta, G.P., Misra, M., Garg, K.: Energy and trust aware mobile agent migration protocol for data aggregation in wireless sensor networks. J. Netw. Comput. Appl. **41**, 300–311 (2014)
17. Dorigo, M., Blum, C.: Ant colony optimization theory: a survey. Theor. Comput. Sci. **344**(2–3), 243–278 (2005)
18. Dorigo, M., Gambardella, L.M.: Ant colony system: a cooperative learning approach to the traveling salesman problem. IEEE Trans. Evol. Comput. **1**(1), 53–66 (1997)

A Transition Model from Web of Things to Speech of Intelligent Things in a Smart Education System

Ambrose A. Azeta, Victor I. Azeta, Sanjay Misra and M. Ananya

Abstract Several terms have been used to describe Internet of Things; Web of Things (WoT) is a term which can be used interchangeability and it is referred to as the capability of devices to interconnect to the World Wide Web and sharing the information and data to one another. WoT has been mentioned in the literature to improve interconnection between devices at all times. In WoT, two different modes of communication which are generally mentioned in previous studies include person-to-thing (or thing-to-person) and thing-to-thing. This paper presents an architecture for transiting from WoT to speech-enabled WoT known as Speech of Intelligent Things (SoIT). The system employs a combination of technologies such as system design, server-side scripting, speech-based system tools, and data management in developing the SoIT prototype system as a third mode of communication. This paper illustrates a scenario whereby remote monitoring and controlling of WoT devices within the university campus might be difficult to manage by only using the modes discussed in the literature. An evolution of WoT to SoIT was realized using speech technology to provide a prototype system. Technical implications involve using a telephone by connecting an object telephone number (OTN) and dial WoT objects and establish a control mechanism. The research limitation is mainly the cost of dialing an OTN number. The contribution of this paper is to favor and encourage the use of speech technology to enhance the convenience of communication between WoT devices within the school campus.

A. A. Azeta · S. Misra (✉)
Covenant University, Ota, Nigeria
e-mail: sanjay.misra@covenantuniversity.edu.ng

A. A. Azeta
e-mail: ambrose.azeta@covenantuniversity.edu.ng

V. I. Azeta
National Productivity Center, Kaduna, Nigeria
e-mail: victaazeta@gmail.com

M. Ananya
Technical University of Munich, Munich, Germany
e-mail: ge25daj@mytum.de

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_47

Keywords e-Campus · OTN · Speech interface · SoIT · WoT

1 Introduction

Among the web technologies that are presently gaining attention in the research community is the Web of Things (WoT). Web of Things transforms the connection between millions of devices and smart devices and interactions with humans and one another. In WoT, sending and receiving of information exists between cloud and one another, with the possible aim of collection and analysis of huge amount of data accurately [1] suggesting that distinct approaches exist across a variety of technologies such as vehicles which are able to sense fatigue of the user and execute maintenance which is self-scheduled and provide dynamic analysis, execute, and inform the forecasted time of arrival of waiting passengers.

The literature has failed to discover that the concepts of WoT have been applied to e-Learning and school environment to enhance teaching and learning results thereby providing greater achievements than in other fields. In other words, educational institutions are a good place in which WoT can be researched [2].

There are two different modes of the interface in WoT usually discussed in the studies. Thing-to-person or person-to-thing in which there is an interaction between people and things and thing-to-thing in which day-to-day objects communicate with each other. In this research, it is envisaged that control and monitoring of WoT objects remotely might be burdensome with the use of two modes only. Some uncertainties do exist on how the WoT will be controlled and monitored. Some years ago, Internet Protocol Version 6 (IPv6) was developed to facilitate a high number of systems with a distinct IP address that aided the transmission and receiving of information [3]. The entire world seems to approach a global village where the exchange and sensing of data are done through the connection to different types of networks (private and public). Data is collected, processed, and used from time to time from the connected device to provide more information about the devices. This process provides intelligent data for control, plan, management, and decision-making.

A method has been proposed on how to control the WoT objects involving mobile telephony. This method involves dialing the Object Telephone Number (OTN) through the use of a mobile device and connection to any of the WoT objects, which will begin this mechanism. In this environment, different SoIT devices are assigned a value for OTN which is used instead of an IP address for communication among the devices. The communication occurs when an OTN value is dialed via a mobile device. A lot of objects and entities such as traditional office systems, educational programmes, vehicles, mobile devices, and students and staff residents can be replaced by an SoIT-based campus. When universities are compared to this system, universities consist of, accommodations, clinics, athletic facilities, health institutions, hospitals, offices, parking areas, classrooms, restaurants, and libraries [4].

With Speech of Internet Things (SoIT), wearable computer can be used by students to explore the points of the university transportation which can be done by dialing an OTN number, attend their classes online, perform online bank transactions, be up-to-date with the library transactions, in other words, check the availability of a book, be aware of the prices of food items in canteens, and order items in the storage. In an educational environment, a large portion of information needs is shared by students and staffs which include information regarding classrooms, schedules, locations, seminars, lab equipment, student residences, assignments, presentations, sports events, etc. This information is highly relevant and related to people and objects found in the institutional environment. When wireless technology such as WLAN (Wireless Local Area Network) is implemented in conjunction with SoIT platform, technologies, and smart devices, it helps to detect devices or places, thereby making this feasible for fulfilling information needed by the individuals in a typical learning atmosphere in an efficient and effective way [5].

A discussion is made regarding the provision of accommodation for students and staffs of the higher institution in Nigeria whereas higher educational institutions which can provide this facility have not yet implemented the required technology in various other units such as transportation, storage, commercial institutions, cafeterias, classrooms, hospitals and clinics, dormitories, library, etc. In other words, manual processing of transactions is still done in many units. Supervision of a human is needed to control objects such as electrical appliances used in homes and offices. Devices have no interconnection and remote manipulation between them. Mostly, learning takes place 40% online and 60% involves the presence of an individual and this illustrates that ICT and the Internet are not being used with their full capabilities for administration, management, learning, and teaching.

Institutions which have been able to provide adequate accommodation for the staff and students still have the problem of providing exact locations of students to the facility as well as the visiting parents. There are several students' engagement activities on campus. Students can be found in places within the campus such as the library for the borrowing of books, storage, cafeterias, banks, and lecture halls for lectures. They sometimes tend to go home for personal reasons. The availability of transportation services for students and staffs not fully on ground. The wide use of SoIT devices should be encouraged for different systems of the campus to work together for successful management and easy access for everyone by an electronic device (remote control) and manipulation of OTN devices. A framework of SoIT based on the above is proposed in [27]. In this paper, we are extending the SoIT framework by developing a real system with a prototype and implementing it on a system.

This paper has six sections. In Sect. 2, the related work is provided. Section three presents the SoIT framework. The fourth section discusses the potential advantages and disadvantages of using WoT/SoIT. In section five, the SoIT prototype and evaluation is presented while section six concludes the paper.

2 Related Work

A research was carried out in [6] which gives a summary of the difficulties faced in the areas such as reliability, security, mobility, and connectivity of WoT with the use of IPv6 to achieve Internet of Everything (IoE). The main challenges and solutions of IPv6 were also addressed. Some of the future works which discuss the emergence of IoE to make applications mobile, distributed, secure, and powerful. An example of such application is smarte cities which were also highlighted. The study in [7] discussed the necessity of business organizations to develop WoT and show the power of digitalization to entities for the creation of more business opportunities. The design and evaluation of the model are presented that will give researchers and industrialists to record, see, and analyze the current and future commercial areas in WoT.

An attempt is made by studies to give an account of the architecture of WoT. The research that was carried out in [8] introduces key enablers of WoT and gives details of the main components of WoT such as the concept of machine-to-machine (M2M), portability and mobility of devices, protocols for communication that are appropriate for the WoT environment. It also discussed the difficulties faced in WoT. A prototype of navigation-based solutions for WoT and an architectural framework of WoT were also mentioned. In [9], an architecture and cloud-centric vision are given for WoT implementation. This study presented major technological development and applications which will make research in WoT feasible in the future. The authors in [10] showed the perception of the future by Rambus. As the number of devices is increasing dramatically and connection to the Internet is possible, the world is deviating from a PC-based model and going toward interconnected systems.

3 The SoIT Architecture

The SoIT system signifies a vision by which the Internet interconnects on a daily basis using OTN. Physical entities are now connected to the virtual world which can serve as access points to Internet service and can be managed from a distance by phone calls. SoIT is a ubiquitous computing platform with a lot of possibilities for the advancement in society. Despite the numerous benefits, there are a lot of technical challenges and risks. SoIT can only continue to grow if the technological advances in the information and communication area are seen in the upcoming years. Due to the creation of smaller versions of technological devices, reduction in energy consumption and price devices can easily be adapted to daily lives. Smart devices play an important role in SoIT because information and communication technology are causing radical changes in the world [11].

According to authors of [12], smart campus comprises other smart components such as smart technology, lecturers, students, and environment. The framework by Abuelyaman needs a smart campus to contain a smart cafe, smart storage, smart residential areas, smart commercial banks, smart lecture rooms, smart hospitals and clinics, smart transportation, smart devices such as smartphones, tablets, laptops, etc. In SoIT, sensors interact with one another automatically and provide real-time response rather than interacting with humans only. Local area networks (LAN) will be used to mitigate physical devices. In WoT, each device will possess a unique IP address in IPv6 format only whereas in SoIT in addition to the IP address, the device will also have the OTN address.

The phone of the user would keep track of the physical and mental requirements of states through the interconnection of the components which contains the main system. Similarly, the complete set used at the higher institution may make efficient use of resources by sharing and distributing the resources with other components [13]. For example, during the arrival and departure from home, an automation system can control the front porch light according to the needs of the user [14]. Figure 1 shows the SoIT framework for a smart campus. Objects have been placed in a circular manner from 1 to N. Every object has a sensor attached and all of them (sensors) are interconnected to each other by wireless sensor network (WSN). This will aim to detect the transmission signals from the neighboring sensors which form the network. WSN which works for information and communication is also called revolutionary collection method. These sensors enhance the effectiveness and improve the efficiency of the architecture for a smart campus [15].

In SoIT, all the objects have an OTN value on every node in the network. Message transmission between users can be connected simply by calling an OTN for a particular system, for example, staff can call the OTN of a commercial bank to

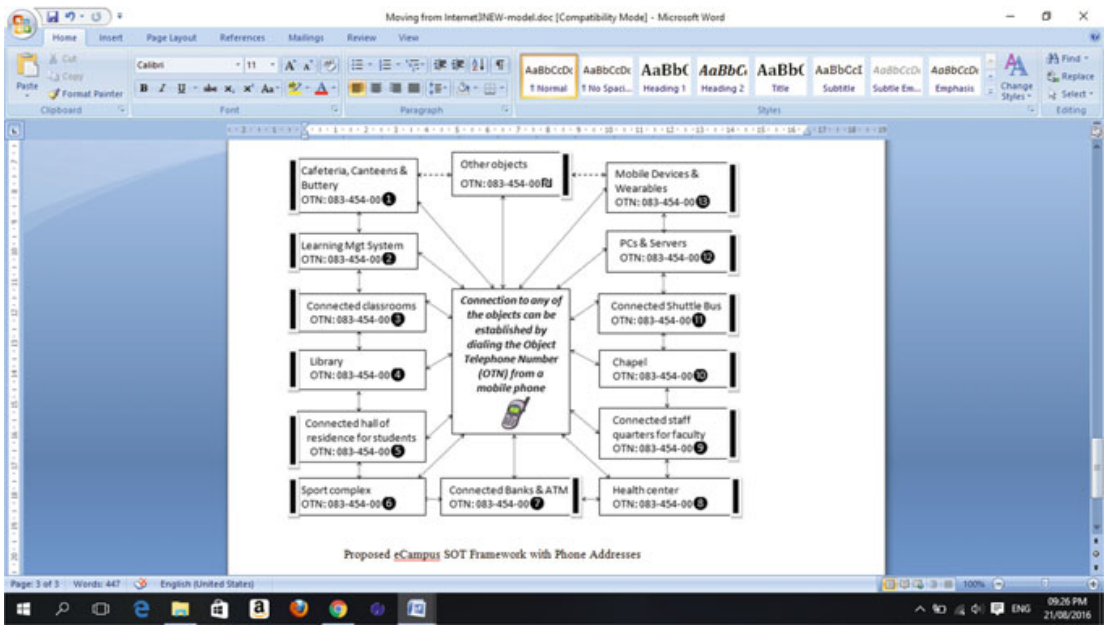


Fig. 1 The SoIT framework for smart campus

perform certain transactions such as checking the bank balance of the account. The student may want to check resources in the library or lecturers might want to connect with the students for giving lectures, answer questions, or make discussions on a particular topic. All these activities can be made possible when objects are interconnected to each other using the SoIT framework.

4 SoIT Prototype and Evaluation

In order to understand the SoIT concept, a simple prototype speech-based application was developed and deployed using VoiceXML [21] for speech interface, PHP for server side, and MySQL for the database. A database with a field named power status was created and assigned the value 1 to put off the computer system, and 0 to restart the computer system. This system allows a user to dial a phone number as an OTN address that is mapped to an object. The system responds through speech interface by saying say 1 to put off or 0 to restart the computer system. Once a 1 is selected, the computer will shut down. The web interface uses an auto-refresh command that constantly checks for the value in the database field of the power status button. The web interface of the system is shown in Fig. 2.

When OTN is dialed from a mobile phone, the system will say "Welcome to shutdown utility", and request for authentication information from the user. After successful authentication, the user responds to the system request by saying 1 or 0. A 1 will shut down the system, and a 0 will restart the system. In experimenting

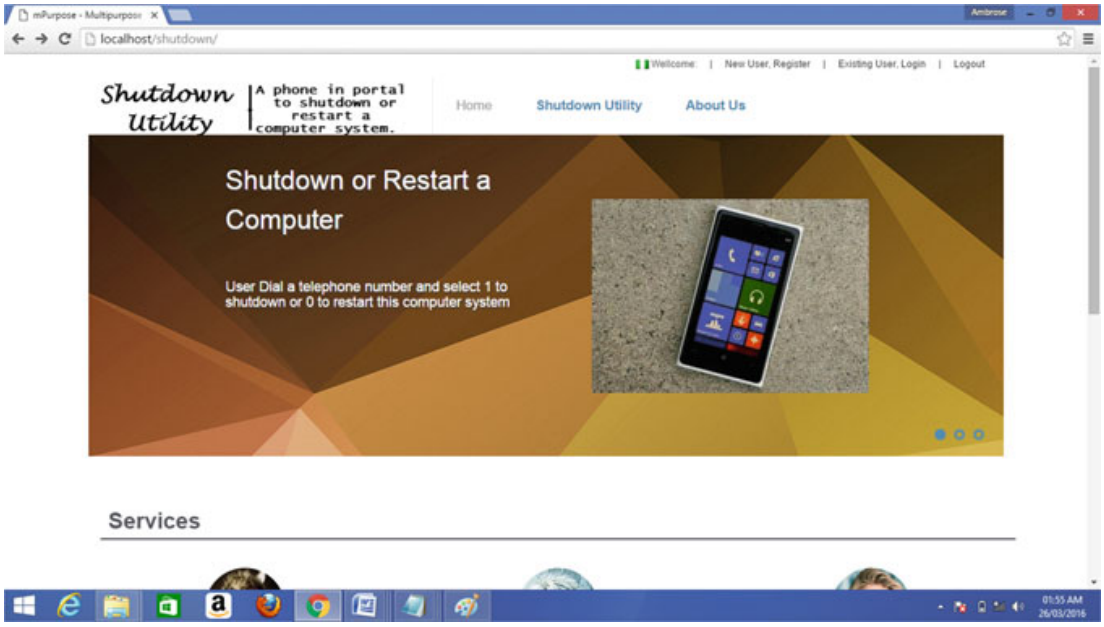


Fig. 2 The web interface. *Source* The researchers

with SoIT with respect to architectural considerations, hardware, software, and deployment requirements, the design and implementation testbed in [22] may be considered.

Typical Scenario

The call sequence in Fig. 3 exemplifies the interaction of the user with the system.

A prototype for an application called VoiceXML (speech user interface) is developed for the web application on a Voxeo voice server and can be retrieved from a mobile phone and as well as in fixed phone-line connections by adopting the following structure: <source country int. dial out #><destination country code><destination area code><generated voice network 7 digit #> [23].

A cognitive method called Cognitive Walkthrough Strategy was developed by Rieman and Redmiles [24]. Cognitive Walkthrough Strategy [24] makes use of one or a group of evaluators which are responsible for checking the user interface by following a certain number of tasks, making learning easy, and accessing the system's understandability.

A questionnaire was developed and used as a survey instrument. Twenty questionnaires were distributed to the users and we got responses from 16. The received responses were analyzed and reported, and the results were presented. We use a five-point Likert scale for our study and analysis (1 = strongly disagree and 5 = strongly agree). The findings show good usability of 4 as the mean rating on a 1–5 scale.

User : Dial OTN address 083-454-0012 to access a computer
System: Wellcome to Shutdown Utility.
What is your userid?
User : My userid is kane.
System: What is your pin?
User : My pin is 4109
(System search the database to authenticate the userid and pin)
System: Say 1 to Shutdown your PC, or
Say 0 to Restart your PC.
User : 1
(System process the input value 1 and initiates a shutdown process)
System: Please save your work and shutdown, the system will go off in 1 minute!!

Fig. 3 Sample call flow to shutdown and restart a computer. *Source* The researchers

5 Advantages and Disadvantages of Using SoIT and WoT

WoT and SoIT have numerous unlimited benefits. WoT systems can make our lives better by conservation of resources, the creation of new opportunities, innovation, and improvement of knowledge [16]. In [17], the authors have explained that WoT is developed to improve connections among devices by using WSN everywhere at a given time. WoT is presently being used positively by some business organizations. Users of business organizations can utilize the services of WoT for data analytics to generate more revenue, monitor the activities and assets, modify new business models, and develop an operational performance for innovation [18].

However, among several obstacles and hindrances of WoT security is one of the areas that prevents WoT from gaining its actual potential. The connection of hundreds of billions of devices gives a serious concern regarding the issue of security. Business owners and stakeholders need to take this issue into consideration for WoT to develop and overcome these obstacles [19]. Other major discoveries in [20] regarding WoT are given below.

- The connection between Secure Socket Layer (SSL) connections with cloud in 19% of mobile applications which are used for the control of WoT devices
- No provision of authentication between client and server in the devices
- Some devices lacked strong password and enforcement strategies
- Some WoT cloud interfaces were not compatible with two-factor authentication (2FA)
- Lack of lock-out or delaying measures against brute-force attacks in some WoT services which hindered the protection of user accounts
- Some devices were not protected from account harvesting mechanism
- Many WoT cloud platforms such as web applications were exposed
- No provision of encrypted firmware updates by majorly all WoT services
- Shortage of IP addresses.

SoIT and WoT and Smart Campus

The basic idea of IoT is the pervasive presence around us of a variety of things or objects connected via enabling technologies such as Radio-Frequency IDentification (RFID) tags, sensors, actuators, and mobile phones, which through unique addressing schemes, can interact with each other and cooperate with their neighbors to reach common goals [33]. WoT includes the interconnection of devices and smart campus includes different smaller systems to interact with other systems to make the bigger system work.

Spoken dialog system has become very popular in the industry with the advent of continuous speaker-independent speech recognition technology over the last few years [28–30]. Speech technology is now combined with human spoken language capabilities to boost learning in students and to make communication between human and computers easier [31, 32]. This incorporation between spoken language and speech technology will make it possible for machines to take over the trivial tasks performed by humans. Intelligent things, on the other hand, will be able to

perform communication among humans and devices and choose the appropriate decision for any situation. Since smart campus consists of smart devices, addition of speech technology will make the device smarter. Current research claims that SoIT and WoT are being seen to work together for a better communication environment.

6 Conclusion

In this paper, a review of WoT and the ability to develop SoIT in a school environment was presented. The office or home appliances, and SoIT devices may be developed which will respond by calling on an OTN address when called by any telephone, cell phone, or mobile. This study has contributed to the area of transiting from WoT to SoIT. WoT is an emerging area of research and scientists are working worldwide by utilizing speech for monitoring systems in the WoT environment which will sharpen the technology and realize the expectation of a smart academic institution system.

With SoIT, students are able to dial an OTN address of a cafeteria and order for food, dial an OTN address of a classroom and receive lecture remotely while in the dormitory. Lecturers and students are able to dial the OTN address of any bank within the school premises and check balance, and also transfer money online. Campus shuttle bus service can be called by dialing the corresponding OTN number. The university library service and health centre can be linked by dialing the corresponding OTN address to ascertain the availability of a certain book and pharmaceutical drugs, respectively. In the case of home appliances, a user may forget to power off a television, fridge, or electric bulbs or any electrical home appliances when leaving home for work; all he needs do is dial the OTN address of the object and say shutdown to power off the system.

Dramatically, changes in information and communication technology also affected the style of education and the choice of students. Thus, a smart campus can be realized by making all the facilities in its jurisdiction smart. Educational institutions are some of the key centers in a smart city, which play an important role in preparing students for tomorrow [25]. In a smart academic environment, worldwide educational resources are shared to all faculty and students, and this will make the academic life of students and staffs easier, comfortable, and attractive. These potentials will be fully realized when IP for everyone in the case of IoT [26] changes to OTN for everyone in the case of SoIT.

Acknowledgements We acknowledge the support and sponsorship provided by Covenant University through the Centre for Research, Innovation, and Discovery (CUCRID).

References

1. David, N.D.: How the internet of things is revolutionizing healthcare. In: White Paper. Healthcare Segment Manager, Freescale Semiconductor (2013)
2. Adamkó, A., Kádek, T., Kollár, L., Kósa, M., Pánovics, J.: New challenges in smart campus applications. In: Recent Advances in Computer Science (2015)
3. Royer, M.: The Internet of Things (IoT). A trends white paper—August 2013. Bellevue College Economic & Workforce Development
4. Donald, S.: Parking on a Smart Campus: Lessons for Universities and Cities. Published by University of California Transportation Center (2005)
5. Rohs, M., Bohn, J.: Entry points into a smart campus environment—overview of the ETHOC system. This work was conducted as part the Entry Points project, which is funded by the ETH World Program (2003)
6. Jara, A.J., Ladid, L., Skarmeta, A.: The internet of everything through IPv6: an analysis of challenges, solutions and opportunities. JoWua **4**(3), 97–118 (2013)
7. Stefanie, T., Christoph, S.: A business model type for the internet of things. In: Research in Progress. 22nd European Conference on Information Systems, Tel Aviv (2014)
8. Katole, B., Sivapala, M., Suresh, V.: Principle elements and framework of internet of things. Int. J. Eng. Sci. **3**(5), 24–29 (2013)
9. Gubbia, J., Buyy, R., Marusic, S., Palaniswami, M.: Internet of things (IoT): a vision, architectural elements, and future directions. Future Gener. Comput. Syst. **29**, 1645–1660 (2013)
10. Rambus: The Internet of Things. How Rambus Sees the Future. 2014 Rambus Inc
11. Mattern, F., Floerkemeier, C.: From the internet of computers to the internet of things. In: Distributed Systems Group, Institute for Pervasive Computing, ETH Zurich (2010)
12. Abuelyaman, E.S.: Making a smart campus in Saudi Arabia. EDUCASE Q. **31**(2), 10–12 (2018)
13. Roman, R., Najera, P., Lopez, J.: Securing the internet of things. Computer **44**(9), 51–58 (2011)
14. StaReport: The Internet of Things: Privacy and Security in a Connected World. Staff Report January 2015. A Workshop Hosted by FTC
15. IEC: Internet of Things: Wireless Sensor Networks. White Paper (2014)
16. Ruggieri, M., Nikookar, H.: Internet of Things: From Research and Innovation to Market Deployment. Rivers Publisher's Series in Communication (2014)
17. Ravikanti, S., Preeti, G.: Future's smart objects in IOT, based on BigData and cloud computing technologies. Int. J. Innov. Res. Comput. Commun. Eng. **3**(7), 6808–6817 (2015)
18. Azure, M.S.: Get Started with the Internet of Things in Your Organization. Introducing Microsoft Azure Internet of Things services. Executive Summary. © 2015 Microsoft Corporation
19. IBM: Saving the future of the IoT. IBM Institute of Business Value. Device Democracy (2005)
20. Barcena, M.B., Queest, C.: Insecurity in the IoT. Candid Wueest, Version 1.0 (2015)
21. Azeta, A.A.: Developing A computerized VoiceXML-based application for E-education: design, implementation and evaluation. Lambert Academic Publishing, Germany (2013)
22. Nati, M., Gluhak, A, Abangar, H., Headley, W.: Smartcampus: a user-centric testbed for internet of things experimentation. In: IEEE WPMC (2013)
23. Voxeo: Voice voice server, available online at: http://community.voxeo.com (2003)
24. Rieman, M.F.J., Redmiles, D.: Usability evaluation with the cognitive walkthrough. In: CHI '95 Proceedings, ACM (1995)
25. APKM: Smart Campus Guidelines-Draft. APKM—Smart Campus Draft Version 123/05/2015
26. Esaki, H.: Smart Campus Implementation Based on Internet-by-Design. Green University Tokyo Project. Internet Society (2015)

27. Azeta, A.A., Igbekele, E.O., Azeta, V.I.: Moving from Web-of-Things to Voice-of-Intelligent-Things in e-Campus. In: AFRICON, 2017 IEEE. IEEE (2017)
28. Nuance: http://www.nuance.com, 2002, as of 15 March 2002
29. SpeechWorks: http://www.speechworks.com, 2002, as of 15 March 2002
30. TellMe: http://www.tellme.com, 2001, as of 15 March 2002
31. Beck, J.E., Jia, P., Mostow, J.: Automatically assessing oral reading fluency in a computer tutor that listens. Technol. Instr. Cognit. Learn. **2**, 61–81 (2004)
32. Rickel, J., Johnson, W.L.: Task-oriented collaboration with embodied agents in virtual worlds. In: Cassell, J., Sullivan, J., Prevost, S., Churchill, E. (eds.) Embodied Conversational Agents, pp. 95–122. MIT Press, Cambridge (2000)
33. Iera, A., Morabito, G., Atzori, L. (eds.): The Internet of Things. Springer, Berlin (2010). ISBN: 978-1-4419-1673-0

Intrusion Detection and Prevention Systems: An Updated Review

Nureni Ayofe Azeez, Taiwo Mayowa Bada, Sanjay Misra, Adewole Adewumi, Charles Van der Vyver and Ravin Ahuja

Abstract The evolution of Information Technology (IT), cutting across several divides in our daily endeavors allows us to interact with all forms of data at different OSI model layers from application to physical. These data are susceptible to intrusion, aimed at compromising its integrity; thus, the need to protect these data, maintain its integrity, confidentiality, and availability cannot be overemphasized. Intrusion Detection and Prevention System (IDPS) is a device or software application designed to monitor a network or system. It detects vulnerabilities, reports malicious activities, and enacts preventive measures to keep up with the advancement of computer-related crimes using several response techniques. This paper presents an updated review on IDPSs given the fact that the most recent review found on the subject was done in 2016. It will also discuss the use of IDPSs to identify vulnerabilities in various channels through which data is accessed on a network or system and prevention mechanisms applied to mitigate against intrusion.

Keywords Vulnerabilities · Malicious activities · Prevention · Network · IDPS

N. A. Azeez · C. Van der Vyver
North-West University, Vaal Triangle Campus, Vanderbijlpark, South Africa
e-mail: nurayhn1@gmail.com

C. Van der Vyver
e-mail: charles.vandervyver@nwu.ac.za

T. M. Bada
University of Lagos, Lagos, Nigeria
e-mail: badataiwo@gmail.com

S. Misra (✉) · A. Adewumi
Covenant University, Ota, Nigeria
e-mail: Sanjaymisra@covenantuniversity.edy.ng

A. Adewumi
e-mail: wole.adewum@covenantuniversity.edy.ng

R. Ahuja
University of Delhi, New Delhi, India

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_48

1 Introduction

During 1984 and 1986, more research on intrusion detection system was done by several researchers. James P. Anderson's [1] presented a research on Intrusion Detection System (IDS). In the mid-1990s, IDS products were first commercialized by two companies, Internet Security System Inc (ISS) and Wheelgroup. They designed a network-based IDS called RealSecure and Netranger, respectively. ISS Inc. released the first version of RealSecure 1.0 for Windows NT 4.0. RealSecure used a knowledge base by matching signatures, however, it was ineffective for new attacks which became a major setback. Wheelgroup's Netranger was a known network-based IDS back in 1995; it functioned by scanning network traffic. Wheelgroup was acquired in by Cisco in February 1998; today, it forms an intrinsic part of Cisco's security.

Many researchers identified the setback in using the knowledge-based technique of matching signatures because it required continuous update of the database to recognize new attacks; more so, network and packet switching began to rise to a high speed from megabits to gigabits per sec. This was a major challenge as it became more difficult to scan through, analyze traffic, and detect attacks in real-time; thus, researchers were burdened with designing an IDS fit for high-speed networks. This led to the invention of host-based IDS, for example, TCP Wrappers, Tripwire, and Snort which provided analysis of system logs in real time. Snort is a free IDS tool, known for its multi-functionality as a network-based and host-based IDS. It was first released by Marty Roesch on December 22,1998 for UNIX systems; later in 1999, a version of Snort (version 1.5) was released; it was effective in analyzing and logging packets in real-time; it was later modified for Windows system by Michael Davis in the year 2000 [2].

Today, as the functionality of IDS advances, attackers now explore means of detecting, bypassing, and disabling IDS before penetrating the infrastructure, resulting in denial of service (DoS). Security experts aim to curb these attacks by using Intrusion Detection and Prevention System (IDPS) architectures which are not visible to attackers by restricting communication permitted among various security components on a network. Due to the gradually increasing number of vulnerabilities, the identification of attack is essential. To this end, a number of reviews have been done on IDPSs in the literature [3] with the most recent one being [4] which was conducted in 2016. A lot has happened since that period that is worthy of reporting. For instance, it was in 2016 that the biggest DDoS attacks powered by a Botnet [5] were recorded. An example is Mirai, a Botnet primarily composed of infected routers and security cameras, low-powered and poorly secured devices which caused a lot of major DDoS attacks [6].

Internet attacks thus must be defined to measure security. Also, in recent times, infrastructure has evolved from a network of systems, private cloud infrastructure to the Internet of Things (IoT) offering several cloud-based services and solutions. While this has provided limitless opportunities on the choice of where to store data, the risk that accompanies these opportunities is also considered enormous because

these data can be compromised via several intrusion methods irrespective of the platform on which the data is stored [7].

2 Motivation

Intrusion is a criminal act committed against an information system, e.g., computer system, network or web infrastructure, such that security on the system is breached or compromised thereby putting it in an insecure state which allows for unauthorized access to the data being hosted by the system. This is done either by bypassing, disabling or exploring vulnerabilities on the system, typically leaving traces which can be discovered by the intrusion detection system.

The most recent review [4] focused majorly on network intrusion detection systems. The elements of security here are basically availability, accuracy, access control, confidentiality, integrity, and identification [8]. Intrusive attacks can be classified into passive [9] and active [5] attacks. An attack is classified as active when data is being altered with the intent to corrupt, destroy the data or the entire network hosting the data [5, 10]. An example of an active attack is interruption. This can include Denial of Service (DoS) [9], Distributed Denial of Service (DDoS) [6], SQL Injection [11], fabrication, replay attack [9], masquerading [12], and modification [5]. Some examples of passive attacks include traffic analysis [9], sniffing [5], and keyloggers [13]. The drawback of the study includes the fact that it mentioned but did not show the classification of intrusion detection systems. Also, network IDPS are known to identify abnormal behavior in network nodes only after the damage has been done to network resources [14]. Furthermore, with the increasing growth of the Web—a global network—network intrusion detection systems are limited in capacity at detecting anomalies on the web.

Newer models comprising a combination of machine learning techniques are being applied to combat attacks on the Web [15]. Some of the models also respond to security threats by detecting various malware intrusions and protocol authentication based on human behavior.

In addition, IDPSs are also being developed for cloud-based environments/systems; hence, the deployment of distributed IDPSs in cloud systems raises many challenges due to the diversity of its services and the complexity of its infrastructure.

3 Intrusion Detection and Prevention Systems

Security systems are designed in practice, to detect, identify and respond to malicious attacks against, a computing system, network or in general, information systems. These attacks are aimed at undermining the integrity of these systems,

steal information and in some cases cause damage to the systems thereby making the system unavailable.

IDS is either a software or hardware that automates intrusion detection, monitors network traffic for suspicious activities, and sends notifications to an administrator [7]. Intrusion Prevention System (IPS) is a software or hardware that prevents an intruder from gaining access to a network, let alone attack a network [16].

Today, security experts are trending with security appliance combining both intrusion detection and prevention capabilities which identify, log possible incidents, prevent attack, and send report to an administrator [16, 17]. Intrusion Detection and Prevention Systems (IDPS) ensures that the protection, availability, integrity, and confidentiality of information systems are guaranteed.

IDPS has become important when putting the security of information systems into context, preserving data, protecting data from unauthorized access or theft, and ensuring continuous availability of services that these information systems provide. Until recently, attackers' focus was on bank customers, where accounts were raided through fraudulent acquisition of personal details either by sending phishing emails or keyloggers, and credit cards were stolen [18].

3.1 Classification of Intrusion Detection and Prevention Systems

IDPS can be classified based on the following criteria [19]:

1. **Type of Intruder**: This is either external or internal. An external intruder is one who does not have any form of access to a network or service, while an internal extruder is one who has authorized access to a network but has restricted permissions on the network.
2. **Type of Intrusion**: There are various types of intrusions which are discussed in chapter two.
3. **Detection Technique**: Three different types of techniques are normally adopted for intrusion detection, misuse detection, anomaly detection, and stateful protocol analysis [20].

    - **Misuse Detection**: It is a signature-based detection method which matches intruder attack patterns, represented by signatures against a knowledge base of known exploit on software and system vulnerabilities. Signature-based detection analyzes and identifies specific patterns of events or behavior that portray an attack using either static, dynamic or hybrid approach. The static approach in misuse detection analyzes intrusion activities against a program and its code before execution. Dynamic approach analyzes attack patterns during or after the execution of a program while hybrid combines both static and dynamic approach to detect malicious attacks [21].

- **Anomaly-Based Detection**: It is a behaviur-based detection technique which gets its input from audit logs generated from the operating system. This type of technique looks for variations in behavior which might indicate masquerading. Anomaly-based detection uses profiles created through monitoring the behavior of typical activity or models of intended behavior of users and applications over a period. It analyzes malicious attacks by using these profiles which represent the normal behavior of users, hosts, network connections, or applications against profiles of monitored activities; any deviation from the norm is triggered via an alert system [21].
- **Stateful Protocol Analysis**: Stateful [22] protocol analysis detects changes of protocol state. Unlike the anomaly detection method, this adopts predetermined universal profiles created based on accepted definitions of protocol behavior created by vendors and industry leaders [23].
- **Rule-based**: This involves making decisions based on rule sets which are defined by domain experts. They can detect known attacks but are incapable of detecting novel attacks. Also, with increase in network traffic, finding and coding rule sets is both difficult and time-consuming.
- **Supervised Machine Learning (ML)**: It does not require model building as in the case of anomaly-based detection. Rather it is able to learn complex malicious and normal models.
- **Unsupervised Machine Learning**: An example is clustering-based IDPS. This approach to intrusion detection involves building models with unlabeled data; however, their performance is not as good as the supervised models (Table 1).

3.2 Types of Intrusion Detection and Prevention Systems (IDPS)

IDPS will be discussed based on the way they are deployed and the type of activities they monitor [16].

- Network-Based Intrusion Detection and Prevention System (NIDPS)
- Wireless Intrusion Detection and Prevention System (WIDPS)
- Network Behavior Analysis (NBA)
- Host-Based Intrusion Detection and Prevention System (HIDPS)

3.2.1 Network-Based IDPS

Network-Based IDPS (NIDPS) technology is designed to analyze packets at the network, transport, and application layer of the Open System Interconnection (OSI) model. NIDPS is most efficient when deployed within a network

Table 1 Advantages and disadvantages of intrusion detection techniques

Detection techniques	Advantages	Disadvantages
Signature-based detection	Effective and simple method of detecting known attacks since it uses signatures of known attacks	Cannot track unknown attacks and variants of known attacks
	Analyzes and identifies attacks by matching malicious signatures against known knowledge base	Attackers can make adjustment to attacks to avoid matching known attack signature
	Detection accuracy for known attacks is high	Requires continuous update of signatures or patterns
	Low computational cost	Newer attack signatures may not be in the signature database
	Rate of false alarm is very low	Detect only the attacks for which they are configured
Anomaly-based detection	Ability to detect and reduce the false alarm rate of unknown attacks	Detection accuracy is based on the amount of collected behavior or features
	Can detect new and unforeseen vulnerabilities	Well-known attacks may not be detected if they fit established a profile
	Dependency on the operating system is minimal and it is able to detect privilege abuse	Intruder can change profile slowly over a period
	Uses statistical test on collected behavior to identify intrusion	Configuring profiles is time-consuming
	No need for priori knowledge of security flaws	Less effective in the dynamic environment due to constant changes in monitored events
	System can also detect attacks from inside a network	
Stateful analysis	Adds stateful characteristics to regular protocol analysis	Resource intensive for protocol state tracing and analysis
	Distinguishes unexpected sequences of commands	Cannot detect attacks that do not violate the characteristics of generally accepted protocol behavior
	Identifies unexpected sequences of commands	
Rule-based	Can easily detect known attacks	Unable to detect unknown attacks
		Finding and coding rule sets is both difficult and time wasting
Supervised ML	Ability to learn complex and malicious models	They are hardly ever used in a real-world scenario owing to the fact that they require sufficient supply of labeled branding data
		Training data is labeled by domain experts which is both costly and time-consuming
Unsupervised ML	Works with unlabeled data on domain specialist may not be required	Performance is not as good as supervised ML

infrastructure with a specific design where it is able to monitor and analyze real-time packets for intrusion and take a decision on any suspicious activity. While NIDPS is effective in analyzing and detecting suspicious network packets in real time, it cannot analyze encrypted traffic, traffic over Virtual Private Network connection (VPN), SSH or HTTPS sessions, and traffic on mobile computing networks [17].

NIDPS has broad intrusion detection capabilities. An example of an NIDPS is KEMP Loadmaster which can detect intrusion and prevent intrusion by shutting down the device.

3.2.2 Wireless IDPS

WIDPS is a variant of NIDPS which monitors and analyzes packets and protocols on a wireless network. Despite its ability to analyze network traffic, WIDPS cannot detect abnormal activities within an application [17].

Advantages

- It is effective for monitoring and analyzing intrusion on a wireless network.
- WIDPS can identify various problematic issues like policy violations and mis-configurations at the WLAN protocol level.

Disadvantages

- It is vulnerable to DoS attacks.
- It cannot monitor and analyze packets on transport layer, network layer, and application layer.
- It is susceptible to evasion technique when an intruder attacks channels that are not currently monitored.

3.2.3 Network Behavior Analysis (NBA)

NBA is also a variant of NIDPS with the ability to monitor and analyze network traffic to detect unusual activities that may emanate from violation of policy, DDoS attacks or malware intrusion [17].

Advantages

- It is effective in detecting DoS attacks.
- It is effective for monitoring packets on transport, network application TCP/IP layer, etc.
- It can monitor and detect threats caused by malware, policy violation, and DDoS.

Disadvantages

- Packets are analyzed in batches, thus delaying the rate of intrusion detection.

3.2.4 Host-Based IDPS

Host-Based IDPS technology is designed for Application level and Operating System intrusion detection and prevention by monitoring the events on a single host on which it is installed. Aside from having the capability of monitoring and analyzing network traffic, HIDPS can analyze system-specific settings such as software calls, local security policy, and audits logs within the host for suspicious activities. HIDPS functionality can be divided into four categories [17]:

- **File System Monitoring**: Every system has a file system to detect and prevent intrusion; HIDPS monitors file systems regularly by checking variations in files size and file content against a known knowledge base. Whenever a system or user file shows a significant deviation, an alert is triggered which sends a notification to an administrator, indicating the detected intrusion and the action taken to prevent access or damage to the file [23].
- **Log File Analysis**: System events are generally logged in a file. These files (event logs) are analyzed constantly by HIDPS for changes or abnormal activities; a typical event log changes in the login information.
- **Connection Analysis**: HIDPS monitors and analyzes network packets (TCP/IP) for suspicious activities such as the ratio or sequence of TCP/IP connections on the host on which it is installed [24].
- **Kernel-Based HIDPS**: The kernel is provided with extra security capability which allows it to identify and prevent intruder activities itself.

Advantages

- It can detect intrusions on host applications, operating system, and network layer traffic.
- It can monitor and analyze suspicious activity on encrypted communication.
- It can detect intrusion on host systems by monitoring its file system, file access, system calls, etc.
- It does not require additional hardware since it is deployed on the host system.
- It can detect misuse of profile because it interacts with the user, as well as server installed application.
- It can prevent intrusion at the system level and detect attacks which NIDPS cannot detect.

Disadvantages

- It does not use a predefined database, therefore, detection accuracy is limited.
- Its uses more host resources, therefore, impacting on the system host performance.
- It must be deployed on each host which is expected to monitor.
- Its monitoring is restricted to the host on which it is deployed.
- There is a possibility of conflict with preexisting security configuration.

4 IDPS, Design, and Architecture

Information systems today have become a target for hackers whose only aim is to undermine the integrity, availability, and confidentiality of data. Therefore, proper design consideration must be put in place when designing an IDPS to increase its capacity to detect a threat and prevent it from gaining access to an information system. In designing an IDPS, the following must be considered [25].

4.1 Speed and Accuracy

These are highly desirable features. The sensitivity of an IDPS in terms of its speed and accuracy determines the rate of false negatives and false positives reported by the system. If the sensitivity of an IDPS is too low, it will have a high rate of false negative where intrusive activities are not detected, thus, no alert is triggered. Whereas if the sensitivity of an IDPS is too high, there is a high tendency of reporting false positives where an alarm is triggered for nonintrusive activities. False negatives and false positives can be triggered by several factors discussed below [26].

Causes of False-Negative Alerts:

- Improper spanning of switch ports which can cause network traffic to overwhelm the switch which can contribute to events with false-negative triggers.
- Flaws in the design of encrypted traffic which are usually not clear to the IDPS.
- A poorly written signature which does not have the capacity to detect an attack even though the attack is known.
- Improper communication of change management on network and server infrastructure to the information security team.
- Intrusive attacks caused by unpublicized or new attacks thus making it invisible to existing signatures.

Causes of False-Positive Alerts:

- A reactionary traffic alarm caused by equipment failure can trigger a false positive alert. For example, an ICMP flood caused by unreachable destination can trigger a false positive alert.
- An equipment-related alarm, e.g., a load balancer can trigger an alert generated from unrecognized packets from the equipment itself.
- A poorly written client software can trigger alerts of policy violation, e.g., alerts triggered by software bugs.
- Alerts triggered by unmalicious events.

4.2 Logging Capabilities

The logging capability of an IDPS is also very important because it facilitates its ability, identifies, detects, and reports malicious activities. The following are the logging capabilities to consider when designing an IDPS:

- IDPS must be able to log time stamps which include the date and time the malicious activity occurred.
- IDPS must be able to log connection ID, usually a unique number assigned to a session or a TCP connection.
- IDPS must effectively log an alarm type, set its severity rating, impact, and the priority of attack.
- IDPS must have the capacity to analyze protocols like TCP, UDP, ICMP at the network, application, and transport layer.
- IDPS must be able to identify the source and destination IP of connections and determine the number of bytes transmitted over the connection.
- IDPS must effectively understand the characteristics of application request and responses.

4.3 Information Gathering Capabilities

For an IDPS to be effective in detecting and preventing malicious activities on an information system, it must be able to gather information about the system upon which it is deployed.

- IDPS must be able to gather information on host profiles which include host IP and their corresponding MAC address.
- Ability to determine the OS version to enable it to determine the type of vulnerability it is susceptible to.
- Ability to identify network characteristic by gathering information on changes in network configuration.

4.4 Architecture of IDPS

Depending on the expected outcome, IDPS can be deployed using the following architecture [20]:

- Centralized: This architecture collects data centrally, sends it to a single location for analysis. Data collection is either from a single host or from several hosts.

- Hierarchical: This architecture collects data from several hosts which are analyzed according to the layers of the deployed IDPS.
- Distributed: This architecture collects data host by host and it is analyzed.

5 Conclusion

Security threats and incidents have evolved and pose a great challenge to information systems; thus, the importance of deploying an IDPS cannot be overemphasized and efforts to create more security techniques must continue to ensure that the integrity, originality, and confidentiality of information systems is guaranteed, thus making it accessible to everyone when the need arises.

IDPS in its various forms according to Chap. 3 is essentially beneficial in that it has the capacity to identify and detect vulnerabilities and prevent all forms of intrusion discussed in Chap. 2. However, consideration must be given to the type of deployment method to get the best of IDPS.

Furthermore, IDPS has extensive logging capacity which makes it effective in intrusion detection, most especially against signatures of various attacks against network systems; this has made a necessity for enterprise environment to protect data, as data sharing of all sorts has evolved to a global trend.

Lastly, when multiple IDPS technologies are combined into a single protection solution, it reduces management costs considerably because IDPS engages several techniques in intrusion detection and prevention; thus when developing a security strategy, it is important that it is comprehensive to stay ahead of the next threat.

References

1. Anderson, J.P.: Computer Security Planning Study. Washington (1972). Retrieved from https://pdfs.semanticscholar.org/0735/6c5477c83773bd062b525f45c433e5b044e8.pdf
2. Bruneau, G.: The History and Evolution of Intrusion Detection, vol. 7 (2001). Retrieved from https://www.sans.org/reading-room/whitepapers/detection/history-evolution-intrusion-detection-344
3. Patel, A., Taghavi, M., Bakhtiyari, K., Júnior, J.C.: An intrusion detection and prevention system in cloud computing: a systematic review. J. Netw. Comput. Appl. **36**, 25–41 (2013)
4. Amudhavel, J., Brindha, V., Anantharaj, B., Karthikeyan, P., Bhuvaneswari, B., Vasanthi, M., Nivetha, D., Vinodha, D.: A survey on intrusion detection system: State of the art review. Indian J. Sci. Technol. **9**, 1–9 (2016)
5. Ahmad, K., Verma, S., Kumar, N., Shekhar, J.: Classification of internet security attacks. In: Proceedings of the 5th National Conference; INDIACom-2011. New Delhi (2011). Retrieved from https://www.researchgate.net/publication/262494946_Classification_of_Internet_Security_Atta
6. Symantec.: Internet Security Threat Report. Mountain View, CA 94043 (2017). Retrieved from https://www.symantec.com/content/dam/symantec/docs/reports/istr-22-2017-en.pdf
7. Kemmerer, R.A., Vigna, G.: Intrusion detection: a brief history and overview, pp. 27–29 (2002). Retrieved from https://www.computer.org/csdl/mags/co/2002/04/r4s27.pdf

8. Persa, S.: Network Security (2003). Retrieved from https://utcluj.ro/pub/docs/cursuri/prc_eng/stallings/securitarea.ppt
9. Pawar, M.V., Anuradha, J.: Network Security and Types of Attacks in Network, pp. 504–506 (2015). Retrieved from https://ac.els-cdn.com/S1877050915006353/1-s2.0-S1877050915006353-main.pdf?_tid=84c3d323-6ab4-4ca1-86f5-7eafc2cbfb30&acdnat=1530640171_af15fb42c5d503b379a7da902477f68d
10. Bloomberg, J.: Cybersecurity Lessons Learned From 'Panama Papers' Breach. The Little Black Book of Billonaire Secrets (2016). Retrieved from https://www.forbes.com/sites/jasonbloomberg/2016/04/21/cybersecurity-lessons-learned-from-panama-papers-breach/#453217b12003
11. Clarke, J.: SQL Injection Attacks and Defense, 2nd edn. Elsevier, Waltham (2012)
12. Salem, M.B., Stolfo, S.J.: Data collection and analysis for masquerade attack detection: challenges and lesson learned. Columbia University, Computer Science. New York: Department of Computer Science, Columbia University (2011). Retrieved from https://doi.org/10.7916/D8D50VV1
13. Wood, C.A., Raj, R.K.: Keyloggers in Cybersecurity Education. New York: Rochester Institute of Technology (2010). Retrieved from https://pdfs.semanticscholar.org/d1d4/628a22e8d27c6cd202839d7bf0e3a7c7ea91.pdf
14. Yerur, S.V., Natarajan, P., Rangaswamy, T.R.: Proactive hybrid intrusion prevention system for mobile adhoc networks. Int. J. Intell. Eng. Syst. **10**, 273–283 (2017)
15. Johnson Singh, K., Thongam, K., De, T.: Entropy-based application layer DDoS attack detection using artificial neural networks. Entropy **18**, 1–17 (2016)
16. Chee, J.: Host Intrusion Prevention Systems and Beyond, vol. 26 (2008). Retrieved from https://www.sans.org/reading-room/whitepapers/intrusion/host-intrusion-prevention-systems-32824
17. Letou, K., Devi, D., Singh, J.Y.: Host-based intrusion detection and prevention. Int. J. Comput. Appl. **0975–8887**(69), 27–32 (2013)
18. NSS Labs.: Security Value Map. Next Generation Firewall (NGFW), p. 1 (2017). Retrieved from https://www.nsslabs.com/research-advisory/security-value-maps/2017/ngfw-svm-graphic/
19. Santos, K.B., Chandra, S.T., Phani, R., Ratnakar, M., Baba, D.S., Sudhakar, N.: Intrusion detection system- types and prevention. Int. J. Comput. Sci. Inf. Technol. 77–82 (2013). Retrieved from http://ijcsit.com/docs/Volume%204/Vol4Issue1/ijcsit2013040119.pdf
20. Singh, A.P., Singh, M.D.: Analysis of host-based and network-based intrusion detection system. Comput. Netw. Inf. Secur. 41–47 (2014). Retrieved from http://www.mecs-press.org/ijcnis/ijcnis-v6-n8/IJCNIS-V6-N8-6.pdf
21. Ghafir, I., Husak, M., Prenosil, V.: A survey on intrusion detection and prevention (2014)
22. Williams, T., Shirley, M.: Next Generation Intrusion Prevention System (NGIPS) Test Report. NSS Labs (2017). Retrieved from https://www.forcepoint.com/resources/reports/nss-labs-2017-next-gen-ips-report
23. Masaryk University, Faculty of Informatics. Brno, Czech Republic: Researchgate. Retrieved from https://www.researchgate.net/profile/Ibrahim_Ghafir/publication/305957314_A_Survey_on_Intrusion_Detection_and_Prevention_Systems/links/57a75a2708aefe6167bc1de0/A-Survey-on-Intrusion-Detection-and-Prevention-Systems.pdf?origin=publication_detail
24. Scarfone, K., Mell, P.: Guide to Intrusion Detection and Prevention Systems (IDPS). National Institute of Standards and Technology Special Publication 800-94, 127 (2007). Retrieved from https://nvlpubs.nist.gov/nistpubs/Legacy/SP/nistspecialpublication800-94.pdf
25. Sharifi, A., Zad, F.F., Noorollahi, A., Sharifi, J.: An overview of intrusion detection and prevention systems (IDPS) and security issues. IOSR J. Comput. Eng. (1OSR-JCE) **16**(1), 47–52 (2014)
26. Stawowski, M.: The Principles and Good Practices for Intrusion Prevention Systems Design, vol. 25 (2006). Retrieved from https://pdfs.semanticscholar.org/8cac/01d90c44ae710719df662f858ca17ffc96d1.pdf

Simulation-Based Performance Analysis of Location-Based Opportunistic Routing Protocols in Underwater Sensor Networks Having Communication Voids

Sonali John, Varun G. Menon and Anand Nayyar

Abstract Recently, Underwater Wireless Sensor Networks (UWSNs) have emerged as a prominent research area in the networking domain due to their wide range of applications in submarine tracking, disaster detection, oceanographic data collection, pollution detection, and underwater surveillance. With its unique characteristics like continuous movement of sensor nodes, limitations in bandwidth and high utilization of energy, efficient routing and data transfer in UWSNs have remained a challenging task for researchers. Almost all the protocols proposed for terrestrial sensor networks are inefficient and do not perform well in an underwater environment. Recently Location-Based Opportunistic Routing Protocols have been observed to perform well in UWSN environments. But it is also observed that these protocols suffer from performance degradation in UWSN networks with communication voids. The objective of this research paper is to discuss the working of major Location-Based Opportunistic Routing Protocols in UWSNs with communication voids and to highlight their issues and drawbacks. We analyzed the Quality of Service parameters, packet delivery ratio, end-to-nd delay, throughput, and energy efficiency of two major Location-Based Opportunistic Routing Protocols, i.e., Vector-Based Forwarding (VBF) and Hop-by-Hop VBF (HH-VBF) in UWSNs with communication voids using NS-2 simulator with Aqua-Sim extension. Simulation results state that both VBF and HH-VBF protocols suffered from performance degradations in UWSNs with communication voids. In addition to this, the paper also highlights open issues for UWSN to assist researchers in designing efficient routing protocols for UWSNs having multiple communication voids.

Keywords Aqua-Sim · Communication void · Hop-by-Hop Vector-Based Forwarding (HH-VBF) · NS-2 · Opportunistic routing · Performance analysis · Quality of Service (QoS) · Underwater Wireless Sensor Networks (UWSNs) · Vector-Based Forwarding (VBF)

S. John · V. G. Menon (✉)
SCMS School of Engineering and Technology, Ernakulam, India
e-mail: varunmenon@scmsgroup.org

A. Nayyar
Graduate School, Duy Tan University, Da Nang, Vietnam

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_49

1 Introduction

Recently, Underwater Wireless Sensor Networks (UWSNs) [1, 2] have emerged as a prominent research area in the networking domain. The increasing interest in applying sensor networks into the underwater environment has found numerous applications in civilian and military fields such as coastline surveillance, pollution detection, and underwater surveillance. In UWSNs, a group of sensors is placed at different depths in the ocean to gather information. This collected information is then forwarded to the target devices positioned at the surface via a network of intermediate nodes. The data is then stored, processed, and passed to different applications for appropriate utilization. As radio signals have many limitations due to its mediocre propagation through water, the acoustic medium is used between the sensor nodes in UWSNs [3]. Most of the applications are highly sensitive and their success relies on the accuracy of the collected data. Many recent applications of underwater sensor network ranging from aquaculture to the oil industry, instrument monitoring, climate recording, natural disturbances prediction, and study of marine culture depend heavily on the accuracy of the collected data. Thus, efficient routing and data transfer among the nodes is vital in UWSNs.

Routing and data transfer from the sender to the sink stations, through the network, have been a tough task for researchers. Incessant movement of the nodes placed in the water due to the difference in the ocean environment and ocean currents is a major challenge. Limitations in bandwidth, frequent link interruptions, increased delay of data transfer, interference caused by marine mammals, etc., are some of the other elements influencing the efficiency of routing in UWSNs [4, 5]. Due to these unique characteristics, almost all the protocols drafted for conventional sensor networks do not perform well in UWSNs [6]. Recently, Opportunistic Routing Protocols (ORPs) [7–10] have found to give better performance in data transfer in UWSNs. The major advantage of this routing strategy is that it dynamically selects one best forwarder device from a group of candidate devices. This selection is based on the present scenario of the network, which leads to better performance of this latest category of protocols. Many ORPs have been designed for ad hoc, terrestrial sensor, and UWSNs [11–16]. Among all categories in opportunistic routing, Location-Based Opportunistic Routing Protocols (LBORPs) has found to give better performance. LBORPs make use of the knowledge on the present location of the devices to dynamically route message packets from the sender to the target. A few Location-Based ORPs have been efficiently used for data forwarding in UWSNs [17–23].

Communication holes or voids have been a major problem in most of the dynamic sensor networks. Often defined as the unreachability problem, the source node suffers from lack of adequate forwarder nodes in its transmission range. Frequent movement of sensor nodes in underwater increases this problem further. Limited research has been performed to determine the working and performance of LBORPs in UWSNs with communication voids [24–27]. In large UWSNs with a limited number of sensors, it is very significant to have a protocol that can handle

communication voids efficiently. In this article, we analyze the working and performance of two major LBORPs in UWSNs with communication holes. Vector-Based Forwarding (VBF) [21] and Hop-by-Hop VBF (HH-VBF) [20, 28] have been used by many applications for data transfer in UWSNs. VBF is an LBORP that constructs a virtual vector pipe between the sender and the sink node for routing. Only the devices within the virtual vector are selected for forwarding the data packet. Extending the VBF protocol, the HH-VBF protocol makes use of different virtual pipes for each node and the direction of the virtual pipe changes during the entire time of transmission.

Objectives of this research paper are

- To study the functioning of major LBORPs designed for UWSNs and to determine their issues and drawbacks.
- To analyze the simulation-based performance comparison of VBF and HH-VBF protocols in UWSNs with communication voids on parameters—PDR, throughput, end-to-end delay, and energy efficiency.
- And, to discuss the issues, challenges, and future directions in UWSNs routing research.

This research article is structured as follows. Functioning of the major Location-Based ORPs in UWSNs is discussed in detail in Sect. 2. The section also discusses the issues and drawbacks with these protocols. Section 3 describes the working of VBF and HH-VBF in a comprehensive manner. Section 4 presents the simulation-based analysis of VBF and HH-VBF protocols in UWSNs having communication voids using NS2 + Aqua-Sim simulator. Section 5 enlists open issues and challenges of existing ORPs in UWSNs. The paper concludes in Sect. 6 with future research directions.

2 Literature Survey

This section examines the functioning of some of the major LBORPs proposed for UWSNs. One of the most accepted protocols in UWSNs is Vector-Based Forwarding (VBF) [21]. VBF uses the details on the present position of the sensor devices for routing the information packets. In VBF, within a fixed virtual pipe the information packets are forwarded between the source and destination devices. These information packets are then transmitted along the redundant paths and thereby remain stable against packet loss. The major limitation with VBF is that the construction of a solitary virtual pipe will decrease the performance of routing in various node density areas and the productivity drops with communication holes.

Hop-by-Hop Vector-Based Forwarding (HH-VBF) [20] is a variant of the VBF protocol that uses virtual pipes from every intermediate device to the target device. Every device then dynamically makes packet forwarding choices with reference to its present location in the network.

Directional Flood Based Routing (DFR) [29] is a receiver-based, stateless, and LBORP in which each and every node is informed about the position of the destination device and one-hop neighbors. DFR utilizes a limited flooding method, in which every node passes its location details to all the other nodes and achieves more reliability in data transmission in the network. The flooding mechanism results in duplicate transmissions and energy loss. DFR also failed to address the void problem.

One of the earliest protocols that functions with reference to the pressure information proposed for UWSNs was Depth-Based Routing (DBR) [18]. DBR utilizes the knowledge on the depth of devices placed underwater to decide whether to transmit the packets or not. When a packet is received, each node will forward the packet to a smaller depth than that stored in the packet. Otherwise, the packet is discarded. Both communication holes and node mobility are the major problems influencing the limited performance of DBR.

GEographic and opportunistic routing with Depth Adjustment-based topology control for communication Recovery over void regions (GEDAR) [30] is another location-based anycast ORP that sends information packets from intermediate nodes to different target nodes. The protocol uses periodic beaconing to get the position information of every node in the network. The protocol suffers from increased duplicate message transmissions.

Hydraulic Pressure-Based Anycast Routing (HydroCast) [31] also utilizes the knowledge on the depth of the deployed device to choose the candidates for routing from the nearest nodes. HydroCast elects a subgroup of neighboring devices with the largest greedy progress to the target node. The performance of the protocol is not satisfactory with voids.

Multi-Path Routing (MPR) [32] helps to solve the data concussion issue at the sink nodes. The data collision is prevented at the receiving packets by creating a route that consists of various sub-paths between the sender and the sink. Geographic Partial Network Coding (GPNC) [33] is another geographic, partial network coding-based protocol proposed for UWSNs. Void is a major concern for this protocol too.

Focused Beam Routing (FBR) [34] was designed to minimize energy drainage during the routing process by controlled flooding of the packets. Although a preventive void handling technique was used, high delay was incurred. Void-Aware Pressure Routing (VAPR) [33] uses an opportunistic data forwarding technique coupled with information on the pressure of the deployed devices. VAPR selects a subgroup of candidate devices with the greatest progress to target. The holding time details of two-hop neighboring nodes helps to deal with the void areas but imposes a high overhead on the system. In Relative Distance-Based Forwarding (RDBF) [35], the packets are sent via the nodes that are nearest to the target node. RDBF uses a fitness value as an upper limit to supervise the number of transmitting devices to the sink. Moreover, the performance of RDBF also comes down with the presence of communication holes in the network.

In the next two sections, we discuss the working of the two LBORPs, VBF, and HH-VBF, used in UWSNs. VBF and HH-VBF utilize the knowledge of the location

of nodes in the network for routing. These protocols use greedy forwarding mechanism that chooses the node nearer to the target as the next forwarder for every data packet. We analyze the performances of these two protocols in UWSNs having multiple communication voids. We have selected these two protocols for the analysis because they are the two popular protocols used by various applications in UWSNs. Further, the design of these protocols is less complex and also incurs less overhead compared to other protocols.

3 Location-Based Opportunistic Protocols

3.1 Vector-Based Forwarding (VBF)

VBF [21, 36] was proposed to addresses the problem of energy limitation and efficient packet delivery. VBF utilize the knowledge of the location of nodes in the network for routing. Each data packet in VBF contains the three position fields named as OP1, TP1, and FP1. This represents the coordinates of the source, the sink, and the forwarder node. The RANGE field in the packet deals with node mobility. As soon as the information packet advances to the region marked by its TPI, it is flooded in that region controlled by the RANGE field. The virtual routing vector from the source to the sink describes the transmitting path. The RADIUS field inside the packet is a precomputed value that is utilized by intermediate devices to check whether they are near to the vector pipe and suitable for packet transmission. In order to limit the number of intermediate transmitting devices and to conserve energy in the network, VBF makes use of a self-adaptive algorithm. Upon getting a data packet, each node first computes whether it is inside the virtual vector and whether it can act as a potential forwarder to the next node. Each potential member device awaits a limited period of time to check its desirability value in the pipe. It describes the closeness of the current node to the past forwarder node, and the virtual pipe between the sender and the sink. The waiting time depends on the desirability factor. If a node is more desirable, then it waits for less time. During this time, the device pays attention to the channel to observe the number of devices that are transmitting the same information packet as the present device. When the period expires, the node will transmit its packet if and only if the reduced desirability value of the other nodes is less than a precomputed level (Fig. 1).

3.2 Hop-by-Hop Vector-Based Forwarding (HH-VBF)

Hop-by-Hop Vector-Based Forwarding (HH-VBF) [20] is a variant of VBF that constructs virtual routing pipes from each hop in the network as packets travel from

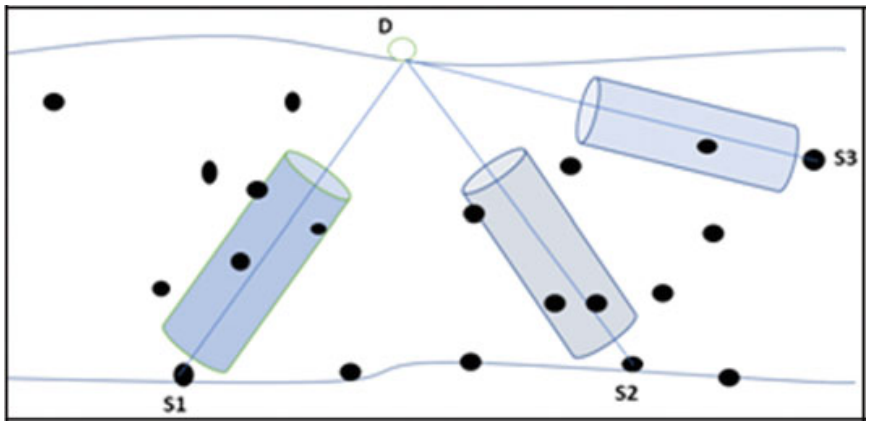


Fig. 1 Virtual vector pipe from sender to receiver node in VBF

the sender device to the receiver. This protocol creates a virtual routing pipe from each hop to the target device. Many studies have discussed the better working of HH-VBF in comparison with VBF. The pipeline radius is similar to the transmission count of the node. The protocol improves the direction of the transmitting pipe hop by hop, in the entire life period and in this way, all transmitting devices can generate a routing choice based on the present local topology details. By changing the direction of flooding pipeline dynamically, the performance can be enhanced. When a stable routing vector radius is set, the protocol shows a reduced performance in the sparse network compared with VBF. However, both the protocols are exposed to interferences caused by marine mammals as the data forwarding happens only inside the pipe. Here, the transmission can be interrupted when marine mammals block the pipe. There is an increased choice of finding a more acceptable forwarder within the hop-by-hop vector pipeline; HH-VBF gives better Packet Delivery Ratio (PDR). However, both the protocols do not succeed to yield energy fairness within the network. Also, they fail to effectively handle the communication hole problem.

The working of HH-VBF is described in Fig. 2. The sender node S2 wants to transmit an information packet to the destination device D2. Here S2 creates a virtual pipe to the destination D2. Once the packet reaches the intermediate sensor node M, it creates a virtual pipe to the destination. Now there are more nodes (N and P) included for sending the data packet to the target location. HH-VBF uses this forwarding strategy until the data reaches the target location. HH-VBF improves the performance in the network by dynamically moving the direction of the flooding pipe. Figure 3 illustrates the communication void problem in UWSNs. We can see that the source node S1 and S2 are trying to forward data packets to the destination D by creating virtual pipes. Data from S2 reaches the destination node located at the surface, but data from source S1 is unable to proceed due to the communication void near to the surface. Table 1 highlights the technical comparison of both the routing protocols.

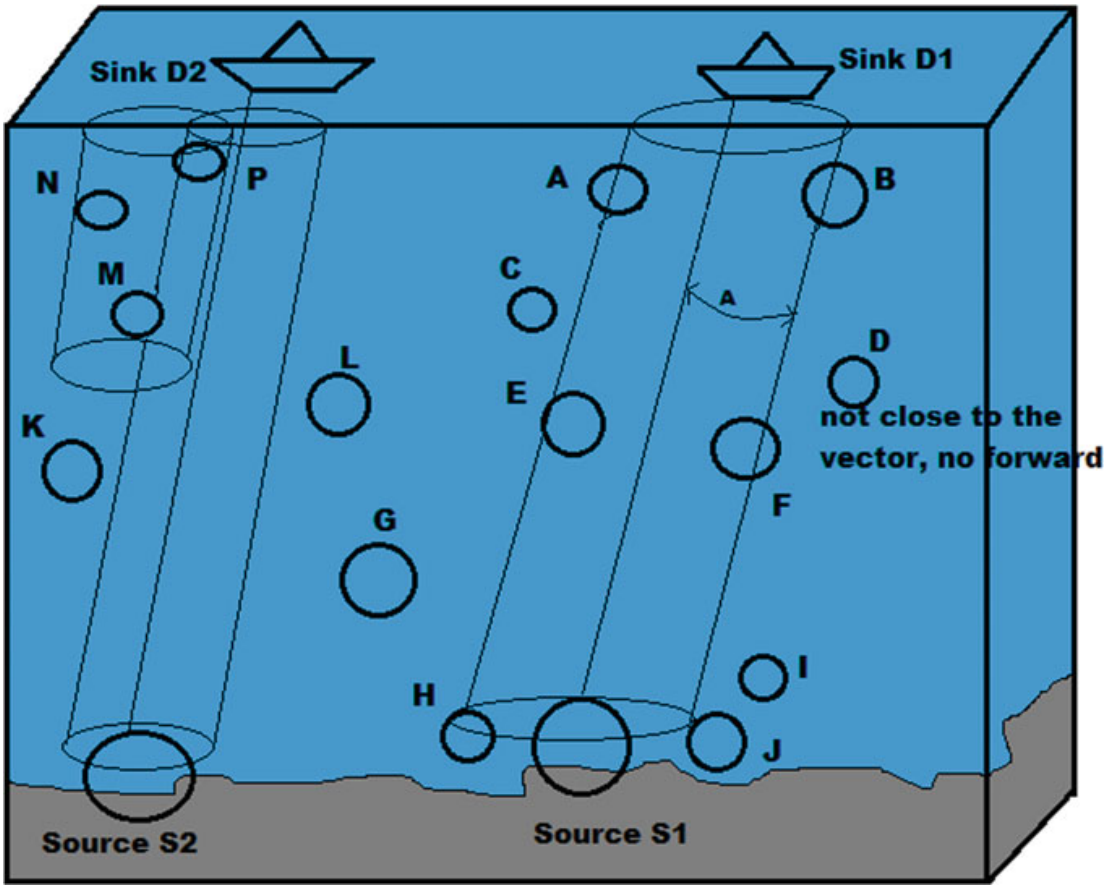


Fig. 2 Illustration of the functioning of VBF and HH-VBF

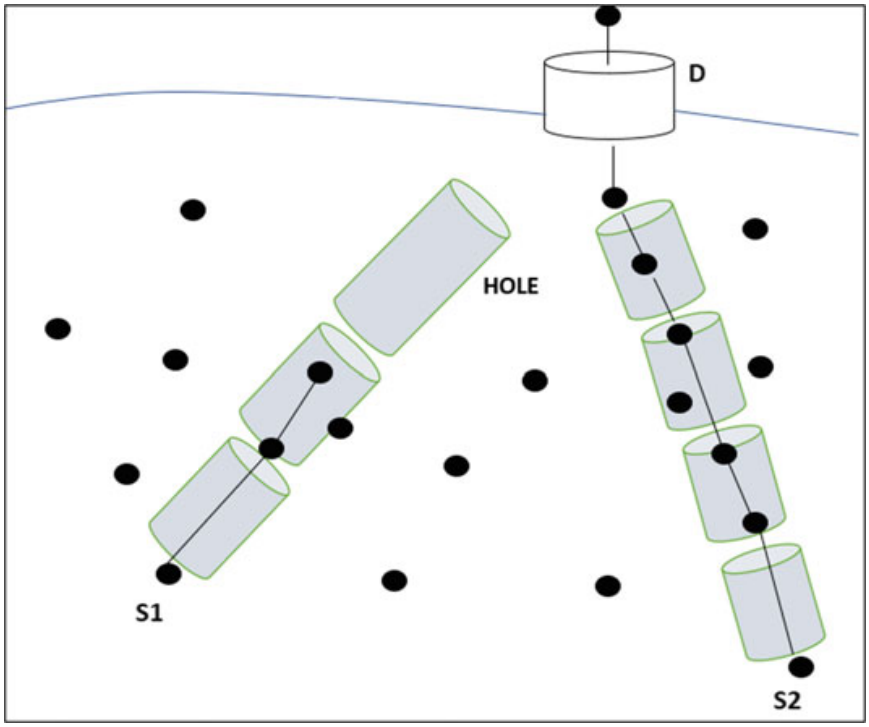


Fig. 3 Communication void problem with VBF

Table 1 Comparison of VBF and HH-VBF

Parameter	VBF	HHVBF
Metric used	Node location	Node location
Candidate coordination	Timer based	Timer based
Void handling	No	No
Routing	Single virtual pipe from the sender to the destination	Individual virtual pipes from each transmitting device to the target device
Advantages	Simple and flexible	Simple and flexible
Limitations	Duplicate messages performance degrades with communication holes	Performance degrades with communication holes

4 Simulation and Performance Analysis

4.1 Simulation Parameters

Performance of VBF and HH-VBF routing protocols are analyzed using simulations in Aqua-Sim [37]. Aqua-Sim is an extended version of NS-2 that efficiently simulates collision behaviors in large delay acoustic networks and also attenuation of underwater sensor networks. Aqua-Sim uses OTcl to model the protocol parameters implemented through C++ algorithms. The simulator creates a 3D environment of packet transmission in the network, which shows that the efficiency of both the protocols comes down with communication holes. Communication voids are created between the source and sink devices to measure QoS parameters in the overall network. Table 2 highlights the Simulation Parameters used to test the performance of the protocols in Aqua-Sim.

Table 2 Simulation parameters

Parameter name	Values
Simulator name	NS 2.35 + Aqua-Sim
Dimension of topology	1500 × 1500 × 1500 m
Transmission range	250 m
Antenna type	Omnidirectional
Data rate	50 kbps
Packet size	25–125 bytes
Number of nodes	100–600
Simulation time	300 s
Number of simulation runs	10
Protocols	VBF, HH-VBF

4.2 Performance Analysis

The performances of VBF and HH-VBF protocols were analyzed in a normal network scenario and in networks with communication void. Figure 4 shows the values of various parameters used in simulations. Four different QoS parameters were measured in the network.

- **Packet Delivery Ratio (PDR)**: PDR is the ratio of the overall packets arrived at the target device to the number of transmitted packets. Figure 4 shows the PDR in UWSN network scenario with varying number of nodes for the two protocols in normal network conditions and also with communication voids. Table 3 highlights the data values of VBF, VBF-Void, HH-VBF, and HH-VBF-Void. From the analysis, it is observed that both VBF and HH-VBF protocols give better performances in networks without communication voids. This proves that communication voids degrade the performance of even the latest LBORPs in UWSNs.
- **Average End-to-End Delay**: It is the time incurred by the information packet in reaching the target device from the sender device. Figure 5 highlights the average delay incurred by VBF, HH-VBF protocols in transferring data packets across the network with varying number of devices in normal and void environments. Table 4 highlights the data analysis of delay observed by routing protocols with void and under normal operating conditions. the data values, it is

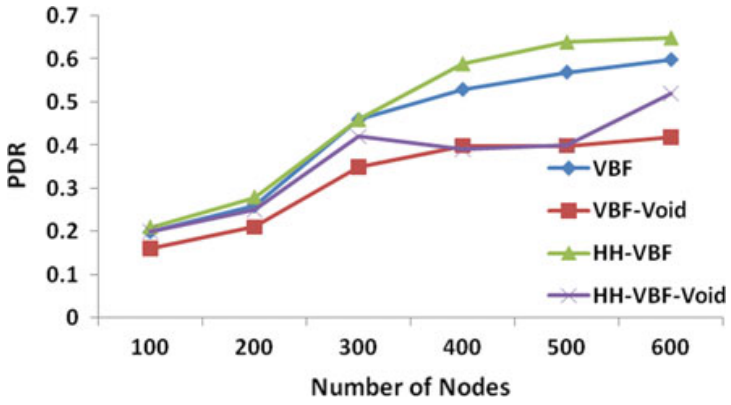


Fig. 4 PDR versus nodes

Table 3 PDR

No. of nodes	VBF	VBF-void	HH-VBF	HH-VBF-void
100	0.2	0.16	0.21	0.2
200	0.26	0.21	0.28	0.25
300	0.46	0.35	0.46	0.42
400	0.53	0.4	0.59	0.39
500	0.57	0.4	0.64	0.4
600	0.599	0.42	0.65	0.52

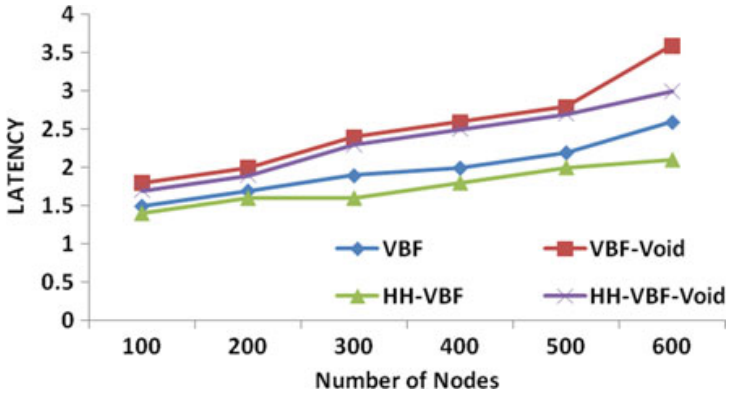


Fig. 5 Latency/End-To-End Delay versus nodes

Table 4 Average end-to-end delay

No. of nodes	VBF	VBF-Void	HH-VBF	HH-VBF-void
100	1.5	1.8	1.4	1.7
200	1.7	2	1.6	1.9
300	1.9	2.4	1.6	2.3
400	2	2.6	1.8	2.5
500	2.2	2.8	2	2.7
600	2.6	3.6	2.1	3

evident that the delay is less VBF and HH-VBF in normal environments as compared to networks with voids.

- **Throughput**: It is the rate or amount of data that is effectively transferred from the sender device to the target in a given time period. Figure 6 highlights the throughput of both VBF and HH-VBF protocols in normal and void environments. Table 5 enlists the data values of throughput. It is observed that HH-VBF protocol has the highest throughput compared to other routing protocols. Both the protocols suffer from degradation in throughput with voids in the network.
- **Energy Consumption**: In order to test the validity of the performance of the routing protocols, energy consumption is regarded as a significant parameter. Energy consumption is regarded as how much energy is consumed by

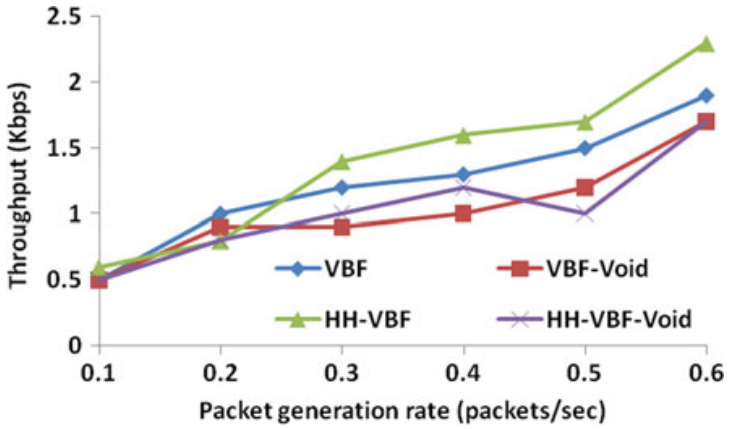


Fig. 6 Throughput versus packet generation rate

Table 5 Throughput

Packet generation rate	VBF	VBF-void	HH-VBF	HH-VBF-void
0.1	0.5	0.5	0.6	0.5
0.2	1	0.9	0.8	0.8
0.3	1.2	0.9	1.4	1
0.4	1.3	1	1.6	1.2
0.5	1.5	1.2	1.7	1
0.6	1.9	1.7	2.3	1.7

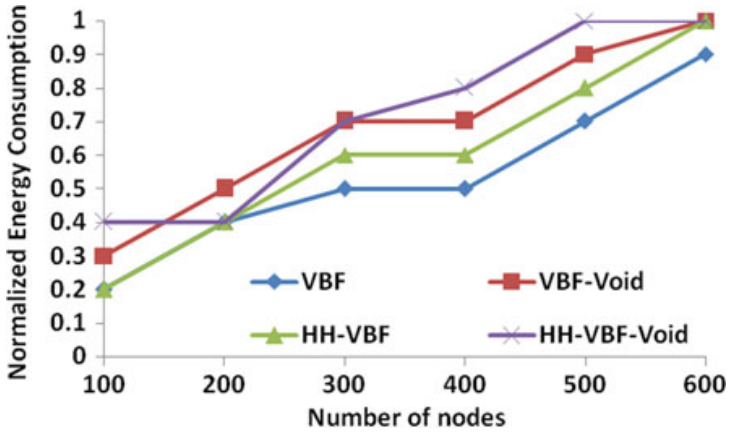


Fig. 7 Normalized energy consumption versus no. of nodes

Table 6 Normalized energy consumption

No. of nodes	VBF	VBF-void	HH-VBF	HH-VBF-void
100	0.2	0.3	0.2	0.4
200	0.4	0.5	0.4	0.4
300	0.5	0.7	0.6	0.7
400	0.5	0.7	0.6	0.8
500	0.7	0.9	0.8	1
600	0.9	1	1	1

transmitting nodes and intermediary nodes to make sure that the packet reaches the destination. Figure 7 elaborates the graphical analysis of the energy consumption of energy nodes in different protocols under UWSN operating scenarios. Table 6 also highlights the data values of protocols and it is observed that both protocols consume more energy in void scenarios.

5 Open Research Areas and Recommendations

In this section, open research problems generated by communications holes in UWSNs are discussed. Researchers designing routing protocols need to address the following issues to come up with optimal routing solutions for UWSN scenarios.

- Quality of Service (QoS): Communication voids drastically degrade the QoS offered by the routing protocols in UWSNs. The data delivery rate is reduced with increased data loss due to communication holes and void nodes. Delay in data transmission increases with frequent loss of data packets.
- Reliability: Reliable data delivery is a vital factor in the success of any routing protocol. Void nodes are often unable to find the next forwarder node in the network and are forced to drop the data packet. Thus, reliability is a big concern in UWSNs with communication voids.
- Scalability: In many UWSNs, the problems of communication holes occur when the network size is increased. With a reduced number of sensor nodes in a large geographical area, the chances are high for the occurrence of communication holes. Solutions to avoid communication voids with a minimum number of sensor nodes are to be found out.
- Mobility: Frequent mobility of sensor nodes is another reason for the communication voids. Good protocols should handle the mobility of nodes effectively for better performance.
- Energy: Loss of energy due to communication voids is another area of concern. With constrained battery and inability in frequent recharges, underwater sensor nodes need to be protected from unwanted drainage of energy.

6 Conclusion and Future Work

Underwater Sensor Networks have come out as a prominent research area in the networking domain with a wide range of aquatic applications. Routing in UWSNs is an area of major concern due to its unique features such as continuous node mobility, frequent disruption of links, and interference caused by other underwater acoustic systems such as marine mammals. Recently in UWSNs, opportunistic routing has been accepted as an efficient routing approach. The paper discussed the functioning of various Location-Based Opportunistic Routing Protocols proposed for UWSNs and evaluated the performance of two major protocols, VBF and HH-VBF using simulations in Aqua-Sim. The performances of these protocols are evaluated in normal network scenarios and also in networks with communication voids. We could observe that the performance of these protocols comes down with communication voids in the network. Insights from this article would enable researchers to work toward designing more effective protocols for underwater environments having voids. In the near future, a novel routing protocol, especially

designed for UWSN operating scenarios would be proposed and compared with existing routing protocols in terms of PDR, Energy, Delay, and Throughput and above all will give good performance in networks with communication voids and that are highly secure from all sorts of intruders and man-in-middle attacks.

References

1. Chen, Y., Jin, X., Xu, X.: Energy-efficient mobile data collection adopting node cooperation in an underwater acoustic sensor network. China Commun. **14**(6), 32–42 (2017)
2. Wang, Z., Han, G., Qin, H., Zhang, S., Sui, Y.: An energy-aware and void-avoidable routing protocol for underwater sensor networks. IEEE Access **6**, 7792–7801 (2018)
3. Akyildiz, I.F., Pompili, D., Melodia, T.: Underwater acoustic sensor networks: research challenges. Ad Hoc Netw. **3**(3), 257–279 (2005)
4. Akyildiz, I.F., Pompili, D., Melodia, T.: State-of-the-art in protocol research for underwater acoustic sensor networks. In: Proceedings of the 1st ACM International Workshop on Underwater networks—WUWNet '06 (2006)
5. Açar, G., Adams, A.: ACMENet: an underwater acoustic sensor network protocol for real-time environmental monitoring in coastal areas. IEE Proc. Radar Sonar Navig. **153**(4), 365 (2006)
6. Partan, J., Kurose, J., Levine, B.N.: A survey of practical issues in underwater networks. In: Proceedings of the 1st ACM International Workshop on Underwater Networks—WUWNet '06 (2006)
7. Biswas, S., Morris, R.: ExOR. ACM SIGCOMM Comput. Commun. Rev. **35**(4), 133 (2005)
8. Bruno, R., Conti, M., Nurchis, M.: Opportunistic packet scheduling and routing in wireless mesh networks. In: 2010 IFIP Wireless Days (2010)
9. Chakchouk, N.: A survey on opportunistic routing in wireless communication networks. IEEE Commun. Surv. Tutor. **17**(4), 2214–2241 (2015)
10. Nayyar, A., Batth, R.S., Ha, D.B., Sussendran, G.: Opportunistic networks: present scenario —a mirror review. Int. J. Commun. Netw. Inf. Secur. (IJCNIS) **10**(1), 223–241 (2018)
11. Menon, V.G., Prathap, P.M.: Comparative analysis of opportunistic routing protocols for underwater acoustic sensor networks. In: 2016 International Conference on Emerging Technological Trends (ICETT) (2016)
12. Menon, V.G.: Opportunistic Routing Protocols in Underwater Acoustic Sensor Networks: Issues, Challenges, and Future Directions. Magnetic Communications: From Theory to Practice, pp. 127–148. CRC Press, Boca Raton (2018)
13. Menon, V.G., Prathap, P.M.: Moving From Topology-Dependent to Opportunistic Routing Protocols in Dynamic Wireless Ad Hoc Networks: Challenges and Future Directions. Algorithms, Methods, and Applications in Mobile Computing and Communications, pp. 1–23. IGI Global, Hershey (2017)
14. Menon, V.G.: Analyzing the performance of random mobility models with opportunistic routing. Adv. Wirel. Mob. Commun. **10**(5), 1221–1226 (2017)
15. Han, M.K., Bhartia, A., Qiu, L., Rozner, E.: O3. In: Proceedings of the Twelfth ACM International Symposium on Mobile Ad Hoc Networking and Computing—MobiHoc '11 (2011)
16. Menon, V.G., Prathap, P.M.: Survey on latest energy based routing protocols for underwater wireless sensor networks. Int. J. Comput. Netw. Wirel. Commun. **6**(6), 52–55 (2017)
17. Ayaz, M., Abdullah, A., Faye, I., Batira, Y.: An efficient dynamic addressing based routing protocol for underwater wireless sensor networks. Comput. Commun. **35**(4), 475–486 (2012)

18. Yan, H., Shi, Z. J., Cui, J.: DBR: depth-based routing for underwater sensor networks. In: NETWORKING 2008 Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet, pp. 72–86 (2008)
19. Jafri, M.R., Sandhu, M.M., Latif, K., Khan, Z.A., Yasar, A.U., Javaid, N.: Towards delay-sensitive routing in underwater wireless sensor networks. Proc. Comput. Sci. **37**, 228–235 (2014)
20. Wang, C., Zhang, G., Zhang, L., Shao, Y.: Improvement research of underwater sensor network routing protocol HHVBF. In: 11th International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM 2015) (2015)
21. Xie, P., Cui, J., Lao, L.: VBF: vector-based forwarding protocol for underwater sensor networks. In: NETWORKING 2006. Networking Technologies, Services, and Protocols; Performance of Computer and Communication Networks; Mobile and Wireless Communications Systems (2006)
22. Ghoreyshi, S., Shahrabi, A., Boutaleb, T.: A novel cooperative opportunistic routing scheme for underwater sensor networks. Sensors **16**(3), 297 (2016)
23. Nicolaou, N., See, A., Xie, P., Cui, J., Maggiorini, D.: Improving the robustness of location-based routing for underwater sensor networks. In: OCEANS 2007—Europe, pp. 1–6. Aberdeen (2007)
24. Ghoreyshi, S.M., Shahrabi, A., Boutaleb, T.: Void-handling techniques for routing protocols in underwater sensor networks: survey and challenges. IEEE Commun. Surv. Tutor. **19**(2), 800–827 (2017)
25. Menon, V.G., Joe Prathap, P.M.: Opportunistic routing with virtual coordinates to handle communication voids in mobile ad hoc networks. In: Advances in Intelligent Systems and Computing, pp. 323–334 (2015)
26. Ghoreyshi, S.M., Shahrabi, A., Boutaleb, T.: An opportunistic void avoidance routing protocol for underwater sensor networks. In: 2016 IEEE 30th International Conference on Advanced Information Networking and Applications (AINA) (2016)
27. Menon, V.G., Prathap, P.M.: A review on efficient opportunistic forwarding techniques used to handle communication voids in underwater wireless sensor networks. Adv. Wirel. Mob. Commun. **10**(5), 1059–1066 (2017)
28. Darehshoorzadeh, A., Boukerche, A.: Underwater sensor networks: a new challenge for opportunistic routing protocols. IEEE Commun. Mag. **53**(11), 98–107 (2015)
29. Hwang, D., Kim, D.: DFR: directional flooding-based routing protocol for underwater sensor networks. In: OCEANS 2008 (2008)
30. Coutinho, R.W., Boukerche, A., Vieira, L.F., Loureiro, A.A.: GEDAR: geographic and opportunistic routing protocol with depth adjustment for mobile underwater sensor networks. In: 2014 IEEE International Conference on Communications (ICC) (2014)
31. Chen, Y.S., Juang, T.Y., Lin, Y.W., Tsai, I.C.: A low propagation delay multi-path routing protocol for underwater sensor networks. J. Internet Technol. **11**, 153–165 (2010)
32. Hao, K., Jin, Z., Shen, H., Wang, Y.: An efficient and reliable geographic routing protocol based on partial network coding for underwater sensor networks. Sensors **15**(6), 12720–12735 (2015)
33. Noh, Y., Lee, U., Wang, P., Choi, B.S., Gerla, M.: VAPR: void-aware pressure routing for underwater sensor networks. IEEE Trans. Mob. Comput. **12**(5), 895–908 (2013)
34. Jornet, J.M., Stojanovic, M., Zorzi, M.: Focused beam routing protocol for underwater acoustic networks. In: Proceedings of the Third ACM International Workshop on Wireless Network Testbeds, Experimental Evaluation and Characterization—WuWNeT '08 (2008)
35. Li, Z.L., Yao, N.M., Gao, Q.: Relative distance-based forwarding protocol for underwater wireless sensor networks. Appl. Mech. Mater. **437**, 655–658 (2013). https://doi.org/10.4028/www.scientific.net/amm.437.655

36. Nayyar, A., Puri, V., Le, D.N.: Comprehensive analysis of routing protocols surrounding underwater sensor networks (UWSNs). In: Data Management, Analytics and Innovation (pp. 435–450). Springer, Singapore (2019)
37. Nayyar, A., Singh, R.: A comprehensive review of simulation tools for wireless sensor networks (WSNs). J. Wirel. Netw. Commun. **5**(1), 19–47 (2015)

A Hybrid Optimization Algorithm for Pathfinding in Grid Environment

B. Booba, A. Prema and R. Renugadevi

Abstract Grid computing has been highly effective in the area of life sciences, financial analysis, research collaboration, and engineering. This paper is a study of existing algorithms like Swarm Intelligence (SI) algorithms such as Ant Colony Optimization (ACO), Particle Swarm Optimization (PSO), Artificial Bee Colony (ABC–PSO), and Parallel Particle Swarm Optimization (PPSO) to opt for the optimal path in a grid computing environment. These algorithms were used to solve the complex optimization problems in finding the path between source node to destination node effectively. Nature computing techniques based on the study of the collective behavior of ants, particle swarms, and bees are used to find the optimal path, improve the optimization methods and scalability in a set of representative problems. The hybridization of a grid computing environment and nature-inspired computing algorithms such as ACO, PSO, ABC–PSO, and PPSO has resulted in a class of solutions that differ in structure and design from the peer-to-peer network algorithms and the evaluated results showed the effectiveness of the pathfinding problem. ACO is implemented on a dynamic grid computing environment to demonstrate scalability and a solution for pathfinding. A class of four algorithms is used to find an optimal path and improve the optimization methods and shorten the computational time in a grid computing environment.

Keywords ACO · PSO · ABC–PSO · PPSO · Grid computing · Nature-inspired computing · Pathfinding · Swarm intelligence

B. Booba (✉) · A. Prema
Department of Information Technology, School of Computing Science, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Pallavaram, Chennai, India
e-mail: boobarajashekar@gmail.com

A. Prema
e-mail: unjanai@gmail.com

R. Renugadevi
Department of Computer Science, School of Computing Science, Vels Institute of Science, Technology and Advanced Studies (VISTAS), Pallavaram, Chennai, India
e-mail: nicrdevi@gmail.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_50

1 Introduction

1.1 Overview of Grid Computing

Grid computing is a platform for solving computational problems in science, engineering, commerce, and other fields. Numerous solutions to grid computing are being developed. Grid computing has been proven to be highly effective in the area of life sciences, financial analysis, research collaboration, and engineering [1].

Grid computing is the accumulation of computer resources to achieve a common goal. The grid provides location-transparent scheduling that is a measure of registering resources required at a determined time interval for computation. Grid capacity has to make more cost-effective utilization of a given measure of computer resources as an approach to tackle problems that cannot be approached without a huge measure of computing power. The fact that it utilizes the resources of numerous workstations might be helpful synergistically to harness and manage as cooperation to a typical destination [2].

1.2 Problem Definition

The objectives are to find the grid computing based-pathfinding, identifying the position of the obstacles being developed on the grid computing environment and comparing, analyzing the performance results of the class of four algorithms in nature-inspired computing. This paper focuses on motion planning for "pathfinding". It addresses the optimal pathfinding problem in a grid computing environment [3]. This paper focuses on the use of SI algorithms such as ACO, PSO, PPSO, and hybrid ABC–PSO to find the optimization problems for selecting the shortest path in a grid computing environment. These algorithms utilize the swarm intelligence in grid environment to obtain an optimal solution over an iterative process. The optimal solutions are evaluated using different kinds of aspects, the hybrid ABC–PSO has been evaluated against other SI algorithms. A conclusion has been shown that the hybrid ABC–PSO algorithm performs better in terms of intensification and diversification concepts. This paper is to demonstrate robustness, scalability of the nature-inspired pathfinding algorithms in a customized grid computing environment.

1.3 Proposed Work

The pathfinding algorithms are analyzed to stumble on the shortest path in grid computing milieu from source node to destination node even if there are any

obstacles. This paper focuses on the use of ACO, ABC–PSO, and PPSO for selecting the shortest path in a grid computing environment [4].

The following approaches are used in the proposed work:

- Election of shortest path using ACO algorithm.
- Selection of shortest path by means of Parallel Particle Swarm Optimization (PPSO) algorithm.

The above mentioned approaches were used to solve the complex optimization problem in pathfinding between source nodes to destination node effectively.

2 Related Works

ACO methods are useful in reducing the problems of finding paths to a goal state. I showed the behavior of real ants, and to provide heuristic solutions for optimization problem. Cellular-DPSABC algorithm is able to apply dynamic optimization problems. The principle of this algorithm was implemented in PSO to fitness value population in Artificial Bee Colony Algorithm (ABC). This projected algorithm can be applied for solving optimization problems. The new algorithm is simple and flexible swarm-based one. It is also robust, it was established on a limited set of test problems [5].

An abstraction mechanism which maintains memory competence using a large-scale grid. It is also effective in reducing planning costs. This mechanism called a Minimal-Memory (MM) abstraction distinguishes the abstraction representation of the techniques which use the abstraction for pathfinding. Memory-efficient way of precomputing subgoals reduces the main obstacle in applying real time search systems in video games. Bee Colony Optimization (BCO) algorithm is focused on obstacle prevention and shortest safe path from the source to the target. Field D* can be used to approximate a distance find on the nodes of a complex, and the distance field can be used to weight the complex [6].

3 Contribution of the Paper

3.1 Design of ACO Algorithm for Pathfinding

ACO is a streamlining algorithm in light of the activities of strategies which are to diminishing the issues of finding the paths to a goal state. ACO basically is an artificial system that takes motivation from the conduct of genuine states. It locates the most efficient routes from their nests to food sources [7]. The attribute of ACO algorithms is the permutation of the structure of a solution. ACO algorithms are

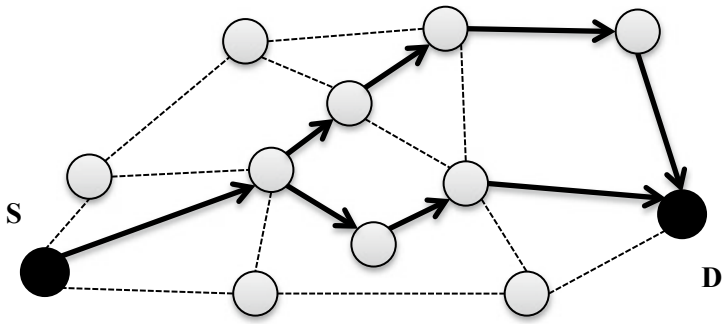


Fig. 1 Grid pathfinding

capable of solving minimum cost path problems on more complicated graphs as shown in Fig. 1.

ACO is capable of solving the large computational problem in gird computing [8]. When all the jobs are assigned, the resources are allocated by the pheromone trails. The pheromone trail can be proposed from the source node while traveling through several nodes to the target node.

The trails are updated by

$$\tau_{ab} \leftarrow (1-p)\tau_{ab} + \sum_{m} \Delta_{\tau_{ab}}^{m} \tag{1}$$

Equation (1) is the amount of pheromone deposited for a state transition, *ab* is the pheromone evaporation coefficient and τ_{ab} is the amount of pheromone deposited by *m*th system [9].

3.2 Design of ABC–PSO Algorithm

The benefit of ABC over other swarm intelligence approaches is that in the ABC algorithm the imaginable solutions indicate food sources. A total number of employed bees are identical to the number of nourishment sources, with each working honey bee speaking to a sustenance source [10]. Nourishment source signifies an answer for the issue that is fundamental to be upgraded. Every sustenance source is produced as pursues

$$X_{ij} = X_{\min j} + \text{rand}[0,1]\left(X_{\min j} - X_{\min j}\right) \tag{2}$$

ABC–PSO, two algorithms' components smoothly with the end goal to pick up profit by their specific qualities Fig. 2. The incorporating PSO with ABC is to build goal of the pbests of the particles. The algorithm with segments from both ABC and PSO is shaped with the end goal to have a calculation that easily tackles distinguishable issues as while having iteratively invariant conduct as PSO in the meantime [11].

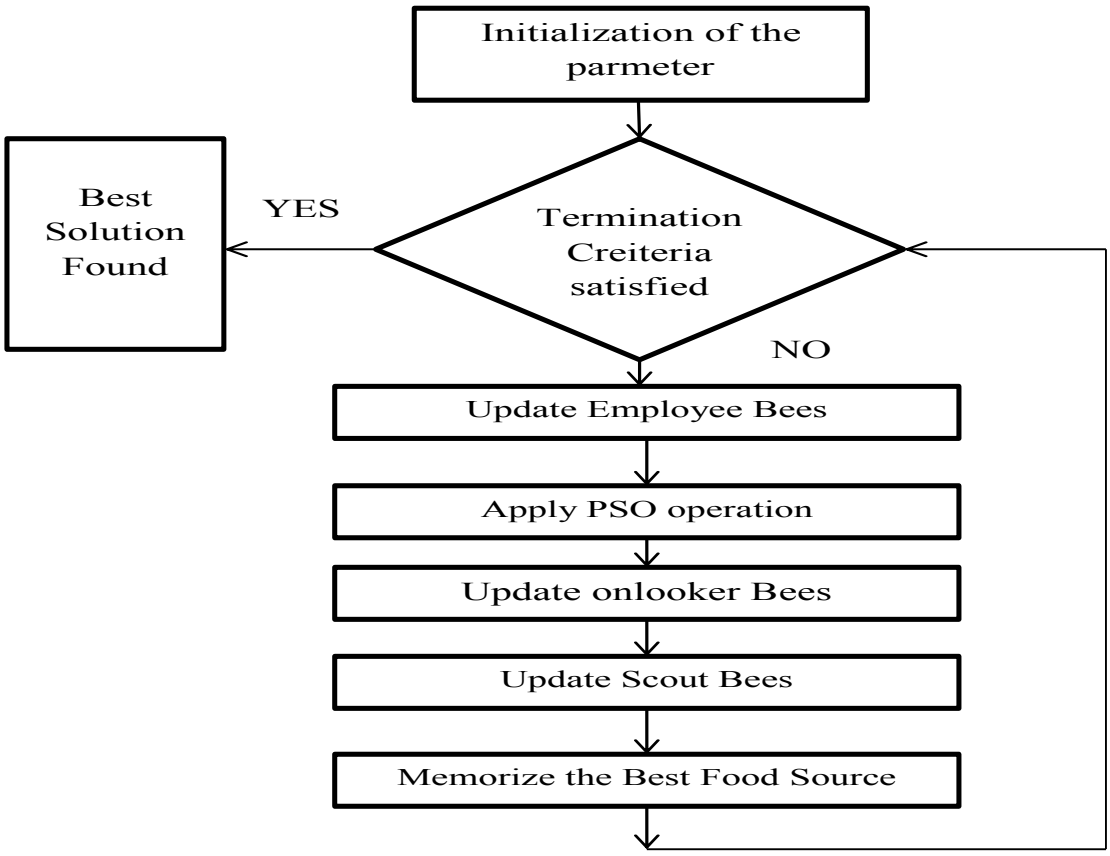


Fig. 2 ABC–PSO algorithm

For every molecule i in the swarm, the ABC calculation method refresh condition is down to business to the individual best $pbest_i$. This is done after iteratively choosing another molecule k, arbitrary issue variable j, hence c$pbest_i$ is refreshed as pursues

$$pbest_{ij} = pbest_{ij} + \phi_{ij}\, X\left(pbest_{ij} - pbest_{kj}\right) \tag{3}$$

3.3 Design of PPSO Algorithm

Parallel processing is to generate the same results and feasibly uses multiple processors to reduce the run time. Each time, all particles are independent it can be easily analyzed in parallel [12]. The PPSO algorithm solves complex functional optimization problems. The particles in PPSO algorithm are attracted to the best locations in the search space [13]. The process of PPSO group is shown in Fig. 3.

The above technique gives the worldwide best and the neighborhood best arrangements [14]. It builds the shot of particles to find a superior arrangement and to expand the nearby ideal arrangements.

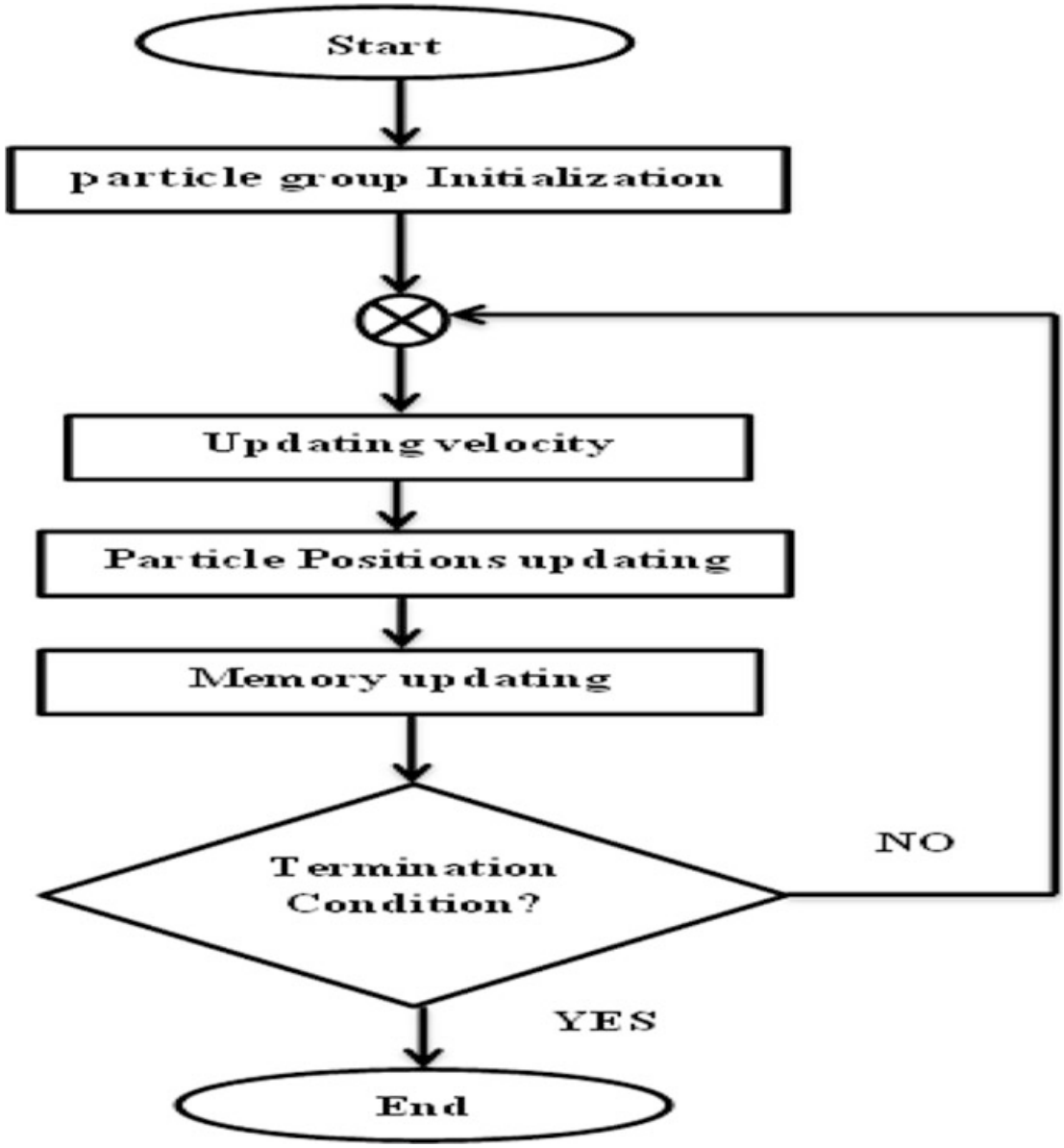


Fig. 3 The process of parallel particle swarm optimization/group

4 Simulation Results

In this work, MATLAB tool is used to build the simulation environment and the proposed techniques are simulated to evaluate the performance of Make span and to find an optimal path [15].

$$\text{Make span} = \text{Completion time of last job} - \text{Starting time of first job}$$

To evaluate the effectiveness of these algorithms, the average deviation of the present best result from the optimum in all iterations is determined. The yield of the dynamic condition gives the premise on which most limited ways can be figured between any two hubs utilizing pheromone esteems at every one of their neighbor hubs [16, 17]. The system show utilized in unique is made out of the hubs and connections that are viewed as bidirectional. It very well may be finished by dissecting the likelihood make length esteems [18].

5 Performance Analysis

The Proposed technique is replicated for various combinations of grid sizes such as 128 × 8, 256 × 16, 512 × 32, and 1024 × 64. The results are shown in Table 1. It demonstrates the performance of the ACO, PSO, ABC–PSO, and PPSO algorithms. This performance measure produces a fast rate of better optimization make span [19].

Figure 4 defines a method to find the solution for the problem of resource with total execution times which equal to the make span. According to this, the optimal solution for each algorithm is found based on the global pheromone.

Table 1 Probability make span of ACO, PSO, ABC–PSO, and PPSO

Probabilistic make span of ACO		Probabilistic make span of PSO		Probabilistic make span of ABC–PSO		Probabilistic make span of PPSO	
No. of iteration in matrice	Value	No. of iteration in matrice	Value	No. of iteration in matrice	Value	No. of iteration in matrice	Value
[0][0]	818	[0][0]	799	[0][0]	470	[0][0]	580
[0][1]	652	[0][1]	550	[0][1]	320	[0][1]	520
[0][2]	563	[0][2]	490	[0][2]	290	[0][2]	450
[0][3]	506	[0][3]	476	[0][3]	281	[0][3]	410
[1][0]	491	[1][0]	359	[1][0]	260	[1][0]	310
[1][1]	404	[1][1]	350	[1][1]	251	[1][1]	280
[1][2]	348	[1][2]	332	[1][2]	234	[1][2]	290
[1][3]	343	[1][3]	312	[1][3]	219	[1][3]	260
[2][0]	309	[2][0]	275	[2][0]	170	[2][0]	230
[2][1]	280	[2][1]	230	[2][1]	140	[2][1]	190
[2][2]	176	[2][2]	141	[2][2]	121	[2][2]	160
[2][3]	163	[2][3]	132	[2][3]	76	[2][3]	99

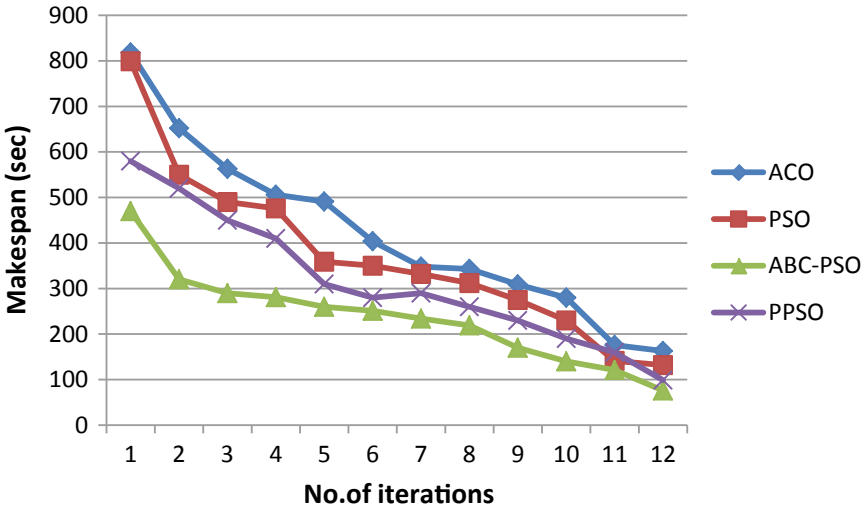


Fig. 4 ACO, PSO, ABC–PSO, and PPSO performance

6 Conclusion

This work reports the application of ACO for pathfinding in a simulated grid computing environment which is suitable for comparison with PSO, ABC–PSO, and PPSO algorithms. The hybridization of grid computing environment and swarm intelligence algorithms such as ACO, PSO, ABC–PSO, and PPSO has resulted in a class of solutions that differ in structure and design from the peer-to-peer network algorithms. The distinction between ant colony and bee colony is also studied through simulation studies in MATLAB. ACO is implemented on a dynamic grid computing environment to demonstrate the scalability of the proposed solution for pathfinding. For the future, this can be implemented in a cloud environment to avoid obstacles freely and improve the scalability.

References

1. Dorigo, M., Stutzle, T.: Ant colony optimization: overview and recent advances. In: Gendreau, M., Potvin, Y. (eds.) Handbook of Metaheuristics, 2nd edition, vol. 146 in International Series in Operations Research & Management Science, pp. 227–263. Springer, Verlag, New York (2010)
2. Reshamwala, A., Vinchurkar, D.P.: Robot path planning using an ant colony optimization approach: a survey. Int. J. Adv. Res. Artif. Intell. **2**, 65–71 (2013)
3. Zafarani-Moattar, E., Haj-Seyed-Javadi, H., Feizi-Derakhshi, M.R.: Parallel reverse niching PSO for multimodal optimization. In: International Conference on Machine Learning, Electrical and Mechanical Engineering (ICMLEME'2014) (2014)
4. Zhang, J.-R., Zhang, J., Lok, T.-M., Lyu, M.R.: A hybrid particle swarm optimization–back-propagation algorithm for feed forward neural network training. Appl. Math. Comput. (2007)

5. Li, M., Duan, H.: Hybrid artificial bee colony and particle swarm optimization approach to protein secondary structure prediction. In: Proceeding of the 10th World Congress on Intelligent Control and Automation (2012)
6. Sood, M., Kaur, M.: Shortest path finding in country using hybrid approach of BBO and BCO. Int. J. Comput. Appl. **40**(6), 0975–8887 (2012)
7. Mundra, P.S.: Ant colony optimization: a technique used for finding shortest path. Int. J. Eng. Innov. Technol. **1**(5) (2012)
8. Baktash, N., Mahmoudi, F., Meybodi, M.R.: Cellular PSO-ABC: a new hybrid model for dynamic environment. Int. J. Comput. Theory Eng. **4**(3) (2012)
9. Kumar, S., Sharma, V.K., Kumari,R.: Comparative study of hybrids of artificial bee colony algorithm. Int. J. Inf. Commun. Comput. Technol. **1**(2) (2013)
10. Uras, T., Koenig, S., Hernandez, C.: Subgoal graphs for optimal pathfinding in eight-neighbor grids. Association for the Advancement of Artificial Intelligence (2014)
11. Bulitko, V., Björnsson, Y., Lawrence, R.: Case-based subgoaling in real-time heuristic search for video game pathfinding. J. Artif. Intell. Res. **39**, 269–300 (2010)
12. Dorigo, M.: The ant colony optimization metaheuristic: algorithms, applications, and advances. International Series in Operations Research & Management Science (2003)
13. Ameur, M.S.B., Sakly, A., Mtibaa, A.: Implementation of real coded PSO algorithms using FPGA technology. In: 2014 15th International Conference on Sciences and Techniques of Automatic Control and Computer Engineering (STA) (2014)
14. Sadhasivam, G.S., Meenakshi, D.K.: Load balanced, efficient scheduling with parallel job submission in computational grids using parallel particle swarm optimization. In: 2009 World Congress on Nature & Biologically Inspired Computing (2009)
15. Bentley, P.J.: Perceptive particle swarm optimization: an investigation. In: Proceedings 2005 IEEE Swarm Intelligence Symposium 2005 SIS 2005, 2005 Publication Chu Shu-Chuan. Parallel Particle Swarm Optimization Algorithms with Adaptive Simulated Annealing, Studies in Computational Intelligence (2006)
16. Chu, S.-C.: Parallel Particle swarm optimization algorithms with adaptive simulated annealing. Studies in Computational Intelligence (2006)
17. Booba, B., Gopal, T.V.: Comparison of ant colony optimization & particle swarm optimization in grid scheduling. Aust. J. Basic Appl. Sci. **8**(7) (2014)
18. Booba, B., Gopal, T.V.: Efficient scheduling of packets in wireless sensor networks using priority based scheduling approach. J. Comput. Sci. (2014)
19. Booba, B., Gopal, T.V.: An efficient distributed computing technique for job scheduling is accepted by the journal. Int. J. Appl. Environ. Sci. (IJAES)

Dynamic Hashtag Interactions and Recommendations: An Implementation Using Apache Spark Streaming and GraphX

Sonam Sharma

Abstract Hashtag, started with Twitter is a keyword with prefix "#" and now being used mostly for all communication on social media. It has been identified as very powerful and effective in organizing communications according to the topic and trend. Hashtag can further help on various analysis, as it links users with their topic of interests. Hashtag aids in building communities of similar interests. With hashtags, we can follow current trend and interest on twitter which can help us in analyzing multiple factors, e.g., sensitivity of the ongoing trend, its spread, people getting affected, its effect on business and so on. Traditionally available approaches help us in analyzing batch data and finding interests and trends on it. Now with the advancements in the field of technology helps us in analyzing a large amount of online data within seconds. In this paper, we will be exploring dynamic hashtag interactions to find correlations among them and propose a methodology which can successfully find relevant hashtags based on the interest in focus. We will propose our methodology of analyzing and exploring tweets in real time with the extent of converting information; we are getting from twitter to knowledge.

Keywords Co-occurrence graph · Community networks · Hashtags · Graph mining · Real-time analytics · Streaming graph · Social media

1 Introduction

GRAPH structures have many application areas named cybersecurity, social network, e-commerce websites, etc. With the abundant data getting generated in wide variety, with different sets of volumes and rates. The increased usage and connectedness of social media have given rise to trends and interests which can be triggered for various reasons. Analyzing social media data whether in batch or in real time with this freely available interesting public information could yield

S. Sharma (✉)
Tata Consultancy Services, Noida, India
e-mail: sonam.sharma2@tcs.com

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8_51

interesting results. People on social media have been expressing their feelings freely, which makes Twitter an ideal source for accumulating wide amount of data for real-time on-time analysis on the vast amount of opinions and on a wide range of topics. Although this information can be easily collected, getting the information from it and representing it in good form has its challenges. These challenges withstand even with the advancement of tools and technologies. Additionally, analyzing twitter data is indeed a challenging problem due to the nature, diversity, and volume of the data.

Whenever we talk about social media, the combination of graphical data and big data is inevitable and is the reason it is popularly known as Social Network Analysis (SNA). There are various types of graph structures that are possible where the most popular ones are user-followers graph, user-mentions graph, etc. But here we are basically building and analyzing the graphs based on hashtags for a particular trend,which can further help us in analyzing the trending hashtags and other related hashtags. We are focusing on open-source software technologies named Spark which provides a graphical layer on top of it.

Data Streams have gain popularity in data mining techniques, where time is the most relevant factor since such data helps in providing *on-time* analysis and focus on *now*. But it is a very challenging task in various aspects [1]:

1. Dynamic nature of the data streams can cause data distribution changes with time (called concept drift) [2]
2. It not feasible to store the data since the data is in infinite volume
3. The analysis of data streams must happen as quickly as possible so as to get quick insights and operate in real time.

We can also see that when it comes to real-time analytics, many canonical analytics approaches may not be suitable for handling real-time data and providing *on-time* analytics. Here, we require a knowledge-based system for social network analysis which can handle streamed data and find interesting patterns on it and deliver the results within seconds [3, 4].

It is expected that the new approaches developed for extracting insights from real-time data will help in advancing the development of more sophisticated analytics allows better decision-making. This will help in gaining insights from your data as you receive with improved decision-making. This will help customers in quick and immediate decisions based on the scenario that is happening now, not what happened weeks, or months ago. Most of the big data that we receive are in real time and it is most valuable at the time of arrival (Fig. 1).

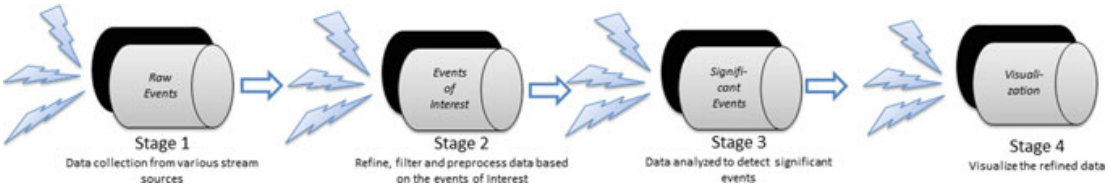


Fig. 1 A canonical real-time event processing pipeline

In any canonical real-time event processing pipeline at every stage, we polish for further step and at every stage we try to refine the data in a way to get new insights from it which can further help end users for decision-making. This streaming data can help in building final products which provide *actionable on-time insights*. The initial stage is to collect data from the various stream sources which is the first stage of any data processing pipeline. The main challenge with stream data preprocessing is that new data continuously arrives and we need to apply fully automated preprocessing methods. Also, we need to be able to update the methods with evolving data and optimize model parameters to deal with rapidly changing data. Streaming data needs to be fully automated with the required preprocessed methods.

In this paper, we present various methodologies on a linguistic process for building knowledge-based social media framework based on hashtag interactions. The framework works seamless with upcoming big data technologies and software. The state-of-the-art methodologies in the data analytic techniques on social media data are provided.

This paper is divided into the following sections: Sect. 2 discusses literature review and critical research gaps highlighted in this paper. In Sect. 3, the detailed problem statement for our analysis is defined. The proposed methodology is stated in Sect. 4. The results and conclusions are explained in Sects. 5 and 6 [5].

2 Literature Review

Activity on social media platform has become a part of daily activity of users because of which it has become one of the most powerful mediums for many researchers and enterprises to perform extensive analysis and derive interesting results. In this section a detailed study about various techniques on knowledge discovery using hashtags on social media. Although the problem of knowledge mining from twitter has been studied extensively during recent years and some we discuss below. Wei Wu designed a system for automatically generating personalized annotation tags to label Twitter user's interests and concerns [6]. However, user-level tags have its limitations as it only focuses on the particular user and their content and Twitter being a very large community, where we can focus on communities or on a large set of people which can help us in building strong knowledge-based system. Other such kinds of implementations have been performed on key-phrase extraction which determines important phrases from single documents or an entire collection of text documents [7, 8]. These approaches are not immediately applicable on Twitter because even if Twitter offers textual information but canonical NLP techniques cannot work on it as people on twitter tend not to follow basic norms of grammar on which we can do language processing. Apart from this, if we focus on single tweet for analysis we will most of the information which is being generated on Twitter.

However, current explorations being dined on social media are more matured than ever before and still it is one of the most chosen topics for researchers because

of its freely available data and whenever there is twitter, hashtag plays the most significant role in tweets and now people have started to understand this and follow it across various social media platforms. It has been identified that hashtag inclusion in the posts are increasing nowadays, because of which hashtags act as the fundamental information facet associated with social network messages [9]. Made a significant contribution by observing that hashtag recommendations aim to encourage the user to use hashtags more often and it is better to use appropriate hashtags and also avoiding synonymous in hashtags. All these observations indicate that the hashtag recommendation is not only interesting but an important problem statement.

Also, most of these existing solutions suffer from certain limitations. One problem is that the majority of approaches are limited to centralized environments. The proposed techniques are very time-consuming and take many computational resources. Most of the available solutions are neither sufficient nor suitable for building total knowledge analytics system on Twitter, since there is a huge mismatch between their processing capabilities and when we get the output from them, we lose the value (the fifth V of big data) of the data since that is not o*n-time.*

Thus, in all, it can be stated that they are solving the purpose of providing knowledge to the user but all the knowledge that is provided is either insufficient for the business users to go with or not on-time. Most of the methodologies that are developed on Twitter are on batch data. With the advent of the technologies, we can modify or build these methods on the top of stream data to help the business in providing on-time and social media presence using this.

3 Problem Statement

(a) *Proliferation of trends on social media*

Trends in social networks are elements of communication which helps in analyzing the popularity of a particular event (news) on social media. The more popular the event is, the more the occurrence of interactions on social media is. On trending events, people express their opinions which further helps them to build community and helps in the spread or proliferate the trend.

Most often hashtags help in the spread of trend and there are various combinations of hashtags which help people in expressing their thoughts clearly and be able to connect to the like-minded people. Sometimes trends can be misleading and biased too as it is basically the opinions which people are expressing what they feel and their personal thoughts. Understanding the dynamics of trend spreading could be of the great value of information for social media research areas.

(b) *Objective*

The objective of this paper is to propose a methodology to analyze hashtag interactions dynamically and graphically detect current ongoing trends on Twitter by building real-time hashtag co-occurrence graph on the top of real-time streaming engine. Real-time co-occurrence graph can help in analyzing popular hashtags. For this study, we have used big data technologies like *Spark Streaming* and *GraphX*. Graph visualization libraries like *graphstream* are used to visualize the graph in real time. Only the activities of Twitter after the connection are made. We analyze the streaming pipeline for each sliding window and perform the analysis. Under such circumstances, it becomes essential to understand each and every aspect of the data beforehand and constantly update the algorithm so as to deal with a future discrepancy.

4 Analytical Methodology for *On-Time* Twitter Analysis Hashtag Recommendations

How should we decide on managing data across an enterprise or an organization which is streamed from multiple sources like various kind of sensors, security data in the form of video from webcam, data from IDS/IPS, telematics data, and so on. Apart from management of data which is getting generated in real time, the other most important task is to analyze and find insights from it the moment it gets generated. Properly managing the data what is the best possible way to find insights from it as soon as it gets generated, so that there is no loss of data value and take necessary "*on-time*" actions on it [10].

Whenever a tweet is being twitted on twitter which contains keyword that starts with '#' called hashtags. Hashtags mostly contain current trends or hot topics in discussion. A single hashtag has the capability to describe the whole tweet. And if there are multiple hashtags available in the same tweet it can help in analyzing and getting interesting results. Multiple hashtags can further help us in analyzing user of the tweet. Multiple hashtags are getting created on Twitter and Trending Hashtags getting retweeted which helps in spread of the news faster not just on Twitter platform, but people use these hashtags on other social media sites to increase awareness. Hashtags refer to specific topics related to particular ongoing event or trend, but only a few of them become too exclusive and popular. People use hashtags to show their interest in a particular topic. Hence, a hashtag can provide valuable information about the topic, its evolution, and what user community of the analyzed hashtag is interested in through time.

Experiments and their results with real-time twitter data and analysis being done here provide real-time recommendations on what hashtag to follow next based on the topics which can help in building the related hashtags network and organization of trend. Finally, we discuss techniques that can describe the way hashtags, related to the same event, evolve through time and what are the most recommended

Table 1 Sample keywords to be crawled from twitter

Tradewar	Tradewars	Trump	trade

Table 2 Combination of hashtags considered

US AND trade	US AND China	Trump AND tradewar	China AND tradewar

hashtags. Keeping this in mind, we now explain the procedure which we followed for our experiment.

(A) *Data*

Tables 1 and 2 have listed the keywords used to crawl the data from Twitter. Where Table 1 shows specific keyword whereas Table 2 shows combination of these keywords to further enhance our analysis.

We have built our analysis from July 16 to Sept 12 2018. And we will be showcasing the results for the 6-min window in spark streaming. We have decided this window based on the amount of data and cluster size so that we can get seamless and fruitful results. These hashtags will be the *Focus Point* for our analysis and will further be enhanced by building the real-time graph and getting related Hash-tags on it.

(B) *Real-time Tweet Statistical Analysis and Data-set Preparation*

The next steps we perform will further provide ongoing real-time analysis on the data. Since our analysis majorly focused on hashtags, so we have filtered hashtags and build our analysis on top of it with some other information like users who have used this hashtag and number of tweets, this hashtag has been used shown in Fig. 2, which gives us an understanding that how important the analyzed topic used by the

Fig. 2 A *Map()* collection displaying the Hashtag -> (#unique tweets, #unique users) for some specific window

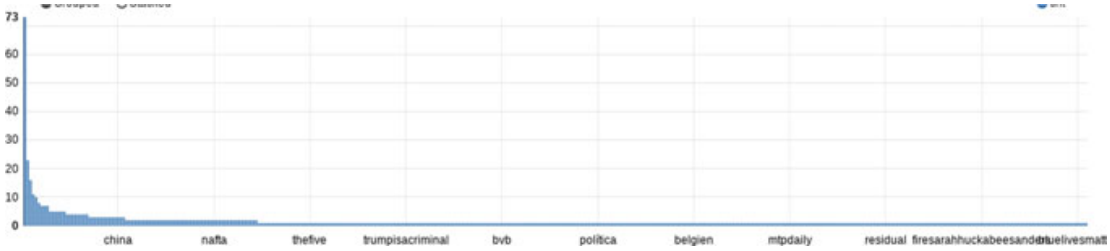


Fig. 3 Bar chart representation of hashtag distribution on some *window()* in Spark Streaming

Fig. 4 Classes of hashtags at a particular time window

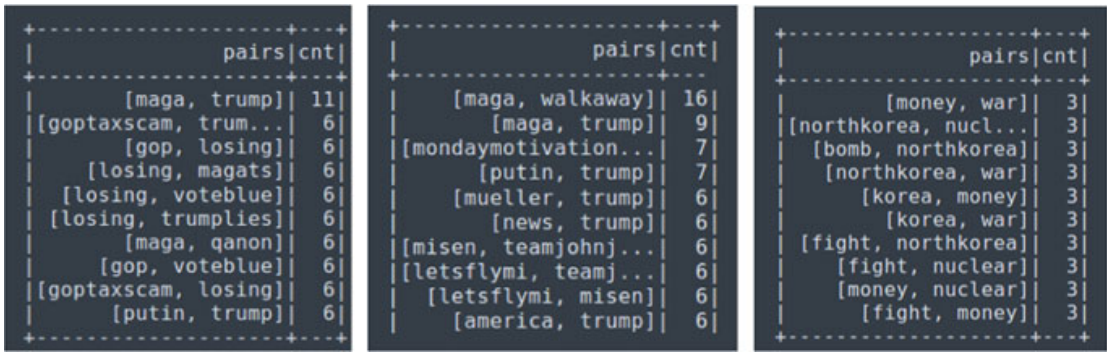
hashtag is for this particular window of stream. In this section, our main goal is to get an aware of the streaming twitter data from different angles to get the full view (Figs. 3, 4).

For further analysis, we have to focus on major top hashtags and their co-occurrences on tweet. And for this analysis, we have filtered the tweets which contain no hashtags even when they can help us in analyzing the content more but these tweets do not fit our use case.

5 Hashtag Interaction and Recommendation

Here, we will look at the major top hashtags and their co-occurrences. For this analysis, we have filtered the tweets which have no hashtags available with them (Fig. 5) [11, 12].

While analyzing the tweets for multiple *window(),* we have identified that some hashtags which are particularly in trend and popularly occur together or individually for a specific period of time. This makes them most relevant among all. We have also added pairs count in the Edge property of the graph for its use in further analysis.



pairs	cnt
[maga, trump]	11
[goptaxscam, trum...]	6
[gop, losing]	6
[losing, magats]	6
[losing, voteblue]	6
[losing, trumplies]	6
[maga, qanon]	6
[gop, voteblue]	6
[goptaxscam, losing]	6
[putin, trump]	6

pairs	cnt
[maga, walkaway]	16
[maga, trump]	9
[mondaymotivation...]	7
[putin, trump]	7
[mueller, trump]	6
[news, trump]	6
[misen, teamjohnj...]	6
[letsflymi, teamj...]	6
[letsflymi, misen]	6
[america, trump]	6

pairs	cnt
[money, war]	3
[northkorea, nucl...]	3
[bomb, northkorea]	3
[northkorea, war]	3
[korea, money]	3
[korea, war]	3
[fight, northkorea]	3
[fight, nuclear]	3
[money, nuclear]	3
[fight, money]	3

Fig. 5 Shows the top hashtag co-occurrences at different time *window()*

6 Architecture Flow of the Analysis

The streaming data we have used for our analysis and provided use case is taken from Twitter using Twitter Streaming API. Figure 6 shows the components we have used for this analysis. Specific to this use case the architecture is kept understandable with least and default configurations. The graph visualization uses *graphstream* library, which is both easy to integrate and visualize the graph in Spark and can help in streaming too if we will be using bigger batched in spark streaming window, and the graph which is getting visualized is small enough to fit the window. With multiple tweaks between the window sizes in *graphstream,* we were able to successfully visualize the graph [13, 14].

Kafka which is a distributed stream processing engine helps us in getting the streaming data and can also act as buffering the data for specified configuration and all the analyses of the data have been done on *Spark Streaming*.

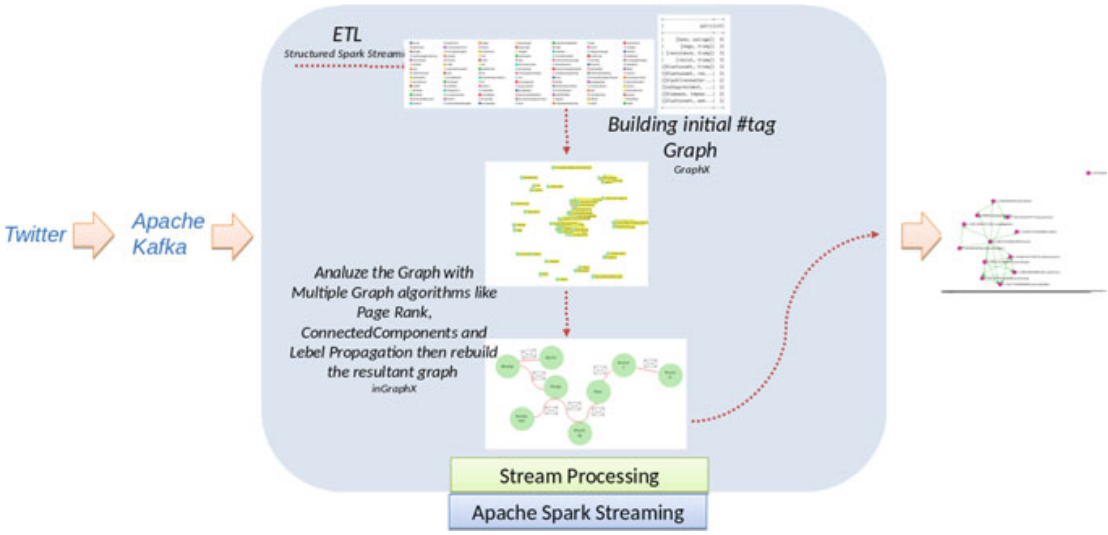


Fig. 6 Architecture defining the flow of analysis

7 Algorithms and Results

In this section, we describe the graph algorithms that we have used to build final hashtag co-occurrence graph for the specific batch window in Spark Streaming.

(a) ***Preliminaries***

(i) For any graph to be built, we have to preprocess and make dataset ready in the desired format for GraphX. For GraphX vertices, we will find *#unique* hashtags and provide them the *id* with *Long* datatype, since GraphX only understands long as a datatype for ID. For this, we have encoded the hashtag string using *MD5 hashing* algorithm so that we get unique id for each hashtag string. We also have to verify that if all hashtags have got different ids since no hashing algorithm can perfectly encode all the strings uniquely. This will make out Vertex RDD in the form of

$$\texttt{vertexRDD} : \texttt{RDD}[(\texttt{Long}, (\texttt{String}))]$$

(ii) Another preliminary component to build a graph is edges. For this, we will create hashtag *combinations* for each tweet and filter out empty tweets. After creating the *combinations,* we will *flatMap* them and create edge RDD by encoding the hashtags in a similar manner as vertices.

$$\texttt{edgeRDD} : \texttt{RDD}[\texttt{Edge}[\texttt{Int}]]$$

With above vertex and edge RDD's, we will build a graph (Fig. 7)

$$\texttt{Graph}[(\texttt{String}), \texttt{Int}]$$

(b) ***Applying Algorithms***

After building the graph we will perform the following analysis to get the final hashtag co-occurrence graph.

(i) To understand structural property of a graph we apply GraphX *Graph* class provides a number of methods to explore the graph.

(a) **Connected Components**

This is the method which can help us in understanding the graph on its connectedness. To describe this, we can say that a graph is called connected when any vertex in that graph can reach another vertex by following a sequence of edges from one vertex to another. If not so, then a graph can have multiple connected components which can act as a connected subgraph (Fig. 8).

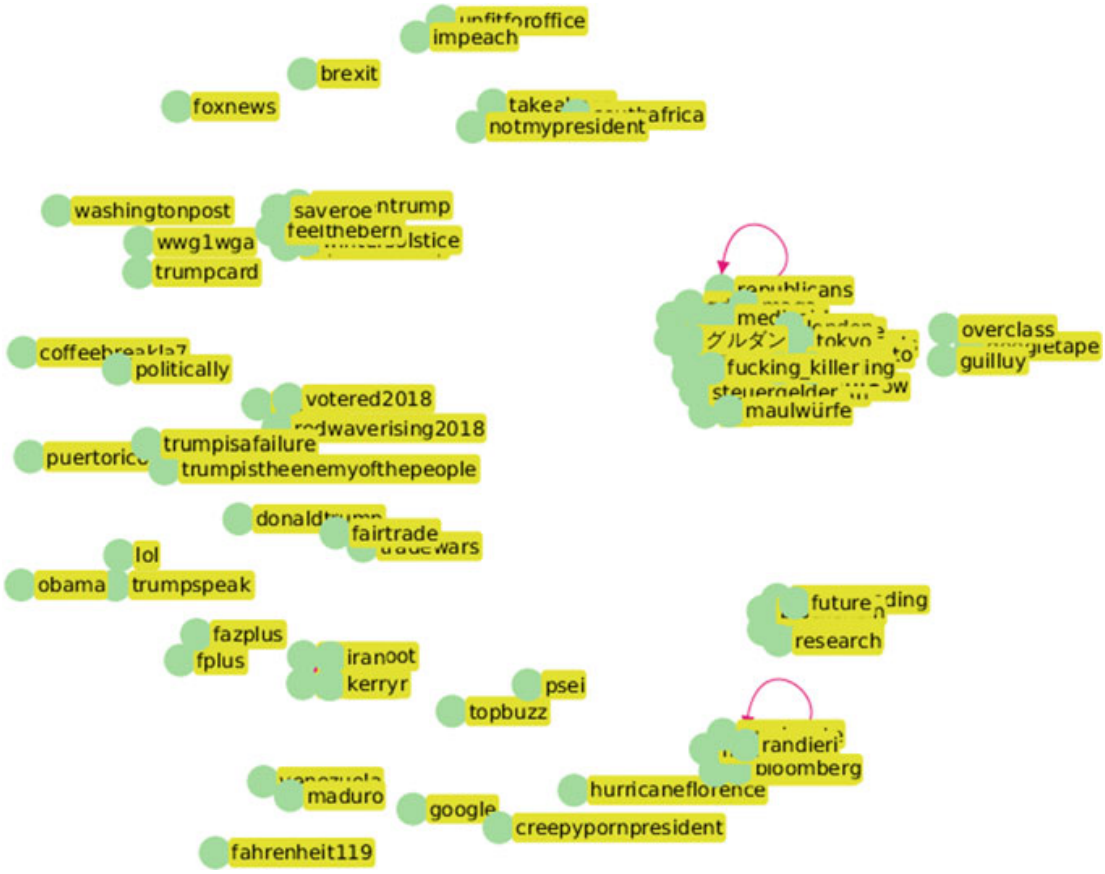


Fig. 7 Result of initial graph created and displayed using graphstream in real-time sliding window

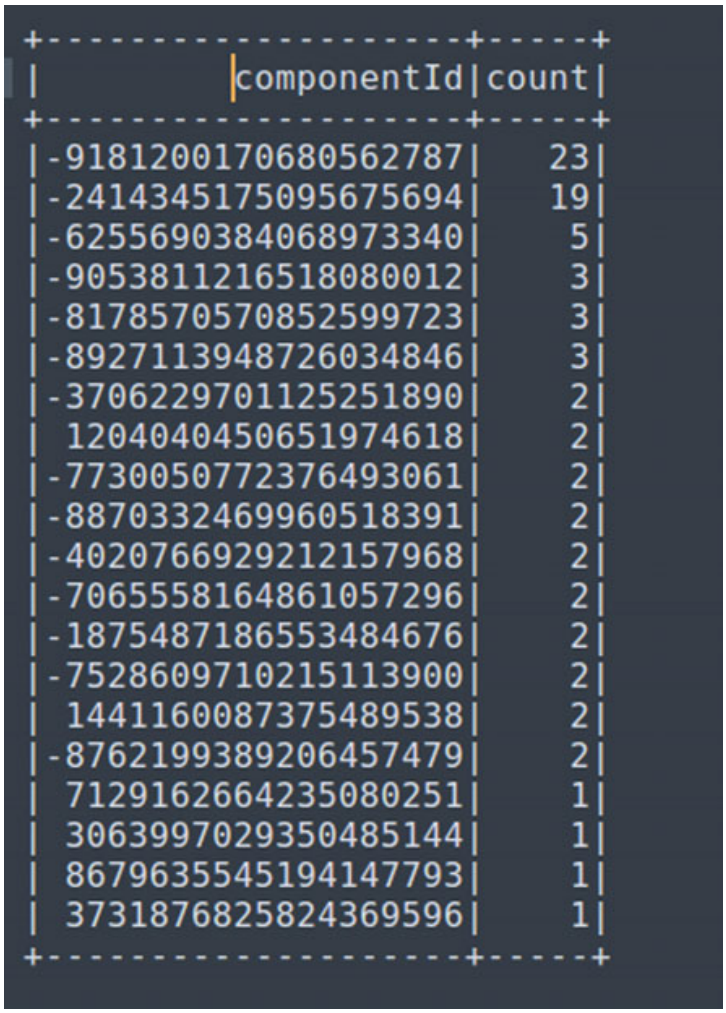
(b) Degree Distribution

A connected graph can have many different structures. For example, there can be one vertex that is connecting the whole graph. Eliminating the total graph becomes disconnected. So, for all these types of insights degree distribution that is *degree* of each vertex, where degree is defined as a number of edges connected to the vertex (Fig. 9).

(c) Community Detection in GraphX

Community detection algorithms usually focus on determining the graph which can either be overlapping or non-overlapping communities. When the communities are disjoint, there are no common elements we find between the communities. But when communities are overlapping or non-disjoint the elements among the nodes have common elements which they share. In this paper, we are focusing on one of the algorithms available on GraphX, i.e., Label Propagation and we are trying to visualize these using graphstream library (Fig. 10).

How Label Detection works is that it propagates the labels of node throughout the network which helps it to form the communities. The propagation of label is called as label flooding. In label flooding at a point of time single label becomes



componentId	count
-9181200170680562787	23
-2414345175095675694	19
-6255690384068973340	5
-9053811216518080012	3
-8178570570852599723	3
-8927113948726034846	3
-3706229701125251890	2
1204040450651974618	2
-7730050772376493061	2
-8870332469960518391	2
-4020766929212157968	2
-7065558164861057296	2
-1875487186553484676	2
-7528609710215113900	2
1441160087375489538	2
-8762199389206457479	2
7129162664235080251	1
3063997029350485144	1
8679635545194147793	1
3731876825824369596	1

Fig. 8 Result of realtime connected components in a GraphX. -9181200170680562787 is the most connected component with a node count of 23. We have identified that #trump helps in building the component with maximum node count

dominant when it is inside the densely populated community which is very well connected where we say that labels are trapped inside. Labels are supposed to be trapped inside whenever it is inside the densely populated network. Labels of each node are treated as opinion. Whenever the opinion is same, therefore the community is same.

In our case whenever the hashtags and their co-occurrence are closely related to trump the form of a community. Apart from trump, there is very less number of communities that are formed which are subsequent across windows of the given timeframe. So for our analysis, we will consider the community containing trump as the most important community and choosing other hashtags connected to it we will move to PageRank which further helps us in ranking the nodes in the graph and within the community.

In the community detection Label propagation algorithm using pregel in spark GraphX we have used a 5 step algorithm since waiting till convergence can hamper the streaming flow of the engine. And we have identified 5 to be the most optimal choice among all.

hashtag	degree
trump	362
maga	149
afp	91
bbc	89
news	87
ft	85
泛亚	85
中國	85
metoo	85
china	85
abc	85
foxnewsus	85
russia	85
岳昕	85
董瑶琼	85
xijinping	85
cnn	85
putin	77
soccer	73
huawei	71

Fig. 9 Result of degree distribution for a streaming graph. Hashtags *trump* and *maga* are among the highest degree vertices subsequently among various stream windows

(d) **PageRank**

Though the original PageRank algorithm was originally developed by Larry Page to rank Google search results. It can be applied majorly to all graphs where we are required to find the influence of the vertices in any graph.

Similarly for our problem statement, we have a graph where we have hashtags as vertices and an edge between the hashtag if two hashtags co-occur together in a tweet. So, our main motive here to apply the PageRank algorithm is to build the subgraph by removing all irrelevant nodes.

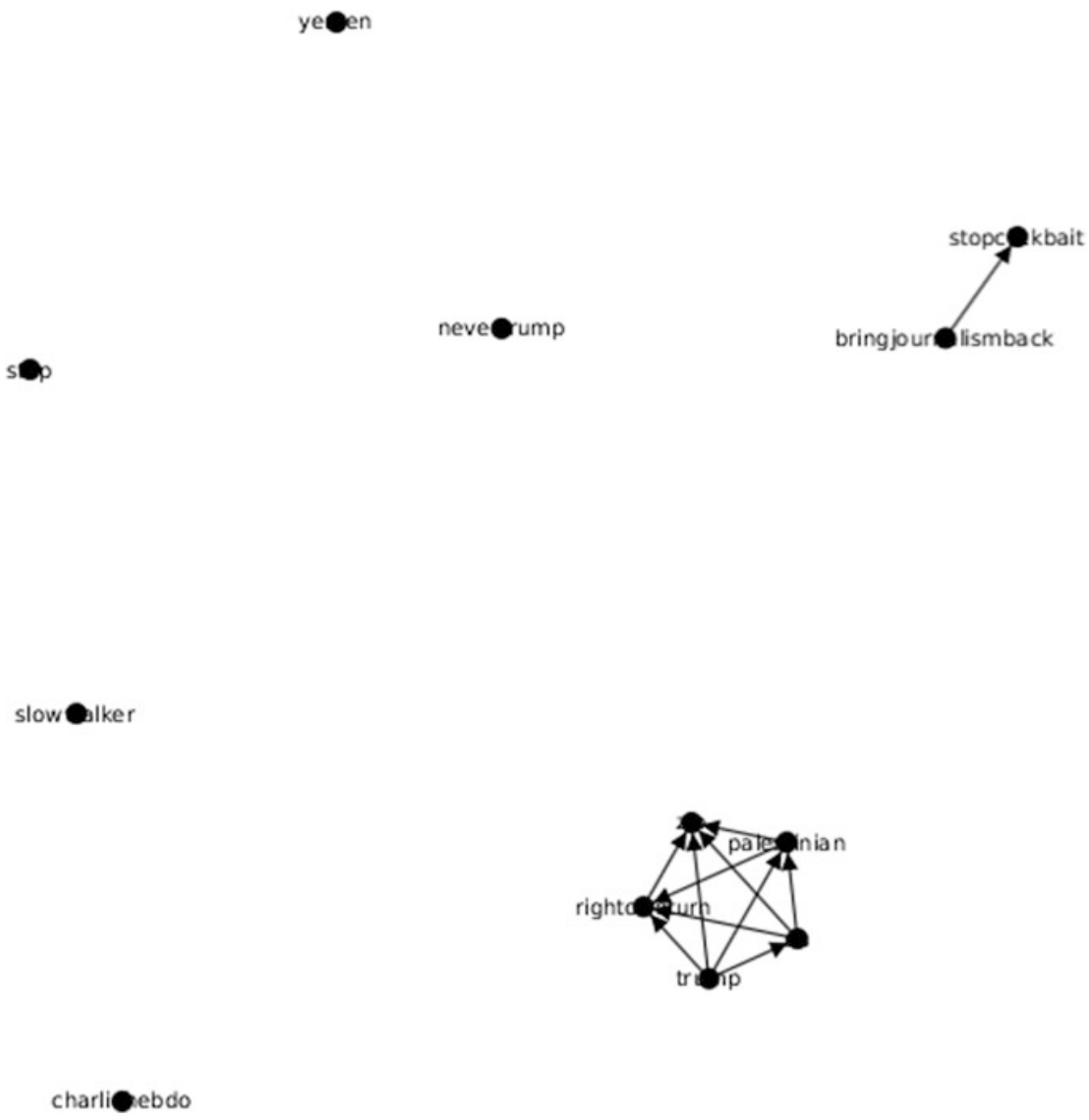


Fig. 10 Result of communities generated by the graph in one of the stream time window. Here, we can see various communities getting created with #trump is the largest community of in a particular sliding window

GraphX property graphs are directed multigraphs, which mean each edge in the graph specifies a unidirectional relationship, i.e., if there is an edge between two vertices in hashtag co-occurrence graph then hashtag1 and hashtag2 occur together in a similar fashion than hashtag2 and hashtag1. So we need to have two directions among two nodes for the PageRank to work correctly. *GraphX* offers static and dynamic implementations of PageRank where static PageRank runs for fixed no. of iterations and dynamic PageRank runs until convergence.

For our analysis, we will be using dynamic PageRank algorithm with a tolerance value of 0.01. We have done our analysis of multiple tolerance values like 0.01,

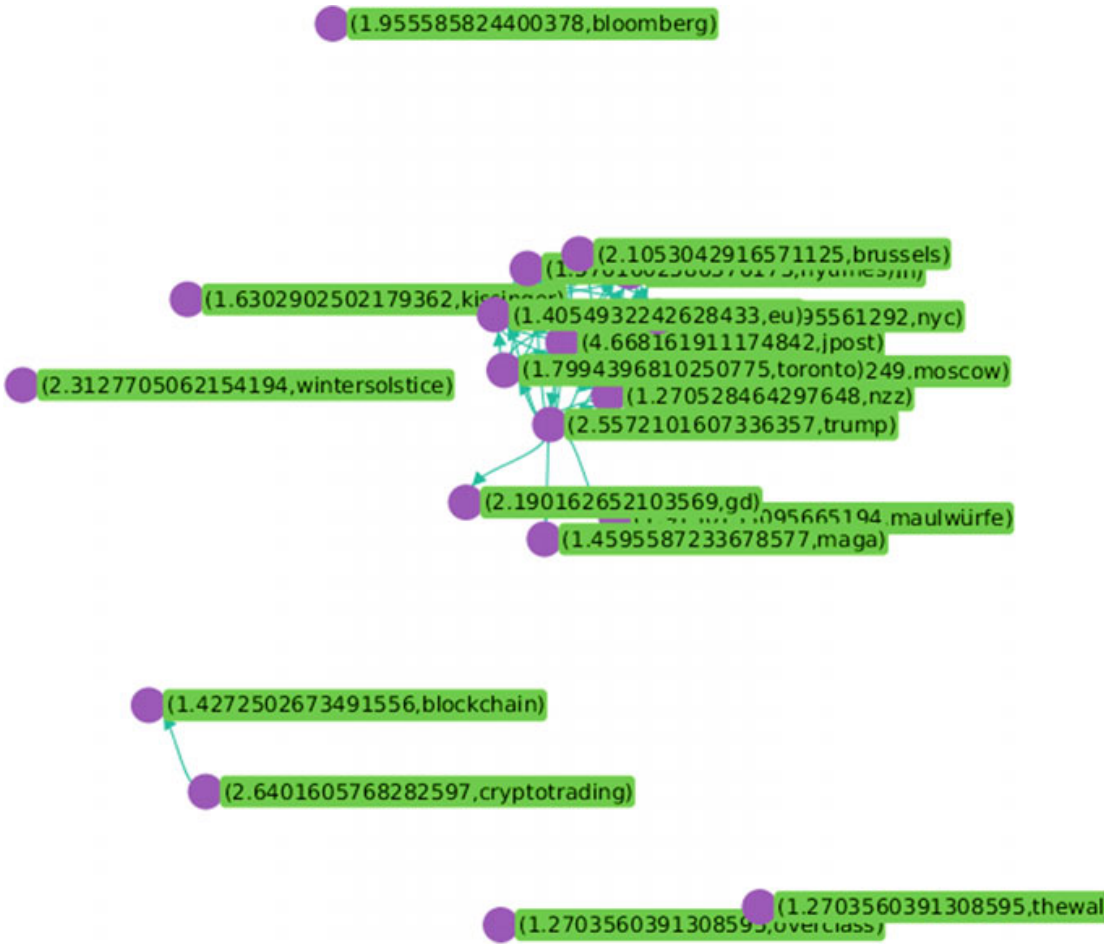


Fig. 11 Final graph visualization using graphstream. A result after community detection and PageRank in spark

0.0001, and 0.001, and found 0.01 to be the best. Since we are using real-time Spark Streaming and other tolerance values have taken longer than expected to converge and which might hamper our results for other subsequent windows of analysis in Spark Streaming [3, 15].

After we have got the PageRank, we will find top 20% of nodes in the graph and discard all irrelevant nodes. A node is called irrelevant. if it is not in topmost 20% of ranks in PageRank and create the new subgraph from this. This final graph will be our output graph for graphstream to visualize (Fig. 11).

> PageRank helps in ranking across the communities in the graph. And after fetching the top 20% of the nodes for every community of Hashtag we can see that PageRank can help us in getting the most important hashtags that can be recommended to the user for further analysis.

8 Conclusion

The practicality of building real-time graphs using Spark Streaming and identifying the related hashtags which help in hashtag recommendation is being investigated in this paper. Twitter analytics has considered to the top priority for most to get public opinions and views and hashtags helps us in understanding it in a better way. From the real-time hashtag co-occurrence graphs, we have seen the relevance of hashtag pairs and how to be so important for a specific period and how their importance fade with time. But some always remain relevant for a specified long period which played an important role in our analysis. Also, we have seen that we can seamlessly build real-time engine with graph functionality on top of it.

This can be proved to be a crucial tool for both social media and real-time applications. In this paper, we have also identified that if we will be able to combine the generated communities by Label Propagation with the PageRank algorithm, we get better results than waiting for both the algorithms to converge since in real-time analytics the one more thing followed is that how soon we are generating the results.

Further steps are required to improvise this.

Work on enhancing PageRank and community detection algorithm and improvement in tolerance of PageRank for better convergence this can further help us in recommending the graphs better.

Improvement in graph creation is required and to enhance we will build the graph by saving previous window data on creating a bigger and better graph which can help in giving better results.

Also, instead of using hashtags we will focus on creating graph with additional properties from the data.

References

1. Sadik, S., Gruenwald, L.: Research issues in outlier detection for data streams. ACM SIGKDD Explor. Newsl. **15**(1), 33–40 (2014)
2. Bifet, A., Holmes, G., Pfahringer, B.: Moa-tweetreader: real-time analysis in twitter streaming data. In: International Conference on Discovery Science. Springer, Berlin (2011)
3. Bifet, A., Frank, E. (2010). Sentiment knowledge discovery in twitter streaming data. In: International Conference on Discovery Science. Springer, Berlin
4. Duan, Y., et al.: Twitter topic summarization by ranking tweets using social influence and content quality. In: Proceedings of COLING 2012, pp. 763–780 (2012)
5. Ferragina, P., Piccinno, F., Santoro, R.: On analyzing hashtags in Twitter. In: International Conference on Web and Social Media (ICWSM). AAAI Press (2015)
6. Wu, W., Zhang, B., Ostendorf, M.: Automatic generation of personalized annotation tags for twitter users. In: Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Association for Computational Linguistics (2010)
7. Turney, P.D.: Information retrieval **2**, 303 (2000). https://doi.org/10.1023/A:1009976227802

8. Tomokiyo, T., Hurst, M.: A language model approach to keyphrase extraction (2003). https://doi.org/10.3115/1119282.1119287
9. Zangerle, E., Gassler, W., Specht, G.: Recommending#-tags in twitter. In: Workshop on Semantic Adaptive Social Web (SASWeb 2011). CEUR Workshop Proceedings, vol. 730, pp. 67–78 (2011)
10. Silberstein, A., Ashwin M., Raghu R.: Feed following: the big data challenge in social applications. In: Databases and Social Networks. ACM (2011)
11. Wang, X., et al.: Topic sentiment analysis in twitter: a graph-based hashtag sentiment classification approach. In: Proceedings of the 20th ACM International Conference on Information and Knowledge Management. ACM (2011)
12. Wang, M., Mizuho, I.: Hashtag sense induction based on co-occurrence graphs. In: Asia-Pacific Web Conference. Springer, Cham (2015)
13. Ramamonjison, R.: Apache Spark Graph Processing. Packt Publishing Ltd, Birmingham (2015)
14. Ryza, S., et al.: Advanced Analytics with Spark: Patterns for Learning from Data at Scale. O'Reilly Media Inc, Sebastopol (2017)
15. Pavlyshenko, B.: Forecasting of events by Tweet data Mining (2013). arXiv preprint arXiv:1310.3499

Author Index

A
Abayomi-Zannu, Temidayo, 339
Adepu, Divya, 93
Adewumi, Adewole, 685
Adnan, Kiran, 139, 301
Agarwal, Mohit, 569
Ahuja, Ravin, 685
Akbar, Rehan, 139, 301
Akila, D., 197
Alfred, Rayner, 151
Ali, Adnan Bin Amanat, 301
Alzahrani, Ahmad, 269
Ananya, M., 673
Andrade, Melvita, 93
Arora, Arushi, 3
Arunkumar, K., 325
Aswale, Neerja, 57
Awirothananon, Thatphong, 255
Azeez, Nureni Ayofe, 685
Azeta, Ambrose A., 673
Azeta, Victor I., 673

B
Bachhety, Shivam, 503
Bada, Taiwo Mayowa, 685
Badgujar, Aishwarya, 93
Baital, Kalyan, 111
Banodha, Umesh, 29
Bharadwaj, Vishnu, 555
Bhattacharya, Sonia, 365
Bhatt, Alok, 555
Bhosale, Shivjeet, 475
Biswal, Monalisa, 41
Booba, B., 713

C
Chakrabarti, Amlan, 111
Chakrabarty, Himadri Bhattacharyya, 365
Chandolikar, Neelam, 475
Chandrakar, Ruchi, 41
Chandrasekaran, E., 197
Chatterjee, Punyasha, 633
Chatterjee, Sankhadeep, 543
Chaudhary, Shankar, 613
Chawra, Vrajesh Kumar, 663
Chutatape, Opas, 125

D
Debnath, Arpita, 633
Deepika, R., 517
Devendran, A., 325
Dey, Anindita, 161, 171
Diwan, Prabhat, 161, 171
Dutta, Nitul, 601

E
Elavarasi, S., 185

G
Garg, Vikas, 69
Gill, Neetu, 161, 171
Goddy-Worlu, Rowland, 339
Govil, Himanshu, 161, 171
Grant, Emanuel, 339
Guha, Subhanil, 161, 171
Gupta, Govind P., 663
Gupta, Nirmal Kumar, 285
Gupta, Suneet K., 569

© Springer Nature Singapore Pte Ltd. 2020

N. Sharma et al. (eds.), *Data Management, Analytics and Innovation*, Advances in Intelligent Systems and Computing 1042, https://doi.org/10.1007/978-981-32-9949-8

H
Haq, Rana Azhar Ul, 229
Hegde, Sarika, 463

J
Jain, Rachna, 503, 583
Jalnekar, Rajesh, 475
Janghel, Rekh Ram, 425
Joglekar, Pushkar, 475
John, Sonali, 697

K
Kamaruddin, Sk., 215
Karthika Renuka, D., 241
Karunya, K., 517
Kashyap, Indu, 583
Kashyap, Smiti, 69
Kaur, Gurkarandesh, 451
Kaur, Varinderjit, 439
Kavita, Choudhary, 15
Khan, Asif Riaz, 139
Khare, Akhil, 15, 613
Khor, Siak Wang, 301
Krishna, Ragini, 489
Kumar, Akshi, 3

M
Mahi, Gurjot Singh, 383
Majeed, Fiaz, 229
Makkar, Garima, 529
Mandal, Prasanta, 633
Menon, Varun G., 697
Misra, Sanjay, 673, 685
Misra, Diganta, 569
Mohanty, Sachi Nandan, 569
Mukul, Kavya, 57

N
Nale, Ruchita, 41
Naonueng, Kitti, 125
Nayak, Jyoti Ranjan, 401
Nayyar, Anand, 3, 503, 697
Niranjani, G., 643

O
Oberoi, Ashish, 439, 451
Odun-Ayo, Isaac, 339

P
Pagare, Reena, 613
Pawade, Dipti, 93
Peddawad, Dipali, 475
Phoophuangpairoj, Rong, 125
Prashanth, C.M., 489
Prema, A., 713

R
Rajamohana, S.P., 517
Rana, Manoj Kumar, 633
Rathore, Yogesh Kumar, 425
Ravi, Vadlamani, 215
Rawal, Ayush, 555
Renugadevi, R., 713
Rohil, Mukesh Kumar, 285
Roy, Santanu, 543

S
Sadaoui, Samira, 269
Saha, Pallavi, 543
Sahu, Anil Kumar, 413
Sahu, Rajkumar, 401
Sakhapara, Avani, 93
Sasi Kumar, A., 197
Saxena, Kanak, 29
Sen, Soumya, 543
Sharma, Sonam, 723
Shaw, Binod, 401
Shetty, Surendra, 463
Shilaskar, Swati, 475
Singh, Gurinder, 69
Singh, Lolita, 601
Sinha, G.R., 413
Soni, Sapna, 413
Sri Vinitha, V., 241
Suseendran, G., 185, 197

T
Thanjunpong, Sathaya, 255
Tomar, Kanika Singh, 69

U
Umamaheswari, K., 517, 643
Upasani, Shubhangi, 3

V
Van der Vyver, Charles, 685
Verma, Amandeep, 383
Verma, Archana, 425

Y
Yadav, Sanjeev Kumar, 15
Yadav, Vineet, 555
Yu, Hin Fuk, 151

International Journal of Mechanical and Production Engineering Research and Development (IJMPERD)
ISSN(P): 2249-6890; ISSN(E): 2249-8001
Vol. 8, Issue 6, Dec 2018, 635-642
© TJPRC Pvt. Ltd.

TRANS STELLAR
•Journal Publications •Research Consultancy

ENHANCEMENT OF SURFACE INTEGRITY USING SPARK ASSISTED ABRASIVE FLOW MACHINING PROCESS

PARVESH ALI[1], R. S. WALIA[2], QASIM MURTAZA[2] & RANGANATH M. S[2]

[1]*Research Scholar, Delhi Technological University, Delhi, India*

[2]*Professor, Delhi Technological University, Delhi, India*

ABSTRACT

In today's competitive environment components having high dimensional accuracy and better functional performance are highly demanded. Abrasive Flow Machining is one of the vital non-conventional techniques used for the finishing of complex geometries which is not easy with other conventional processes. It provides finishing up to nano level. During the finishing, parameters affecting the system efficiency are very important to identify to improve the productivity of the system. This paper discusses a newly developed technique i.e spark assisted Abrasive Flow machining, which uses the spark energy to improve the material removal and providing better surface finish and also parameters such as rotational speed of the electrode, duty cycle and current were optimized for the better surface finish using Taguchi L_9 OA.

KEYWORDS: Finishing, Spark, Hybrid, Current, Duty Cycle & Rotation

Received: Sep 27, 2018; **Accepted:** Oct 17, 2018; **Published:** Dec 01, 2018; **Paper Id.:** IJMPERDDEC201866

Original Article

INTRODUCTION

Abrasive Flow Machining (AFM) is a nonconventional technique used for the finishing of complicated shapes which cannot be obtained by conventional techniques. However, in recent past, many traditional processes such as honing, lapping and broaching were used for finishing. But these processes cannot provide nano level finishing. AFM is a technique that uses a non-Newtonian media (combination of polymer, gel and abrasive) and is forced to flow through the restrictive path at a huge pressure. The advantage of using this media is that it can flow to any unreachable area which is not possible using traditional techniques [1, 2]. This process can be used to finish the burrs, edges, recast layers [3]. The disadvantage of this process is low material removal. So various researchers hybridize this process with other techniques to get better material removal. Kenda et al. [4] used AFM process for finishing the profile of plastic gear and found an increase in the life cycle of the gear with a significant decrease in the finishing cost. Subramanian et al. [5] finished hip joint of Co-Cr alloy using AFM process in the biomedical area and found minimum surface roughness of 39 nm. Sidpara et al. [6] studied the rheological behavior of the MR fluid in the presence of a magnetic field. Sambharia and Mali [7-9] developed a new media based on polymer abrasive gel and identified the effect of parameters such as abrasive concentration, abrasive mesh size, media temperature, an additive percentage on the viscosity of media. Judal and Yadava [10] combined ECM process with Magnetic assisted AFM to finish non-magnetic stainless steel and analyze the effect of parameters such as magnetic flux density, rotational speed and current on the material removal and surface roughness.

METHODS

Although Abrasive Flow Machining process provides a good level of finishing it has a constraint of low material removal. Many researchers have hybridized this process with other nonconventional techniques to improve its material removal. This paper discusses a newly developed technique which uses the spark energy to melt the surface material so that the abrasives can easily take away the softened material. This process uses a power supply system which converts the AC supply to the pulsed DC supply. The positive supply was given to the workpiece while negative supply was given to the electrode inside the hollow cavity. When the supply is given to both the poles electrons start generating between both the poles and start moving to the opposite polarity. This movement of ions can be seen as a spark which melted the surface material. The experimental setup of the developed hybrid process is shown in figure 1.

Figure 1: Experimental Setup of Developed Hybrid Abrasive Flow Machining Process

The experimental setup included a fixture for holding the workpiece and guiding the media from one media cylinder to the other, power supply system, gears for the rotation of electrode to maintain the uniformity of surface, bearings. The rotation of either workpiece or electrode is necessary to maintain the uniformity of the surface otherwise spark will melt the surface at a particular spot and damages the surface integrity of the surface.

PERCENTAGE IMPROVEMENT IN SURFACE ROUGHNESS

To analyze the effect of variable parameters on the response i.e percentage improvement in surface roughness, experiments were planned and analyzed using Taguchi L_9 Orthogonal array. Percentage improvement in surface roughness can be calculated by the equation given below-

% improvement in R_a = (Initial Roughness - Final Roughness)/(Initial Roughness) * 100

The present L_9 OA has three parameters with three levels of each. It has 8 degrees of freedom. Brass was taken as the workpiece having a hollow cylindrical shape of an outer diameter of 10 mm and an inner diameter of 8 mm with a length of 16 mm. The variable parameters with their levels affecting the performance of the system are noted in Table 1.

Table 1: Process Parameters and their Values at Different Levels

Symbol	Parameters	Unit	L_1	L_2	L_3
S	Rotational speed	RPM	150	200	250
D	Duty Cycle	Fraction	0.68	0.73	0.78
I	Current	Ampere	3	6	9
Polymer to Gel Ratio: 1:1, **Workpiece Material:** Brass, **Abrasive Type:** Al_2O_3, **Initial Surface Roughness:** 1.1 - 2.85 microns					

Experiments were performed according to the L_9 OA. A number of 9 experiments were performed with three times each, which means a total 27 experiments were performed to analyze the behavior of the parameters as given in Table 1. The initial surface roughness of the workpiece was in the range of 1.1-2.85 micron. For every sample operation and its replications, % ΔR_a was measured and recorded in Table 2.

Table 2: Experimental Results of Percentage Improvement in Surface Roughness as per TAGUCHI L_9 OA

E.N	Run Order	% Improvement in Ra			S/N Ratio
		R_1	R_2	R_3	
1	1	7.25	8.34	8.79	18.11
2	4	16.1	17.1	19.24	24.78
3	7	15.5	16.85	15.6	24.05
4	2	16.42	17.95	17.23	24.69
5	5	29.08	31.92	32.12	29.81
6	8	20.84	23.74	28.295	27.5
7	3	9.58	10.52	10.93	20.25
8	6	13.94	14.87	13.31	22.92
9	9	14.68	14.27	15.76	23.44
Total		**143.39**	**155.56**	**161.275**	
		T ΔRa = Overall mean of ΔRa=17.04%			

RESULTS AND DISCUSSIONS

Surface roughness was measured through the Taylor Hobson precision machine having the resolution of 0.01 micrometer. The percentage improved R_a value for each trial condition is shown in Table 2. Higher the better quality characteristic was taken for the response and for calculating the S/N ratio as given in equation 1.

$$(S/N)_{HB} = -10 \log[\frac{1}{n} \sum_{j=1}^{R} \frac{1}{Y_j^2}] \quad (1)$$

Figure 2 shows, on increasing the rotational speed percentage improvement in surface finish increased up to 200 rpm and further decreased on increasing the rotational speed. As the rotational speed was increased it forced the abrasive particles towards the surface with a larger force due to the developed centrifugal force which corresponded more material removal and greater surface finish, but further increasing the rotational speed increased the media temperature and caused restriction during the media flow. This reduced the media viscosity which corresponded decrease in viscosity of media due to which abrasive particles did not participate in the abrasion process which reduced the percentage improvement in the surface roughness.

Figure 3 shows surface finish got better on increasing the duty cycle up to 0.73 but decreased after a further increase in duty cycle. Duty cycle means the frequency of spark in the gap. As duty cycle increases, the discharge energy will be available for a larger duration which developed high temperature on the surface and abrasive particles easily took away the melted material. However, on further increasing the duty cycle produced a larger crater on the surface and

degrades the surface quality.

Figure 4 shows the effect of current on the surface finish of the product. As the value of current was increased larger discharge energy was available for melting the surface material, which made easy for the abrasive particles to take away the softened material and corresponded better surface finish.

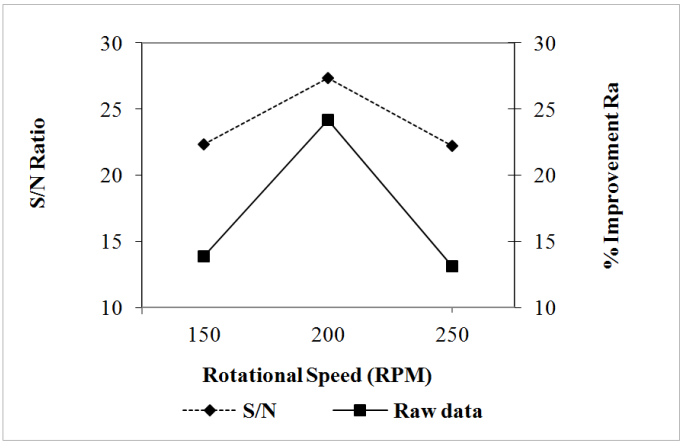


Figure 2: Effect of Rotational Speed on the S/N Ratio and % ΔR_a

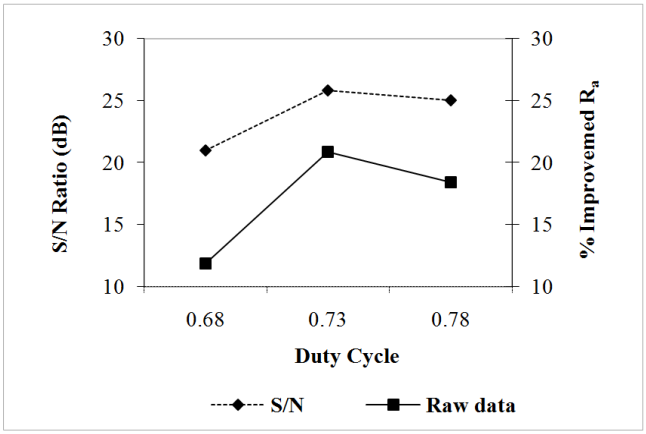


Figure 3: Effect of Duty Cycle on the S/N Ratio and % ΔRa

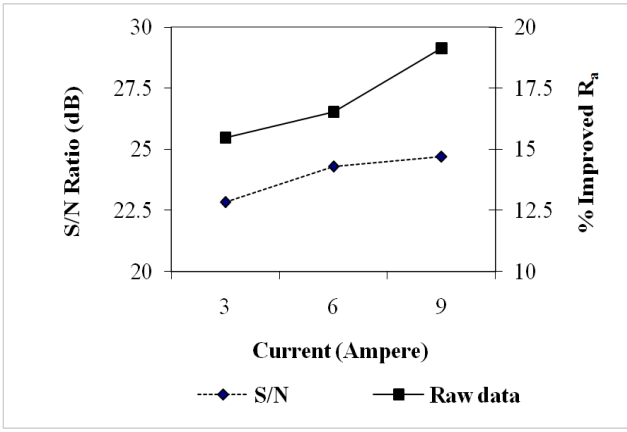


Figure 4: Effect of Current on the S/N Ratio and % ΔR_a

Table 3 shows the average value and main effect for the % Δ R_a.

Table 3: Main Effects of % Δ R_a

Process Parameter	Level	Rotational Speed (S)		Duty Cycle (D)		Current (I)	
Type of Data		Raw Data	S/N Ratio	Raw Data	S/N Ratio	Raw Data	S/N Ratio
Average Values(%Ra)	L_1	13.86	22.31	11.89	21.01	15.48	22.84
	L_2	24.17	27.33	20.85	25.83	16.52	24.3
	L_3	13.09	22.2	18.39	25	19.12	24.7
Main Effects(%Ra)	L_2-L_1	10.31	5.02	8.96	4.81	1.04	1.45
	L_3-L_2	-11.08	-5.13	-2.46	-0.83	2.59	0.4
Difference (L_3-L_2)–(L_2-L_1)		-21.39	-10.15	-11.42	-5.65	1.55	-1.05
L_1, L_2, L_3 shows levels of parameters.							

In this analysis, ANOVA technique was used to analyze the effect of variables on the response, i.e % ΔR_a. Table 4 and Table 5 represent the ANOVA for the raw data and S/N data. It can be observed from the Tables that the parameters affect the mean and variability in the R_a value.

Table 4: Pooled value of ANOVA for Raw Data (ΔR_a)

Source	SS	DOF	V	F-Ratio	SS'	P%
Rotational Speed	689.3	2	344.65	97.11	682.2	56.99
Duty Cycle	386.04	2	193.02	54.38	378.94	31.91
Current	63.11	2	31.55	8.89	56.01	5.21
E	70.98	20	3.54	------	92.27	5.86
Total (T)	**1209.44**	**26**	*	-------	**1209.44**	**100**
Fcritical =3.4928 (Tabulated Value)						

Table 5: Pooled ANOVA (S/N Data)(ΔR_a)

Source	SS	DOF	V	F-Ratio	SS'	P%
Rotational Speed	51.56	2	25.78	207.87	51.31	52.97
Duty Cycle	39.77	2	19.88	160.33	39.52	40.86
Current	5.74	2	2.87	23.17	5.50	5.90
e	0.24	2	0.12	------	0.99	0.25
Total (T)	**97.33**	**8**	--------	-------	**97.33**	**100**
F critical = 19(Tabulated Value)						

The optimum %ΔR_a was determined and related confidence intervals were predicted.

For the mean of %ΔR_a, optimum condition for the response is estimated as:

$\mu = \bar{S}_2 + \bar{D}_2 + \bar{I}_3 - 2\bar{T}$ (1)

$\bar{T}$ = overall mean of the R_a obtained = 17.04 %

$\bar{S}_2$ = Average value of % age improvement in R_a at the 2nd level of rotational speed = 24.17%

$\bar{D}_2$ = Average value of % age improvement in R_a at the 2nd level of duty cycle = 20.85 %

$\bar{I}_3$ = Average value of % age improvement in R_a at the 3rd level of Current =19.12 %

Substituting these values, % improvement in R_a = 30.06 %

(CL_{CE}) and (CL_{POP}) can be calculated by using the equations (2) & (3):

$$CI_{CE} = \sqrt{F_a\,(1, f_e)\, V_e \left[\frac{1}{n_{eff}} + \frac{1}{R}\right]} \quad (2)$$

$$CI_{POP} = \sqrt{[F_a(1, f_e)\, V_e]/ [n_{eff}]} \quad (3)$$

Where CI_{CE} = Confidence interval for confirmation experiments

CI_{POP} = Confidence interval for population

$F_\alpha(1, f_e)$ = The F- ratio having confidence level of (1-α)

f_e =3.49 (Value Tabulated)

f_e = Error Degree Of Freedom = 20

N= Total number of experiments =27

R = Size of Sample = 3

V_e = Error variance = 3.55

By using equation no. (4), one may achieve

n_{eff} = 3.86

So, $CI_{CE} = \pm 2.70$

and $CI_{POP} = \pm 1.79$

The Confirmation interval for predicted response is:

27.36< %age improvement in ΔR_a <32.76

The 95% confirmation interval of predicted mean is

28.27<%age improvement in ΔR_a <31.85

The optimum level of the parameters for percentage $\%\Delta R_a$ is $S_2D_2I_3$.

S_2 = Rotational speed at second level =200 RPM

D_2 = Duty Cycle at second level = 0.73

I_3 = Current at Third level =9 Ampere

The results are presented in Table 6 for % improved R_a.

Percentage improved R_a from the confirmation experiments are inside the specified range.

The mean values of percentage improved R_a and S/N ratio for every parameter at all the levels was calculated and given in Table 6.

Table 6: Values Obtained from Confirmation Experiments

Response	Optimum Condition	Predicted Response	Confidence Intervals 95%	Real Value
%Improvement ΔR_a	$S_2D_2I_3$	30.06 %	CI_{CE}:27.36<%ΔR_a<32.76 CI_{POP}:28.27<%ΔR_a<31.85	27.75 %

SEM ANALYSIS

As the power supply is provided between both poles, a spark is generated in the gap. This developed high temperature on the surface and caused oxide layer formation on the surface. The surface materials started melting and became soft. The melting of the surface can be easily studied by figure 5, which clearly shows the spots developed by the spark on the surface.

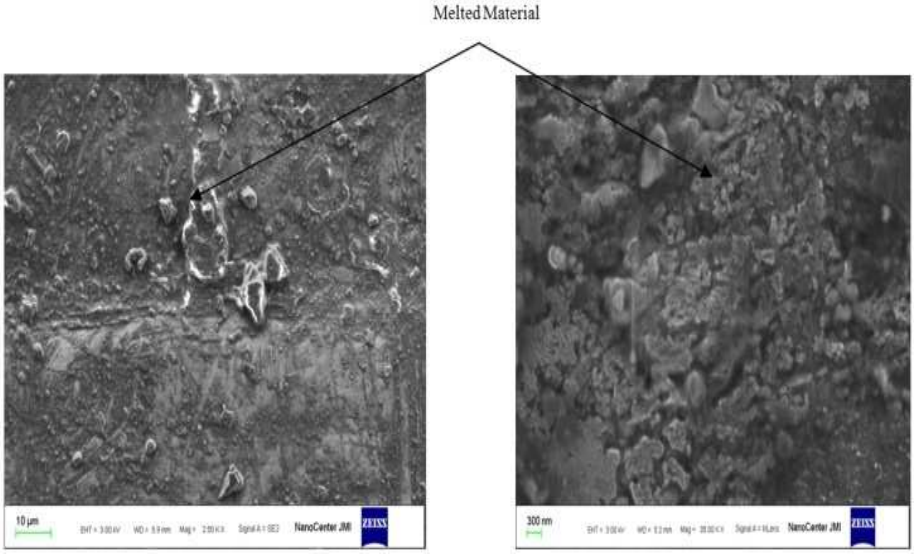


Figure 5: SEM Images of Surface During Finishing

XRD ANALYSIS

X- ray diffraction technique for the finished surface was analyzed by using "X' Pert High Score" software. XRD of the brass workpiece after finishing can be observed from figure 6. The figure confirms the formation of Copper vanadium fluoride oxide at the plane (004) for a diffraction angle of 49.48 degrees [Ref- 00-043-0082], and formation of Copper magnesium oxide at the plane (4400) for a diffraction angle of 43.418 degrees [Ref- 01-074-1921]. Copper Praseodymium oxide was formed at a plane (105) with a diffraction angle of 43.47 degrees [Ref- 00-049-1891].

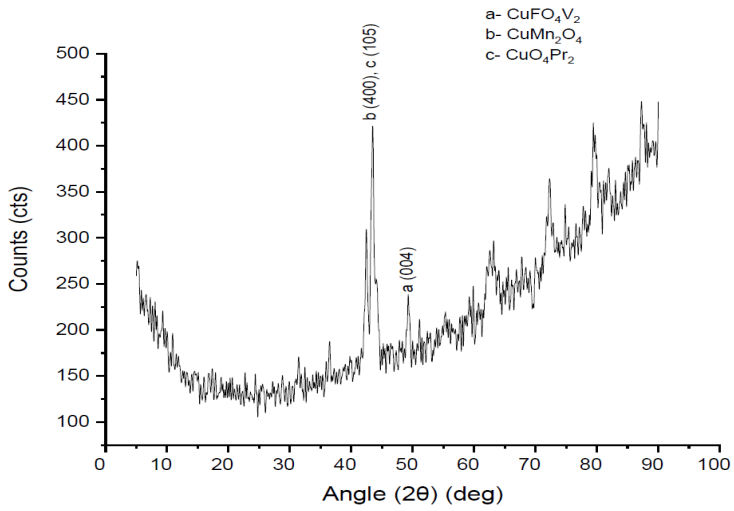


Figure 6: XRD Graph of the Finished Brass Workpiece

CONCLUSIONS

The following conclusions can be drawn on the basis of the above experimental study-

- The experimental results show that rotational speed is contributing 56.99 %, a duty cycle is contributing 31.91% and current is contributing 5.21 % towards percentage improvement in surface roughness.
- The 95% confidence interval of the % ΔR_a is $27.36 < \%\Delta R_a < 32.76$.
- The experimental results show optimum result at the second level of rotation speed (200 rpm), a second level of the duty cycle (0.73) and third level of current (9 Ampere).
- Melted material observed by the SEM image shows less requirement of abrasion force for removing the material.
- The XRD results confirm the oxide formation on the surface of the materials.

REFERENCES

1. *Tan, K. L; Yeo, S. H; & Ong, C. H. (2016). Nontraditional Finishing Processes for Internal Surfaces and Passages: A Review. Proc. Institution of Mechanical Engineers, Part B, 1–15. DOI:10.1177/0954405415626087.*
2. *Singh, G; Singh, A. K; & Garg, P.(2016). Development of Magnetorheological Finishing Process for External Cylindrical Surfaces. Materials and Manufacturing Process, 32, 1–8. DOI:10.1080/10426914.2016.1221082.*
3. *Perry, K. E. (1977). Abrasive Flow Machining Method and Tooling. US 4005549 A.*
4. *Kenda, J; Duhovnik, J; Tavčar, J; & Kopac, J. (2014). Abrasive Flow Machining Applied to Plastic Gear Matrix Polishing. Int. J. Adv. Manuf. Technology, 71(1–4), 141–151. DOI: 10.1007/s00170-013- 5461-4.*
5. *Thirumalai K; & Natchimuthu, S. (2016). Nanometric Finishing on Biomedical Implants by Abrasive Flow Finishing. J. Inst. Eng. (India): Series C, 97(1), 55–61. DOI:10.1007/s40032-015-0190-0.*
6. *Sidpara, A; Das, M; & Jain, V. K. (2009). Rheological Characterization of Magnetorheological Finishing Fluid. Mater. Manuf. Process, 24, 1467–1478. DOI:10.1080/10426910903367410.*
7. *Mali, H.S; & Jaikishan.(2014). Developing Alternative Polymer Abrasive Gels for Abrasive Flow Finishing Process. 5th International & 26th All India Manufacturing Technology, Design and Research Conference (AIMTDR 2014), 1–8.*
8. *Sambharia, J; & Mali, H. S. (2015). Characterisation and Performance Evaluation of Developed Alternative Polymer Abrasive Gels for Abrasive Flow Finishing Process. Int. J. Precis. Technology, 5, 185–200. DOI:10.1504/IJPTECH.2015.073824.*
9. *Sambharia, J; Ram, J; & Mali, H. S.(2015). Rheological Investigation and Comparative Study on Synthesized Low Cost Consumables for Abrasive Flow Finishing Processes. International Conference on Precision, Meso, Micro and Nano Engineering (COPEN-9), IIT Bombay.*
10. *Nagpure, S. S., & Tufail, M. Concept of automated machining and inspection in FMS.*
11. *Judal, K. B; & Yadava, Y.(2013). Experimental Investigations into Electrochemical Magnetic Abrasive Machining of Cylindrical Shaped Nonmagnetic Stainless Steel Workpiece. Mater. Manuf. Process, 28(10), 1095–1101. DOI:10.1080/10426914.2013.7924.*

Experimental and Computational Investigation on Mechanical Properties of Reinforced Additive Manufactured Component

Nagendra Kumar Maurya[1,2*], Vikas Rastogi[1], Pushpendra Singh[1]

[1]Department of Mechanical Engineering, Delhi Technological University, Delhi, India, India

[2]Department of Mechanical Engineering, GL Bajaj Institute of Technology and Management, Gr. Noida, U.P, India

E-mail: nagendramnnit@gmail.com

(Received Jun 25, 2019; accepted September 20, 2019).

Abstract: Fused deposition modelling (FDM) is the one among various additive manufacturing techniques which can fabricate component of multi material. However, the tensile strength of FDM component is naturally low and hence it is difficult to use for the engineering application. For enhancing the mechanical properties, this study explores a novel methodology by adding the high strength material in between matrix material as reinforcement in the additive structure. In this study, ABS and PLA based reinforced composites are prepared. ABS and PLA are used as matrix material whereas PETG is used for the reinforcement. Unidirectional tensile test is performed to check the effect of reinforcement on tensile strength of the fabricated component. Experimental results depict that addition of PETG reinforcement material improves the tensile strength of ABS by about 70% and tensile strength of PLA material is improved by about 8%. Computational analysis is also performed under similar boundary condition to physical model. It is observed that computational results are close approximation to the experimental results. The methodology proposed in this study could be very useful for the fabrication of high strength component by using low cost FDM technology.

Keywords: Multi materials; reinforced additive manufacturing; mechanical properties; computational analysis; fused deposition modelling

1. Introduction

Over the past few years, industrial demand of biodegradable materials has suddenly increased due to the biodegradability, renewability of the base materials and less greenhouse gas emissions and recycling [1)]. ABS, PLA and PETG have been extensively used in the medical industry, packaging industry, automobile industry, aerospace industry etc due to its biodegradability[2)]. However, poor mechanical, thermal properties and higher cost limit to use in large scale in the industry. Therefore many researchers have made effort to improve the mechanical properties of ABS and PLA material by varying the process parameters like raster angle, orientation, infill pattern, infill density, layer thickness, shell thickness etc. Some studies were carried out to improve the mechanical properties by mixing the reinforcement fiber into the extruding wire. Significant improvement in the mechanical properties was found due to the addition of reinforcement fiber in the matrix material.

Panes et al.[3)] have performed an experiment to characterize the tensile strength of ABS and PLA part printed by FDM process. The process parameters used in this study were layer thickness, infill density and orientation. Pruska i3 printer was used for the fabrication of test samples and HAYTOM HM-D was used for the tensile test. Mansour et al. [4)] have conducted a series of experiments to find out the dynamic and mechanical behavior of 3-D printed material. Reinforced carbon fibers PETG material was used for the fabrication of specimen. The reinforcement direction of the fibers was 0° and 90°. It was observed that due to addition of carbon fibers, modulus of elasticity and hardness increased by 30% whereas compressive strain increased by 66%. Masood et al.[5)] have performed an experiment to investigate the effect of FDM process parameter orientation on tensile strength and surface roughness of PLA material. Along y-axis uneven surface roughness was observed and tensile strength of material was more along x-axis as compared to y-axis. At an angle of 45°, tensile elongation was nearly same for all the three orientation.

Wu et al.[6)] have highlighted the influence of layer thickness and raster angle on the mechanical properties of 3-D printed component. It was found that for higher layer thickness, tensile strength of ABS and PETG material tends to maximum at 0° degree and 90° orientation. Lanzotti et al.[7)] have highlighted the impact

of process parameters such as infill orientation, layer thickness and number of shell on tensile strength of PLA component fabricated through FDM process. Results shows that on increase of orientation, ultimate tensile strength of the material was found to be decreased. On increasing the number of shell from 2 to 6, ultimate tensile strength was increased from 42 MPa to 52 MPa. Whereas when the layer thickness increased from 0.10 mm to 0.20 mm, minute changes were observed in the ultimate tensile strength of material.

Sugavaneswaran et al.[8] have conducted experimental and analytical investigation for the effect of reinforcement and its orientation on elastic modulus of structures manufactured by polyjet 3D printing technique. Vero black "ABS like" material was used as reinforcement and darus white as matrix material. Instron machine was used for tensile testing and rom equation was used to find out analytical results. The results shows that due to presence of reinforcement, elastic modulus increased by 6.79% in transverse direction (90°), 21.9% along the longitudinal direction at (0°) and 14.88% along inclined direction at (45°). There was good correlation between experimental and analytical results. Afrose et al.[9] have carried out experimental study to investigate the impact of orientation on tensile properties of the PLA thermoplastic material FDM AM technique. The specimens were printed at three orientations namely X, Y and 45°. The shape of the specimen was used similar to dog bone. Result shows that maximum tensile strength was found along the X orientation. Summary of some other research work related to FDM technology is reported in table 1.

Recent development in the field of additive manufacturing is to fabricate heterogeneous material by using multi material. In present work, an attempt was made to investigate the effect of unidirectional reinforcement of PETG material on mechanical properties of ABS (material) and PLA (material). Significant improvement in the tensile strength was found due to the reinforcement of PETG material.

2. Materials and Method

The present study focuses on development of reinforced composite material by using multi material. In additive manufacturing softer material is referred as matrix material and harder material is referred as reinforcement material [17]. Development of AM for the fabrication of multi material had improved the mechanical properties, specific functional requirement and design feature for the new product design. In this work, effect of unidirectional reinforcement of PETG material on mechanical properties of ABS and PLA material was investigated. The next sub section will provide the details of matrix and reinforcement material used in this investigation.

2.1 Matrix and reinforcement material

In this work, two separate cases were considered for selection of matrix material. In first case, ABS was considered as matrix material and PETG was used as reinforcement material. In second case, PLA filament was used as matrix material and PETG was considered as reinforcement material. Three different reinforcement percentages were used for the preparation of test samples viz., 15%, 25% and 35%. ABS was the weakest material and PETZ was the stronger material among all the three material. The schematics arrangement of matrix and reinforcement material is shown in Fig.1. The characteristics of ABS, PLA and PETG materials are shown in table 2.

Table 1 Summary of the characterization of mechanical properties in FDM technology

Investigator Name	Material	Raster angle	Other process parameters	Mechanical properties
Durgun et al. [10]	ABS	0°, 30°, 45°, 60°,90°	Orientation (horizontal, vertical and perpendicular)	Tensile strength
Ravari et al.[11]	PLA	Constant	-	Compressive strength, collapse stress
Onwubolu et al.[12]	ABS	Constant	Layer thickness, part orientation, raster angle, raster width, and air gap	
Dawoud et al.[13]	ABS	0/−90, 15/−75, 30/−60, 45/−45	Raster gap	Tension, bending, impact
Espin et al.[14]	PC	constant	orientation	Young's modulus, ultimate strength, elastic constant, poission's ratio, shear modulus
Chacon et al.[15]	PLA	constant	Build orientation, layer thickness and feed rate	Tensile strength and bending strength
Mohamed et al. [16]	PC-ABS	0°, 45°and 90°	Slice height, air gap, part print direction, bead width and number of shells	Creep displacement (μm)

Fig.1: Reinforcement direction of fiber and matrix material

Table 2 Characteristic of ABS, PLA and PETG

Properties	Acrylonitrile Butadiene Styrene (ABS)	Polylactic Acid (PLA)	Polyethylene Terephthalate Glycol (PETZ)
Nozzle Temperature (°C)	245	200	245
Filament diameters (mm)	1.75	1.75	1.75
Tensile Strength (MPa)	27	37	53.08
Temperature Resistance	Resistant to high temperature.	Poor resistant to high temperature.	Poor resistant to high temperature.
Environmental Friendly / Recyclable	Recyclable, but not environmental friendly, Based on Petroleum	Environmental friendly, made from corn & sugarcane, Recyclable	Recyclable
Heated Bed Temperature (°C)	100-110	50-60	60-70

2.2 Selection of specimen geometry and 3D modeling

American society for testing and materials (ASTM) has fixed some standards for testing of materials under different tests [18)-23)]. The D638 series corresponds for tensile testing of plastic and polymers. ASTM D638 specifies methods for testing the tensile strength of plastics and other resin materials and for calculating their mechanical properties [24)]. Fig. 2 show the ASTM D638 Type-IV specimen which was fabricated through FDM printer. ASTM D638 specimen has the dog bone like structure [25)]. Specimen has shoulders at the both ends, which are wider than other portion of the structure. In this work, Autodesk Fusion @360 was used for the 3D modelling of the specimens shown in Fig. 2. 3D CAD model of specimen is shown in Fig. 3. The next section will give the fabrication method of test specimens.

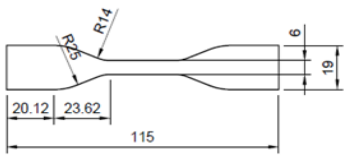


Fig.2: Drawing of ASTM D638 type-IV specimen

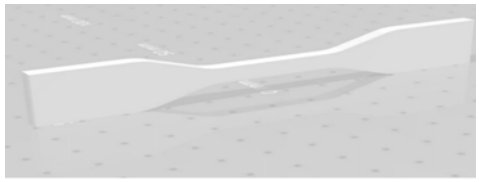


Fig.3: CAD design of specimen in .STL format

2.3 Fabrication method

Fused deposition modelling additive manufacturing technique was used in this research work for the preparation of test specimens. Three different materials were used for the fabrication of the specimen viz., acrylonitrile butadiene styrene (ABS), polylactic acid (PLA) and polyethylene terephthalate glycol (PETG). Printer used for the fabrication of test specimens was PRUSA MK3. This printer can manufacture the component of single material only. Additional machine mosaic pallet 2 was attached with PRUSA MK3 for fabrication of multi materials samples. For fabrication of specimens, all the process parameters were fixed and used at optimum levels. The process parameters used for the fabrication of samples are shown in table 3. Three different reinforcement percentages were used for the preparation of test samples viz., 15%, 25% and 35%. Fig. 4 shows the research methodology used for the improving part strength in the flow diagram. Experimental plan used in this research work is presented in Table 4. Figure 5 shows the schematic representation of dual extrusion nozzle used to fabricate test specimen. Figure 6 shows the fabricated samples, with and without reinforcement.

Table 3 Process parameters used for the sample fabrication

Process parameters	Value
Nozzle Diameter	0.4 mm
Infill Density	100%
Print Speed	60mm/s
Raster Angle	45°
Orientation	0° with respect to x axis
Layer Thickness	0.2 mm
Infill Pattern	Hexagonal

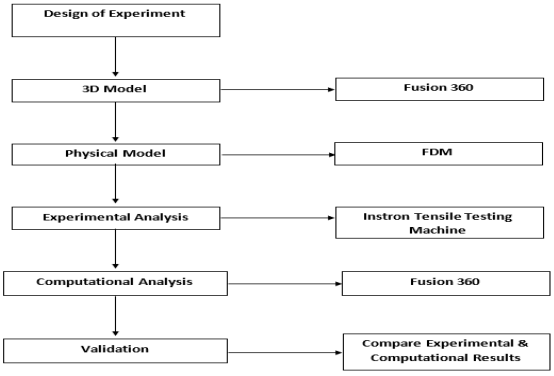


Fig.4: Flow diagram of research methodology

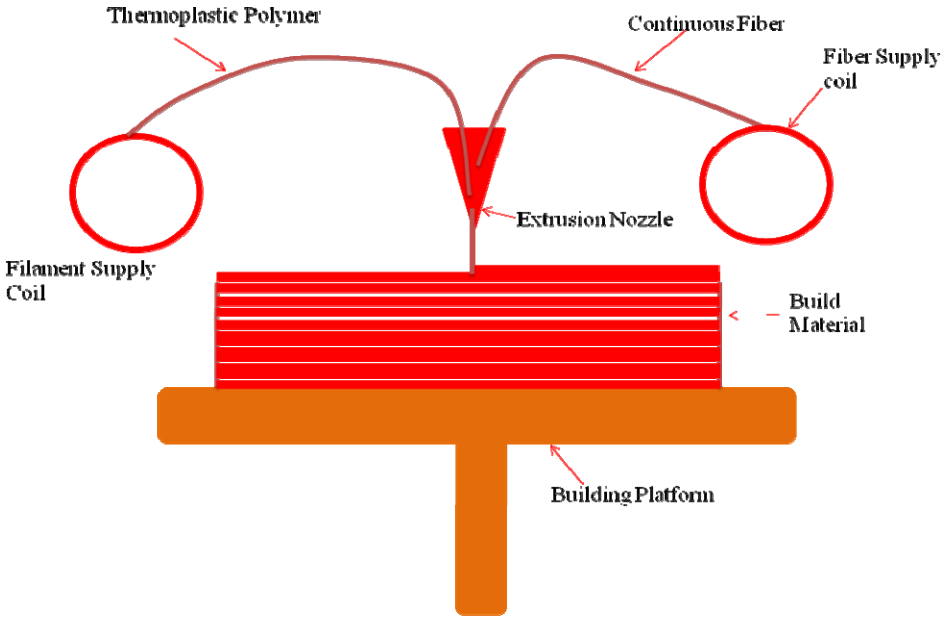


Fig. 5: Schematic layout of FDM printing process

Table 4 Experimental plan for part strength

Experiment No	Matrix material	Reinforcement fiber material	Reinforcement percentage	Tensile strength (MPa)
1	ABS	-	0%	24.05
2	PLA	-	0%	43.50
3	PETG	-	0%	49.50
4	ABS	PETG	15%	38.44
5	ABS	PETG	25%	39.42
6	ABS	PETG	35%	40.64
7	PLA	PETG	15%	45.26
8	PLA	PETG	25%	46.72
9	PLA	PETG	35%	46.85

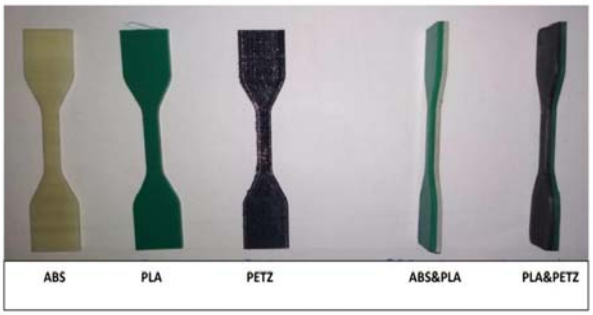


Fig.6: Fabricated specimen along the flat orientation

2.4 Experimental setup for tensile test

The tensile test of the fabricated specimen was conducted on universal testing machine Instron. The actual experimental set-up image is shown in Fig. 7. The specimen was hold between the cross head of UTM machine and axial tensile load was applied on the material at the one end. The corresponding stress-strain diagram was traced on computer according to the load. The specimens were tested at a rate of 1.5 mm/min.

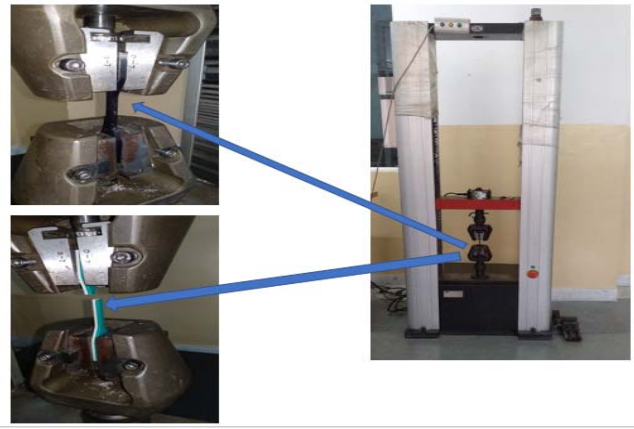


Fig.7: UTM Instron tensile testing machine

3. Computational analysis

For computational analysis, 3D CAD model of ASTM-D638 (tensile test specimen) was created by using Autodesk Fusion 360 software. The CAD model has been simulated by FEA Autodesk Fusion 360 software. A 8-node linear brick, reduced integration, hourglass control element was used for meshing of tensile specimen. After performing several mesh convergence studies 7284 elements with 1.0 seed size has been chosen for current simulation as shown in Fig.8.The bottom end of the specimen is fixed with the boundary condition Ux, Uy, Uz, Rx, Ry, Rz = 0, where Ux, Uy, and Uz are the translation in x, y and z direction respectively. Further, Rx, Ry and Rz are the rotation in x, y and z direction respectively. A velocity of 1.5 mm/min has been applied at the top surface of the specimen as shown in Fig. 9. A tie constraint has been applied to assemble the PLA and PETG in a common specimen for providing the mix effect of properties.

Total nine simulations were carried out to investigate the impact of reinforcement on mechanical properties of ABS and PLA (material). First three simulations were carried out without reinforcement to check the Von mises stress in the ABS, PLA and PETG (material) respectively. In this study, two cases were considered for the analysis. In case–I, ABS material was considered as matrix material and PETG was treated as reinforcement material. Unidirectional reinforcement material was placed between the matrix materials. The direction of reinforcement was considered in the loading direction. Three different reinforcement percentage was used i.e. 15%, 25% and 35%. In case-II, PLA was considered as matrix material and PETG (material) was considered as reinforcement material. However, the reinforcement percentage was used similar to case-I.

Material properties used for the simulation are shown in table 5. For simulation of samples without reinforcement, material was considered as single solid geometry. In case of reinforcement, two different materials system were defined; one for the matrix material and other was used for the reinforcement material

Table 5 Mechanical properties for FEA simulation

Material	Elastic modulus (E)	Shear Modulus (G)	Poisson's ratio (μ)
ABS	2110MPa	14605 MPa	0.39
PLA	30140 MPa	2310MPA	0.33
PETG	2300MPa	1640MPa	0.38

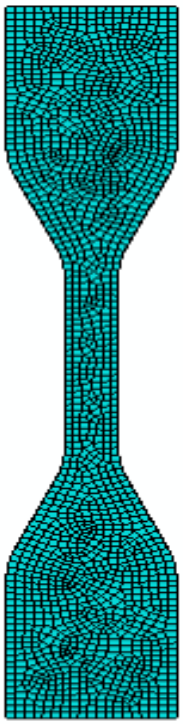


Fig.8: Mesh geometry of tensile test specimen

Fig. 9: Loading condition of test specimen

4. Results and discussion

4.1 Experimental result analysis

The average value of tensile strength for ABS, PLA and PETG materials with and without reinforcement has been calculated using 3 samples each. The effect of PETG reinforcement on tensile strength of ABS and PLA materials can be clearly seen in Fig.10. Due to reinforcement of PETG material, tensile strength of ABS and PLA material have significantly improved. Due to the uncertainty of experimental results, tensile strength is shown within the range of 5% error bar chart. The value of tensile strength of ABS material was improved by about 70% and tensile strength of PLA material was improved by about 8%.

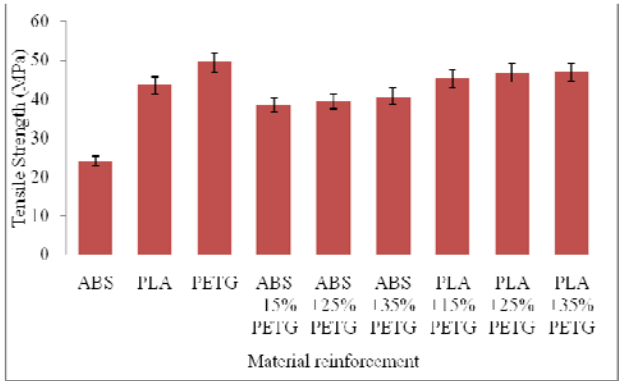


Fig.10: Effect of reinforcement on mechanical properties

4.1 Computational result analysis

Fig. 11 shows the Von mises stress of ABS and PLA materials with and without reinforcement of PETG material. It can be observed that the tensile strength of ABS and PLA materials was improved due to the reinforcement.

Results of computational study depicted that ABS samples was failed at a load of 577.2 N. The value of Von mises stress was observed 25.37 MPa at failure

point. Specimen of PLA material was failed at a load of 1044 N and value of Von mises stress was found to be 46.61 MPa. Sample of PETG material was failed at a load of 1188N and value of Von mises stress was observed 53.06 MPa. Fig. 12(a), (b) and (c) shows the Von mises stress in the specimen of ABS, PLA and PETG materials respectively. The value of Von mises stress was observed maximum in PETG material.

The effect of PETG reinforcement on mechanical properties of PLA was simulated with different reinforcement percentage. Fig. 13 shows the effect of reinforcement on Von mises stress of reinforced composite. Result shows that due to the reinforcement, the maximum value of Von mises stress for the PLA composite was found to be 50.51 MPa at 35% PETG reinforcement. It can be concluded that the tensile strength of PLA material was improved about 8% due to the reinforcement.

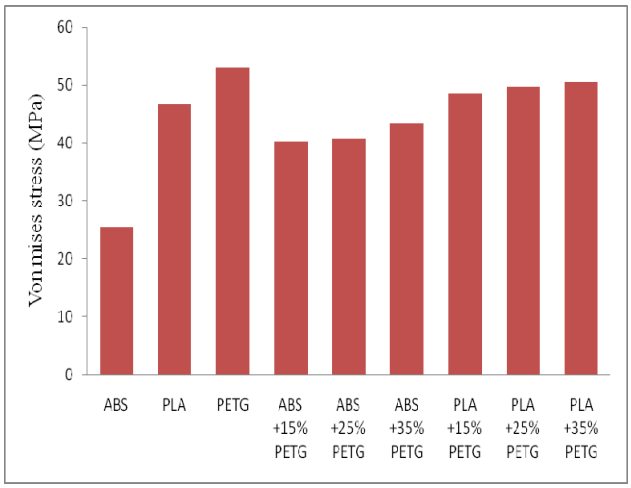


Fig. 11: Von mises stress in ABS and PLA materials

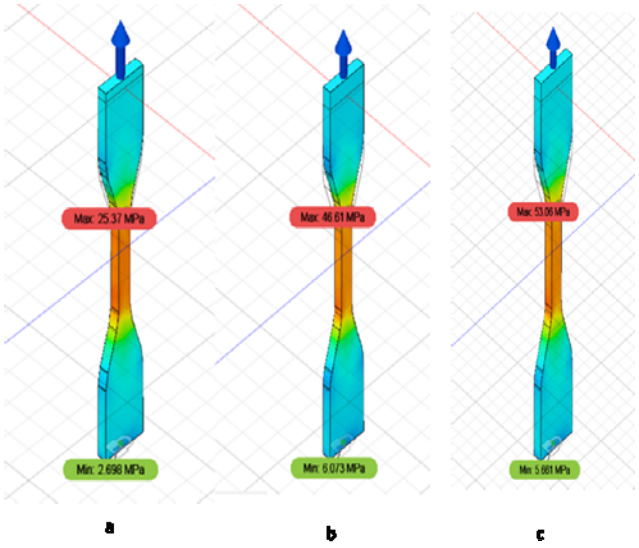


Fig.12: Von mises stress without reinforcement

(a) ABS material (b) PLA material (c) PETG material

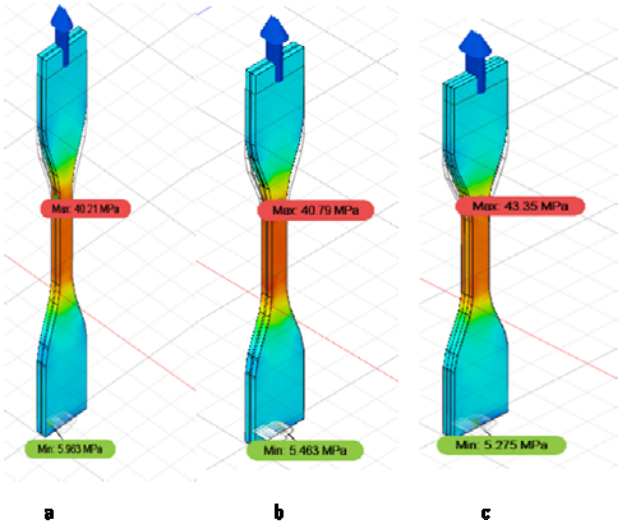


Fig.13: Von mises stress of PETG reinforced PLA composite

(a) 15% PETG (b) 25% PETG (c) 15% PETG

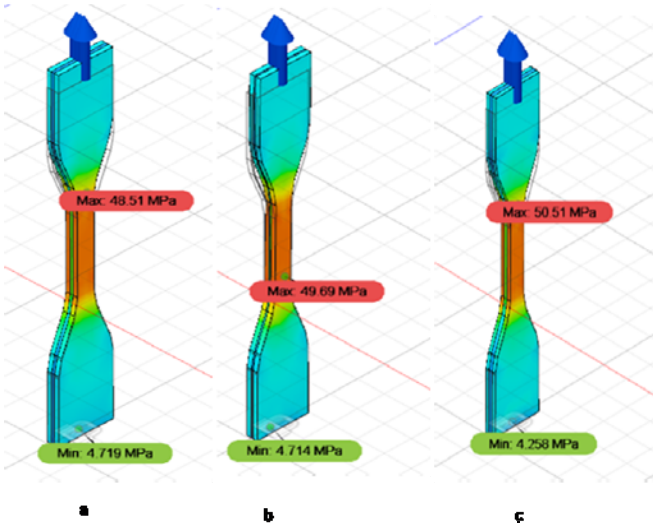


Fig.14: Von mises stress of PETG reinforced PLA composite

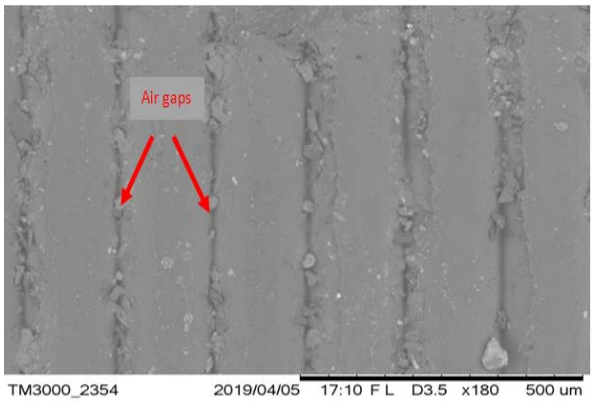
(a) 15% PETG (b) 25% PETG (c) 15% PETG

Fig.14 shows the effect of PETG reinforcement on Von mises stress of ABS material. Owing to reinforcement of PETG material, maximum value of Von mises stress of ABS was observed about 43.35 MPa at a reinforcement of 35%. Result depicted that the mechanical properties of ABS material was significantly improved. Due to the reinforcement mechanical properties of ABS was improved about 70%.

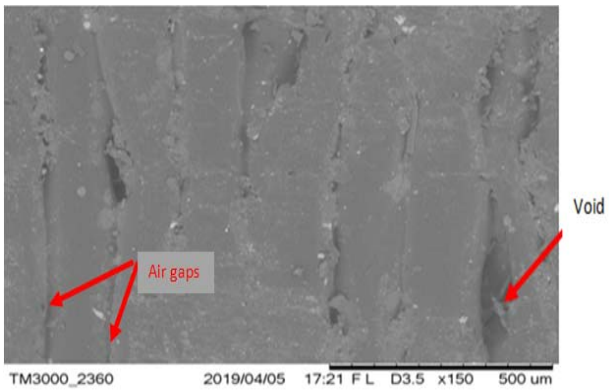
4.3 Surface morphology of fabricated specimen

For selecting the suitable orientation, surface morphology was performed. A scanning electron microscope (SEM) was used to visualize the defects in the fabricated samples. Fig.15 (a) shows the SEM image of ABS sample fabricated along flat orientation and Fig.15 (b) shows the SEM image of ABS samples

prepared along the edge orientation. It was observed that Voids were present in the component fabricated along the edge orientation, which yields poor tensile strength. For analysis of tensile test all the samples were fabricated along the flat orientation.



(a)



(b)

Fig.15: SEM image of ABS samples

(a) Flat orientation (b) Edge orientation

4.3 Comparison of experimental and computational results

Fig.16 shows the results of experimental and computational analysis. It can be concluded that results of computational analysis was close approximation to the experimental results. However, the experimental value of tensile strength was found to be less as compared to computational results. From SEM analysis, it was observed that there were some air gaps and voids present in the fabricated samples, which yields lower tensile strength.

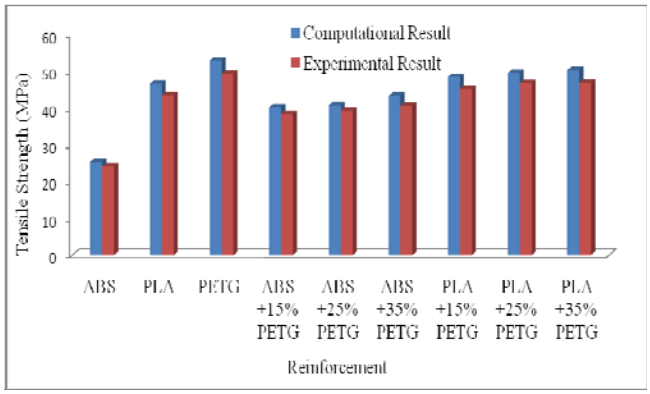


Fig. 16: Comparison of experimental and computational results

5. Conclusions from Present Work

In this study, effect of PETG (material) reinforcement on tensile strength of ABS and PLA materials were experimentally investigated. Computational analysis was also carried out in this investigation. The main conclusions of these studies are as below.

- It was depicted through surface morphology that components fabricated along the flat orientation have less defects as compared to edge orientation.
- Unidirectional reinforcement of PETG material have improved the tensile strength of ABS material by about 70% and strength of PLA material was improved by about 8%.
- From Fig. 16 it can be concluded that computational results were close approximation to the experimental results.
- However, computational results were slightly more as compared to experimental results which may be due to the presence of void and air gap between the layers of fabricated samples.

Reference

1) G. Hu, S. Cai, Y. Zhou, N. Zhang and J. Ren, 2018, "Enhanced mechanical and thermal properties of poly (lactic acid)/bamboo fiber composites via surface modification". *Journal of Reinforced Plastics and Composites*, 37(12) 797–807(2018).

2) R. Yao, Z. Yao, J. Zhou and P Liu, "Mechanical and acoustical properties of polylactic acid based multilayer-structured foam biocomposites", *Journal of Reinforced Plastics and Composites*, 35(10)785–795(2016).

3) A. R. Panes, J. Claver and A. María Camacho, "The Influence of Manufacturing Parameters on the Mechanical Behaviour of PLA and ABS Pieces Manufactured by FDM: A Comparative Analysis", *Materials, 11*(8)1333(2018).

4) S. Mansour, M. Gilbert and R.A. Hague, "A study of the impact of short-term ageing on the mechanical properties of a stereolithography resin", *Materials Science and Engineering A*, 447, 277–284 (2007)

5) O.A. Mohamed, S.H. Masood and J.L. Bhowmik, "Process parameter optimization of viscoelastic properties of FDM manufactured parts using response surface methodology", *Materials Today: Proceedings*, 4, 8250–8259 (2017).

6) W. Wu, P. Geng, G. Li, D. Zhao, H. Zhang and J. Zhao, "Influence of Layer Thickness and Raster Angle on the Mechanical Properties of 3D-Printed PEEK and a Comparative Mechanical Study between PEEK and ABS" *Materials.* 8, 5834-5846 (2015).

7) A. Lanzotti, M. Grasso, G. Staiano and M. Martorelli, "The impact of process parameters on mechanical properties of parts fabricated in PLA with an open-source 3-D printer" *Rapid Prototyping Journal* , 21(5)604-617 (2015).

8) M. Sugavaneswaran and G. Arumaikkannu, "Analytical and experimental investigation on elastic modulus of reinforced additive manufactured structure", *Materials and Design* , 66, 29–36 (2015).

9) M. F. Afrose1, S.H. Masood1, M. Nikzad and P. Iovenitti1, " Effects of Build Orientations on Tensile Properties of PLA Material Processed by FDM" *Advanced Materials Research,* 1044-1045 , 31-34. (2014).

10) I. Durgun and R. Ertan, "Experimental investigation of FDM process for improvement of mechanical properties and production cost", *Rapid Prototyping Journal,* 20 (3) 228–235(2014).

11) M.R. K. Ravari, M. Kadkhodaei, M. Badrossamay and R. Rezaei, "Numerical investigation on mechanical properties of cellular lattice structures fabricated by fused deposition Modeling". *International Journal of Mechanical Sciences*. 88, 154-161 (2014).

12) G. C. Onwubolu and F. Rayegani, "Characterization and Optimization of Mechanical Properties of ABS Parts Manufactured by the Fused Deposition Modelling Process", *International Journal of Manufacturing Engineering.* (2014). "http://dx.doi.org/10.1155/2014/598531.

13) M. Dawoud, I. Taha, S. J. Ebeid, "Mechanical behaviour of ABS: An experimental study using FDM and injection moulding techniques", *Journal of Manufacturing Processes*, 21, 39–45. (2016).

14) J. M. Chacon, M. A. Caminero, E. G. Plaza and P.J. Nunez, "Additive manufacturing of PLA structures using fused deposition modelling: effect of process parameters on mechanical properties and their optimal selection". *Materials and Design*,125,143-157(2017).

15) O. A. Mohamed, S. H. Masood and J. L. Bhowmik, "Experimental investigation of time-dependent mechanical properties of PCABS prototypes processed by FDM additive manufacturing process". Materials Letters, 193, 58-62 (2017).

16) M. Mansour, K. Tsongas, D. Tzetzis and A. Antoniadis, "Mechanical and Dynamic Behavior of Fused Filament Fabrication 3D Printed Polyethylene Terephthalate Glycol Reinforced with Carbon Fibers", *Polymer-Plastics Technology and Engineering,* 58(11) 1234-1244(2018)

17) M. Sugavaneswaran, G. Arumaikkannu, "Analytical and experimental investigation on elastic modulus of reinforced additive manufactured structure". *Materials and Design* , 66 ,29–36 (2015).

18) D. Blanco, P. F ernandez and A. Noriega "Nonisotropic experimental characterization of the relaxation modulus for PolyJet manufactured parts". *Journal of Materials Research*, 29(17) 1876-1882 (2014).

19) M. D. Espin, J. M. P. Forcada, A. A. G. Granada, J. Llumà, S. Borros and G. Reyes, "Mechanical property characterization and simulation of fused deposition modeling Polycarbonate parts". *Materials & Design*, 83,670–677(2015).

20) A. Cazon, P. Morer and L. Matey, "PolyJet technology for product prototyping: Tensile strength and surface roughness properties", *Proc IMechE Part B: J Engineering Manufacture* 228 (12) 1664–1675 (2014)

21) M. W. Garrett, S. S. Jonathon, R. D. Michael and P. C. Jason, "Evaluation of dimensional accuracy and material properties of the MakerBot 3D desktop printer". *Rapid Prototyping Journal*, 21 (5) 618 – 627(2015).

22) A. Bellini and S. Guceri, "Mechanical characterization of parts fabricated using fused deposition modeling" *Rapid Prototyping Journal*, 9(4) 252–264(2003).

23) J.M. Chacon, M.A. Caminero, E. Garcıa-Plaza, P.J. Nunez, "Additive manufacturing of PLA structures using fused deposition modelling: effect of process parameters on mechanical properties and their optimal selection. *Materials and Design*, 124, 143-157(2017).

24) H. Sosiati, Y. A. Shofie and A. W. Nugroho, "Tensile Properties of Kenaf/E-glass Reinforced Hybrid Polypropylene (PP) Composites with Different Fiber Loading". *EVERGREEN Joint Journal of Novel Carbon Resource Sciences & Green Asia Strategy*, 05 (02)1-5(2018).

25) G.W. Melenka, J.S. Schofield, Dawson, R. Jonathon and J.P. Carey, "Evaluation of dimensional accuracy and material properties of the MakerBot 3D desktop printer. *Rapid Prototyping Journal,* 21(5) 618 – 627(2016).

International Journal of Research in Engineering and Innovation (IJREI)
journal home page: http://www.ijrei.com
ISSN (Online): 2456-6934

Experimental investigation of joining of aluminum alloy 5083 by friction stir welding (FSW)

R.S. Mishra, Preety Rani

Department of Mechanical, Production, Industrial and Automobiles Engineering, Delhi technological University, Delhi, India

Abstract

This study represents the Friction stir welding of aluminum alloy 5083 with vertical milling machine as a machine tool and specially designed heat treated high carbon steel as a tool. In the present investigation, three different process parameters i.e. tool rotation speed, transverse speed and pin diameter are selected with their three different levels. The DOE by L9 array is consisted of nine trial runs. The analysis of the results obtained has been performed according to the standard procedure recommended by Taguchi. Analysis of Variance (ANOVA) is employed to analyze the influence of these parameters on % elongation during welding process

© 2019 ijrei.com. All rights reserved

Keywords: Friction stir welding (FSW), Tool rotation speed (TRS), Welding speed (WS) .Green Technology

1. Introduction

In friction stir welding a rotating tool is slowly plunged between the parting line of the two plates that need to be weld. The friction between rotation tool and the work piece produces frictional heat which softens the work piece material without reaching its melting point. The material aluminum alloy 5083 is chosen for the optimization of friction stir welding process parameter, which finds widespread application in aerospace industry, shipbuilding industry, pressure vessel and cryogenic tank in space craft.

1.1 Literature Review

Uzun et al. [1] has joined, aluminum alloy 6013-T4 with stainless steel X5CrNi18-10 by using FSW techanique. The aluminum alloy 6013-T4 plates of 4 mm thickness were welded with the plates of X5CrNi18-10 stainless steel of 4mm using FSW using milling machine. Zhang et al. [2] optimized the parameters of welding magnesium AZ31 by using friction-stir-welded to get defect-free weld. The hot rolled AZ31 (5*100* 200) mm was used as parent metal. The plates were welded by friction-stir-welding at different WS from 40 to 600mm/min and TRS 1000 RPM was same. Zhang and Zhang [3] joined two plates of Al 6061-T6 by using FSW. The tool consisting of the pin and the shoulder of radii were 5 and 22.7mm, resp. The two plates were of (100* 40 * 8) mm. Zhang et al. [4] joined the plate of AA 6061-T6 using FSW having tool of pin radius of 3mm and the shoulder radius of 7.5 mm. The welding plates were of 100*30*3 mm in length, width and thickness respectively. Kulekci et al. [5] made lap joints by FSW of AA 5754 plates by using a semi-automatic conventional milling machine. The AA 5754 plates were of thickness 3 mm, length 200 mm and width 100 mm. Fratini et al. [6] studied the flow of materials using finite element analysis (FEA), which occurs in friction welding and forms a T-joint. A thin piece of brass is used as a marker and placed on the boundary of the two necessary plates for welding. Sakthivel et al. [7] joined two plates of aluminum of dimensions (300*150*6) mm by using friction stir welding at a constant TRS of 1,000 rpm with different WS. The tool having a cylindrical threaded (1mm pitch 6mm) pin of length 5.75-mm with 15 mm shoulder dia. Zimmer and others. [8] Optimized the friction welding process parameters for force and torque generated by AA 6082-T651 with a thickness of 6 mm. Aval et al. [9] found a relationship between heat input and the microstructures of thermo mechanically affected zone of the joint of AA 5086 by friction stir welding. Jonckheere et al. [10] investigated the effect of tool and friction welding tool dimensions on fracture and shear properties at the friction stir point. Husain Mehdi et. al., [11-15], analyzed that the mechanical properties of FSW joint are mainly dependent on chemical composition and processing parameters of alloying element. The

Corresponding author: R.S. Mishra
Email Address: hod.mechanical.rsm@dtu.ac.in
http://doi.org/10.36037/IJREI.2019.3504

microstructural analysis of FSW joints shows the formation of new grain size in the weld zone with different amount of heat input by controlling the processing parameter. Rajakumar et al. [16] predicted the value of hardness and grain size of nugget region of friction stir welded AA-6061 weld joints. AA 6061 has huge applications in the production of light weight structures with good corrosion resistance and high strength/weight ratio. Liu et al. [17] carried out welding of aluminum 2219-T6 alloy using FSW and investigated effects of transverse speed on mechanical and micro structural properties. Ravindra [18] studied the effect of the tool rotation speed (rpm) and the tool probe profile on Friction Stir Processing in AA5083 of 2.5 mm thick. The experimentation was carried out using five different cylindrical, straight, square, triangular and conical pin profiles and using three tool rotation speeds that were 900, 1400 and 1800 rpm with a constant transverse speed of 16 mm / min. Yuvral et al.[19] investigate the effect of parameters and reinforcement of TiC nano particles of size 10-30nm . Shuhan et al. [20] made defect-free butt welds of titanium alloy with steel by FSW. Wwas done at TRS varied at 600 and 950 rpm with a constant weld speed of 47 mm/min. the result showed that increment in TRS increase the thickness of intermetallic compound layer. Wenya et. al [21] improve the tensile properties of magnesium AZ31 joining by friction stir welding zone by using a stationary shoulder.

2. Experimentation

A universal vertical milling machine used as the experimental machine in this study with a fixture used for clamping the workpiece with locking nut. The size of table is 254×1370 mm on which a fixture or backing plate is mounted which is used to hold the work-piece during welding. Two aluminium alloy 5083 plates of size 75mm×60mm×6mm are mounted on the fixture of vertical milling machine for making butt joint by using friction stir welding. The parameters which selected for this investigation are: tool rotation speed, transverse speed and tool pin diameter. Various experiments are conducted for optimize the response parameter i.e. %Elongation. Taguchi's robust design of experiments (DOE) methodology is used to plan the experiments statistically. Table 1 details the different control variables and their levels.

Table 1: Control variables and their levels

Parameters	Level 1	Level 2	Level 3
Tool rotation speed	1200 rpm	1950 rpm	3080 rpm
welding speed	20mm/min	25mm/min	30mm/min
Tool Pin diameter	5mm	6mm	7mm

All the selected parameters have equal levels. L9 and L27 orthogonal array has come out as the possible solutions for designing the experiments. Hence, L9 array was selected for the present investigation. The DOE by L9 array is consisted of nine trial runs as shown in table2.

The welding of two plates of AA 5083 is done according to the entire nine trail run. The joined plates after welding are shown in fig.1.

Figure 1: Plate after friction stir welding

Then two tensile test specimens are extracted from each welded piece. The dimensions of the test specimens are according to American Society for Testing of Materials (ASTM) standard and after extract from the welded plates are shown in fig.2**.** The values of % elongation were recorded and plotted as per Taguchi design of experiments methodology. Table.3 & Table .4 shows the values of % elongation against the input parameter setting for L9 orthogonal array.

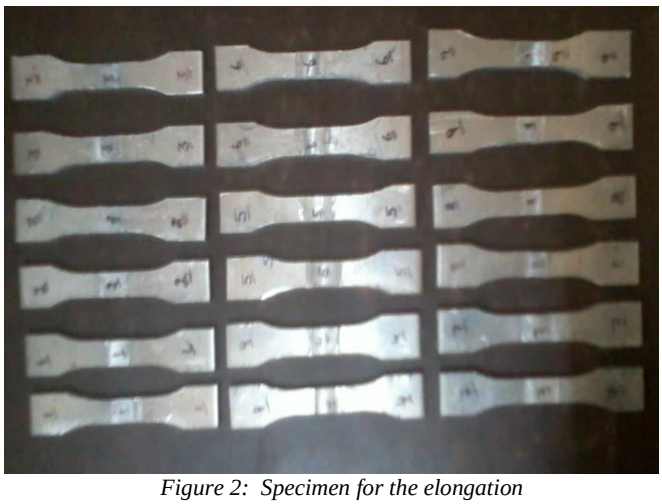


Figure 2: Specimen for the elongation

The analysis of the results obtained has been performed according to the standard procedure recommended by Taguchi. The main effects are studied by the level average response analysis of % elongation raw data or of S/N data. The analysis is done by averaging the % elongation data and/or S/N data at each level of each parameter and plotting the values in graphical form. The factor effects of different input parameters with their corresponding levels on S/N ratio and average response for % elongation are tabulated in table-3 & table 4 respectively.

Table-2: Taguchi's L9 Standard Orthogonal Array for % Elongation

Experiment No.	Design of Experiment			% Elongation Sample 1	%Elongation Sample 2	Average % Elongation
	Tool rotation speed (A)	Welding speed (B)	Tool pin diameter (C)			
1	1200	20	5	3.9	3.95	3.92
2	1200	25	6	3.74	2.98	3.36
3	1200	30	7	5.39	5.19	5.28
4	1950	20	6	4.41	4.4	4.4
5	1950	25	7	5.44	5.44	5.44
6	1950	30	5	7.3	7.27	7.28
7	3080	20	7	2.3	2.35	2.32
8	3080	25	5	2.67	2.77	2.72
9	3080	30	6	3.18	2.9	3.04

Table.3: Response Table for % Elongation (Means)

Level	A	B	C
1	4.187	3.547	4.640
2	5.707	3.840	3.600
3	2.693	5.200	4.347
Delta	3.013	1.653	1.040
Rank	1	2	3

Table.4: Response Table for % Elongation (S/N Ratio)

Level	A	B	C
1	-12.282	-10.682	-12.600
2	-14.941	-11.310	-11.018
3	-8.553	-13.784	-12.158
Delta	6.388	3.103	1.582
Rank	1	2	3

Table..5: Analysis of Variance for % Elongation (Means)

Source	DF	Seq SS	Adj SS	Adj MS	F	P	%Age contribution
A	2	13.6206	13.6206	6.81031	195.4	0.005	67.78
B	2	4.6692	4.6692	2.33458	67.00	0.015	23.24
C	2	1.7252	1.7252	0.86258	24.76	0.039	08.58
Resi. Error	2	0.0697	0.0697	0.03484			
Total	8	20.0846					

2.1 Analysis of Variance (ANOVA)

The purpose of ANOVA is to investigate the parameters, whose combination to total variation is significant. The ANOVA for Means of %elongation are given in table-5.

3. Results and Discussion

The main effect of input parameters on the response parameters (% elongation) are discussed on the basis of mean of response, ANOVA and plots between response parameters and input parameters. The plots show that how response parameters are varying at different levels of input parameters. The plots for mean and S/N ratio for % elongation are shown in fig. 3-4. It can be observed from fig. 3 that the Tool rotation speed, welding speed and tool pin diameter affects the % elongation very significantly. Moreover, the different input parameters used in the experimentation can be ranked in the order of decreasing %elongation as tool rotation speed, welding speed and tool pin diameter.

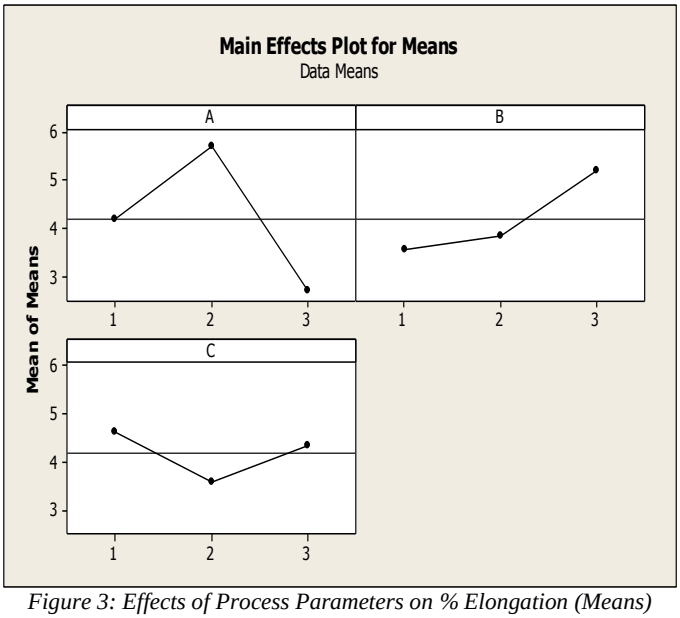


Figure 3: Effects of Process Parameters on % Elongation (Means)

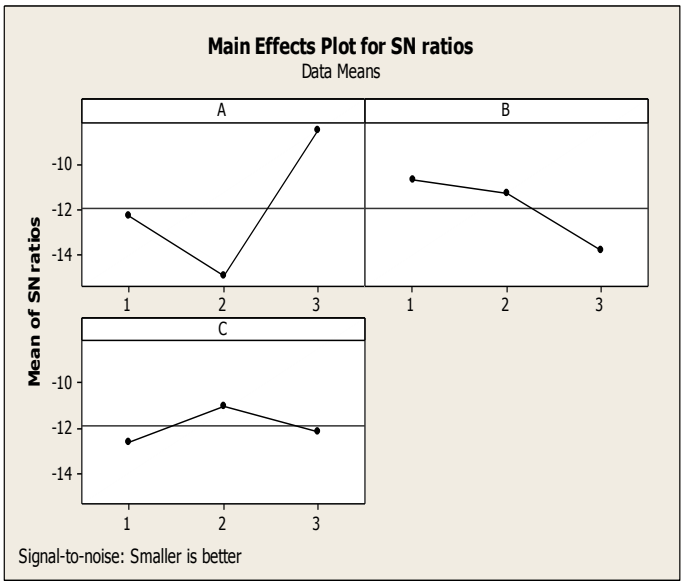


Figure 4: Effects of Process Parameters on % Elongation (S/N Ratio)

4. Conclusion

Following conclusions were drawn

1. In Friction stir welding, the Tool rotation speed is most significant factor for decreasing % elongation, welding speed is the second significant factor and Tool pin diameter is the third significant factor.
2. The most effective parameter with respect to %Elongation is tool rotational speed. Whereas the effect of welding speed and tool pin diameter on the %elongation was also significant.
3. By increasing the tool rotation speed, the %elongation first increases and then decreases. By increasing the welding speed, the %elongation is increased. By increasing the tool pin diameter, %elongation is first decreases and then increases increased.
4. The Computed results using ANOVA also showed that the tool rotational speed is the most effective factor with a percent contribution of 67.78%. The percent contribution tool shoulder dia. and welding speed is 23.24% and 08.58% respectively.
5. The input parameters settings tool rotation speed at 3080 rpm, welding speed at 20 mm/min, and tool shoulder dia. at 6 mm have given the optimum results for %elongation; in friction stir welding on AA 5083.

References

[1] HuUzun, H., Donne, C.D., Argagnotto,A., Ghidini,T., & Gambaro,C. (2005). Friction stir welding of dissimilar Al 6013-T4 to X5CrNi18-10 stainless steel. Materials and Design, 26, 41–46.

[2] Zhang, H., Lin, S., B., Wu, L., Feng, J., C. & Ma, S.L. (2006). Defects formation procedure and mathematic model for defect free friction stir welding of magnesium alloy. Materials and Design, 27, 805–809.

[3] Zhang, Z. & Zhang, H. W. (2007). Material behaviors and mechanical features in friction stir Welding process. Int J Adv Manuf Technol, 35, 86–100.

[4] Zhang, H., W., Zhang, Z., & Chen, J., T. (2007). 3D modeling of material flow in friction stir welding under different process parameters. Journal of Materials Processing Technology, 183, 62–70.

[5] Kulekci, M., K., Şik, A. & Kaluç, E. (2008). Effects of tool rotation and pin diameter on fatigue properties of friction stir welded lap joints. Int J Adv Manuf Technol, 36,877–882.

[6] Fratini, L., Buffa, G., Micari, F. & Shivpuri, R. (2009). On the material flow in FSW of T-joints: Influence of geometrical and tecnological parameters. Int J Adv Manuf Technol, 44, 570–578.

[7] Sakthivel, T., Sengar, G., S. & Mukhopadhyay, J. (2009). Effect of welding speed on microstructure and mechanical properties of friction-stir-welded aluminum. Int J Adv Manuf Technol, 43,468–473.

[8] Zimmer, S. Langlois, L., Laye, J. & Bigot, R. (2010). Experimental investigation of the influence of the FSW plunge processing parameters on the maximum generated force and torque. Int J Adv Manuf Technol, 47,201–215.

[9] JamshidiAval, H., Serajzadeh,S. & Kokabi, A. H.(2010). Theoretical and experimental investigation into friction stir welding of AA 5086. Int J AdvManufTechnol, (2011), 52,531–544.

[10] Jonckheere,C., Meester,B.D., Cassiers,C., Cassiers,M. & Simar,A.(2011). Fracture and mechanical properties of friction stir spot welds in 6063-T6 aluminum alloy.Int J AdvManufTechnol,.(2012), 62,569–575.

[11] Husain Mehdi, Mishra. R.S. (2016). "Mechanical properties and microstructure studies in Friction Stir Welding (FSW) joints of dissimilar alloy – A Review." Journal of Achievements in Materials and Manufacturing Engineering 77 (1) pp.31-40.

[12] Husain Mehdi, Mishra. R.S. (2017). "Influences of Process Parameter and Microstructural Studies in Friction Stir Welding of Different Alloys: A Review." International Journal of Advanced Production and Industrial Engineering, SI-MM 509, pp. 55–62.

[13] Husain Mehdi, R.S. Mishra, Mechanical and microstructure characterization of friction stir welding for dissimilar alloy- A Review, International Journal of Research in Engineering and Innovation, vol-1, issue-5, 57-67.

[14] Husain Mehdi, R.S Mishra, Analysis of Material Flow and Heat Transfer in Reverse Dual Rotation Friction Stir Welding: A Review, International Journal of Steel Structures, 2018, doi: 10.1007/s13296-018-0131-x.

[15] Husain Mehdi, R.S. Mishra, Study of the influence of friction stir processing on tungsten inert gas welding of different aluminum alloy, SN Applied Sciences, 1: 712. https://doi.org/10.1007/s42452-019-0712-0.

[16] Rajakumar,S., Muralidharan,C. & Balasubramanian,V.(2011). Statistical analysis to predict grain size and hardness of the weld nugget of friction-stir-welded AA6061-T6 aluminium alloy joints. Int J AdvManufTechnol, (2011), 57,151–165.

[17] Liu, H. J., Li,J.Q. & Duan,W.J.(2012). Friction stir welding characteristics of 2219-T6 aluminum alloy assisted by external non-rotational shoulder. Int J AdvManuf Technol.

[18] Thube, R.S. (2014). Effect of Tool Pin Profile and Welding Parameters on Friction Stir Processing Zone, Tensile Properties and Micro-hardness of AA5083 Joints Produced by Friction Stir Welding. International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-3, Issue-5.

[19] Rashid, M., Yuvraj, N., Singh, S.,K., (2016) "Micro Structural and Mechanical Behaviours of Nano-TiC-Reinforced AA6082 FSW Joints". V th International Symposium on "Fusion of Science & Technology, New Delhi, 406-411.

[20] Li, S., Chen, Y., Zhou, X., Kang, J., Huang, Y., Deng, H. (2019). "High-strength titanium alloy/steel butt joint produced via friction stir welding". Materials Letters 234 (2019) 155–158.

[21] Li, W., Niu, P.,L., Yan, S.,R., Patel, V., Wen, Q. (2019). "Improving microstructural and tensile properties of AZ31B magnesium alloy joints by stationary shoulder friction stir welding". Journal of Manufacturing Processes 37 (2019) 159–167.

Cite this article as: R.S. Mishra, Preety rani, Experimental investigation of joining of aluminum alloy 5083 by friction stir welding (FSW), International Journal of Research in Engineering and Innovation Vol-3, Issue-5 (2019), 306-309, http://doi.org/10.36037/IJREI.2019.3504

Journal Pre-proofs

Experimentation and Modelling of CNT additive abrasive media for micro finishing

Parvesh Ali, R.S. Walia, Q Murtaza, M.S. Ranganath

PII: S0263-2241(19)30999-6
DOI: https://doi.org/10.1016/j.measurement.2019.107133
Reference: MEASUR 107133

To appear in: *Measurement*

Received Date: 9 March 2018
Revised Date: 24 September 2019
Accepted Date: 4 October 2019

Please cite this article as: P. Ali, R.S. Walia, Q. Murtaza, M.S. Ranganath, Experimentation and Modelling of CNT additive abrasive media for micro finishing, *Measurement* (2019), doi: https://doi.org/10.1016/j.measurement.2019.107133

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier Ltd.

Journal Pre-proofs

Experimentation and Modelling of CNT additive abrasive media for micro finishing

Parvesh Ali[a], R S Walia[b*], Q Murtaza[c*], Ranganath M.S[d]

[a] Research Scholar, Department of Mechanical Engineering, Delhi Technological University, Delhi
[b,c,d] Professor, Department of Mechanical Engineering, Delhi Technological University, Delhi
[b] Professor, Deptt of Production and Industrial Engineering, Punjab Engineering College, Chandigarh

ARTICLE INFO

Keywords:

AFM

ANSYS

Fluent

CNTs

Material removal

Modelling

ABSTRACT

Abrasive flow machining (AFM) is a micro finishing methodology applied to achieve complex geometries which are not easy to obtain using other conventional techniques. AFM removes the material in form of small chips using abrasion mechanism. It gives good surface finishing but has a limitation of low material removal. To overcome the limitation, various hybridizations of AFM process have been developed. AFM process parameters such as pressure, velocity of media and temperature distribution on work piece play a major role in micro finishing. In order to enhance the AFM process capability, it is required to analyze the critical conditions of AFM process parameters using any commercially available software's. Therefore, in present study, three dimensional modelling of the media flow path in abrasive flow machining was developed and analyzed to obtain the effect of the parameters using ANSYS® and Fluent® software. Further to overcome the major limitation of the AFM process that is low material removal, an experiment was conducted using abrasive media with carbon nano tube (CNT) and alumina. To enhance productivity, AFM process was optimized using Taguchi L_9 OA. Also, the work piece wear during AFM process obtained from the experiments was compared with Reye–Archard–Khrushchov wear law.

Nomenclature

AFM	Abrasive Flow Machining		Flow Machining
BP-RBFN	Back-propagation Radial basis functions network	GA- RBFN	Genetic algorithms Radial basis functions network
CFAAFM	Centrifugal force assisted Abrasive Flow Machining	UAAFM	Ultrasonic assisted Abrasive Flow Machining
		CVD	Chemical Vapor Deposition
ECA^2FM	Electrochemical assisted Abrasive		

***Corresponding author.**
E-mail address: waliaravinder@yahoo.com, qasimmurtaza@gmail.com

Journal Pre-proofs

Journal Pre-proofs

1. Introduction

It is a well known fact that finishing of a machine part / component / product generally improves its aesthetics and functional performance i.e. the fine finished parts have better dimensional controls, endurance strength and thus more useful life. Merchant [1] has analyzed the need and trends of future of manufacturing technology with assumption that past and present manufacturing activity can be conveniently mixed. In recent past, due to requirements of the better quality product with high precision, there was a need to develop close dimensional control along with increasing complexity in their shapes. For such finishing requirement, abrasive flow machining (AFM) process has evolved and since its invention, it has been convincingly used for the finishing of complex shapes product [2]. Among all non conventional finishing processes, AFM is most suitable to produce economically better finished surface in the complex geometric shapes. In AFM metal removal mechanism, extrusion of abrasives laden media (made up of a liquid polymer with abrasive particles) through a controlled passage is done that results in abrasion of the required surfaces. The abrasives are oriented randomly that act as flexible grinding tools. Abrasives laden medium is prepared with abrasives and polymer, has low viscosity and good abrading capability [3]. It removes the material layer thickness in the range of 1-10μm and can produce best surface finish achieved up to 50 nm [4].

The general conclusion drawn from the experiments reported in the literature review [4-25], is that the extrusion pressure should be high for better surface finish. If low extrusion pressure is used with low abrasive concentration and smaller grain size, the surface finish will be decreased. If low extrusion pressure is used with larger grains, it causes uneven distribution that result in variation in the surface finish of work pieces. So, for better surface finish optimum combination of extrusion pressure, grain mesh size and abrasive concentration is required [5].

AFM process has a major limitation of low material removal and also it cannot be used for large surface irregularities because it removes material uniformly from all surfaces. Therefore, imperfections in the shape cannot be improved [6]. Various researchers have hybridized AFM process with other non conventional processes to enhance the material removal by using centrifugal force, magnetic force, vibration, chemical action and so on in the media flow path. Walia et al. [7] used centrifugal force in the media flow path of AFM to reduce amount of cycles for removing the same material in comparison to simple AFM process. When centrifugal force is added in the simple AFM process, dynamic abrasive grains in the media is increased. The experimental results showed that CFAAFM produces no affect on the surface layers of work piece during finishing and also produced better surface finish in comparison to simple one way or two way AFM process after specific number of cycles [8-10]. Sushil et al. [11] used AFM nano finishing process to finish EDM generated surface. The result showed that extrusion pressure and material of work piece were the significant parameters in the finishing process and also provided better finish by removing the cracks and recast layer produced during the EDM process. Wang et al. [12] studied three process parameters during the experimentation viz. space between the helical edge and the work surface, helical slot thickness and amount of helical turns. The results show that the surface and helical edge space, amount of helical turns plays a crucial role in the

material removal. Sambharia et al. [13] proposed a unidirectional set up of AFM process and found that finishing time and viscosity of media provided a better finishing performance during the machining process. Venkatesh et al. [14] used Ultrasonic assisted AFM process to finish the bevel gear made of EN 8 material using the input variable as frequency of vibration, extrusion pressure, operating time and flow rate of media and observed that use of vibration during the machining removed higher peaks of the work surface in comparison to the other processes. Brar et al. [15] hybridized AFM process with ECA^{2}FM that enhanced the material removal by producing the electrochemical dissolution effect. Marzban et al. [16] tried to apply a spin motion along with the rotation of the work piece and found significant change in the material removal rate. Wan et al. [17] used zero order semi mechanistic approach for two way AFM process to decrease the time required for machining and found that using zero order methodology would be better for internal passages in Abrasive Flow Machining. Types of abrasive laden media also have an influential effect on material removal rate. SiC based media is found better in comparison to Al_2O_3 based media in terms of material removal rate [18]. A highly concentrated abrasive laden media corresponds to more viscous media which improves the surface roughness due to better abrasive holding capacity [19]. Mali et al. [20] reported that mesh size of abrasive particles was the most effective parameter for material removal. Mollah et al. [21] compared two algorithms for prediction of outputs and found better results in case of GA-trained RBFN as compared to BP- trained RBFN. A simulation analysis was carried out on ANSYS FLUENT software to calculate the material removal which was dependent upon the axial and radial stress. Seifu et al. [22] found that velocity of media was maximum at the converging section and was minimum at the walls of the work piece. Venkatesh et al. [23] observed from the computational approach that the media flow in UAAFM process did not follow the axial path as in conventional AFM, while it provided better axial movement to abrasives in UAAFM process. Chen et al. [24] demonstrated a simulation how to find the optimum design of polygonal passageways, having four helix core and found that polygonal passageways produce more deviation in strain rate in comparison to the helical passageways. Fu et al. [25] observed irregular streamlines near trailing edge during the movement in axial direction and also observed a huge pressure reduction towards the flow. Wei et al. [26] proposed a new media based on guar gum hydrogel having better machining capability at higher shear strain rate conditions. From the experimental results it was found that guar gum based hydrogel media provided better surface finish in comparison to styrene butadiene based media under shear thickening condition. Bouland et al. [27] used combination of numerical and experimental approach to predict the results of the finishing process. It was found that simulation and experimental results were very close for the material removal. Han et al. [28] finished 15-5 PH stainless steel internal channel surfaces using AFM process and studied about the residual stress profiles generated over the surface. It was observed that this process generates compressive residual stress over the surface because temperature rise in AFM process was not comparable to grinding and turning. Lv et al. [29] numerically investigated the erosion mechanism in hard brittle materials. From the results it was observed that the radial and lateral cracks were the main mode of fracture during the finishing process. Wang et al. [30] finished different types of coated and uncoated milling cutter through Abrasive Flow Machining process to

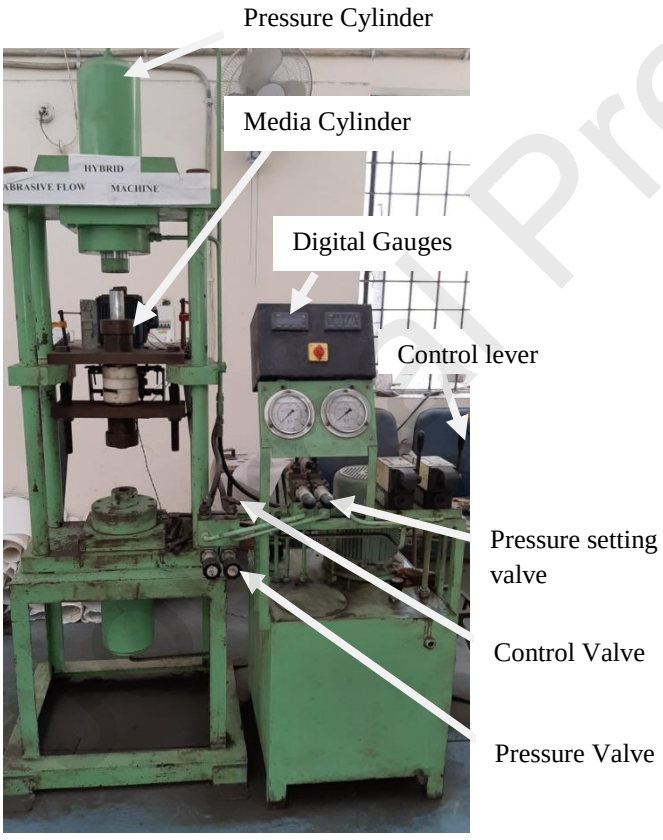
Journal Pre-proofs

sharpen its cutting edges. The researcher found that micro sized diamond films raised the roughness of both rake and flank faces and reduced the radius of cutting edges. Shao & Cheng [31] analyzed the surface roughness and topography profile generation in AFM process and found that abrasive media volume creates largest difference in the machining system.

In present research, CNT additive with alumina has been used as abrasive media to study the critical parameters of AFM such as pressure, velocity and temperature distribution of media on the work piece surface through simulation in ANSYS® and Fluent® software. The material removal obtained from both experimental as well as computational model has also been compared. Further AFM process was also optimized using Taguchi L_9 OA.

2. Materials and Methods

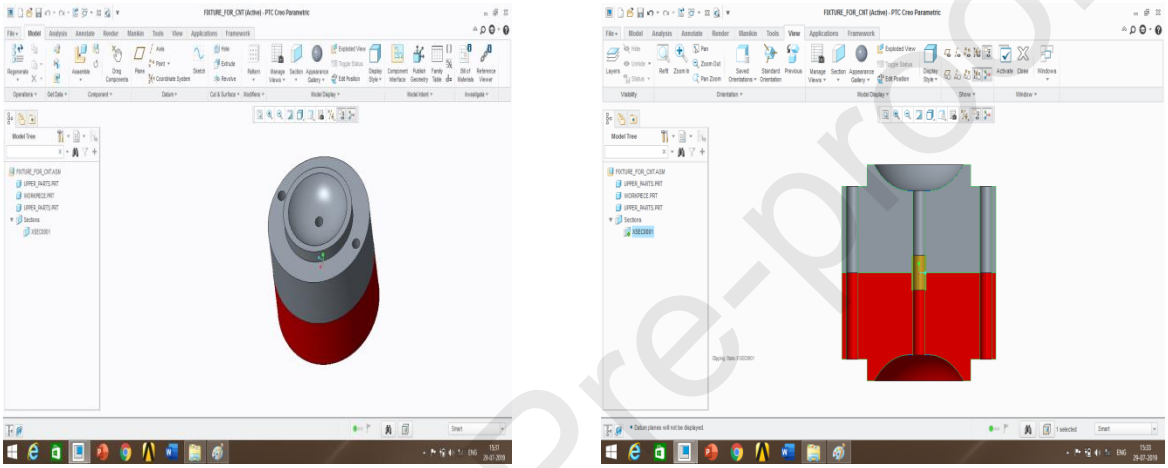
The two way AFM setup was used, which pressurizes the abrasive media to flow through the internal cylindrical surface [32, 33, 34]. The AFM components include hydraulic power pack, hydraulic cylinders, media cylinders, nylon fixture, machine frame (Fig. 1 (a) & (b)). The AFM used a non-Newtonian liquid polymer containing nano tubes CNT particles along with aluminium oxide abrasive particles as shown in fig. 2. The Aluminium work pieces used in the experiment are as described in fig. 2 & 3.



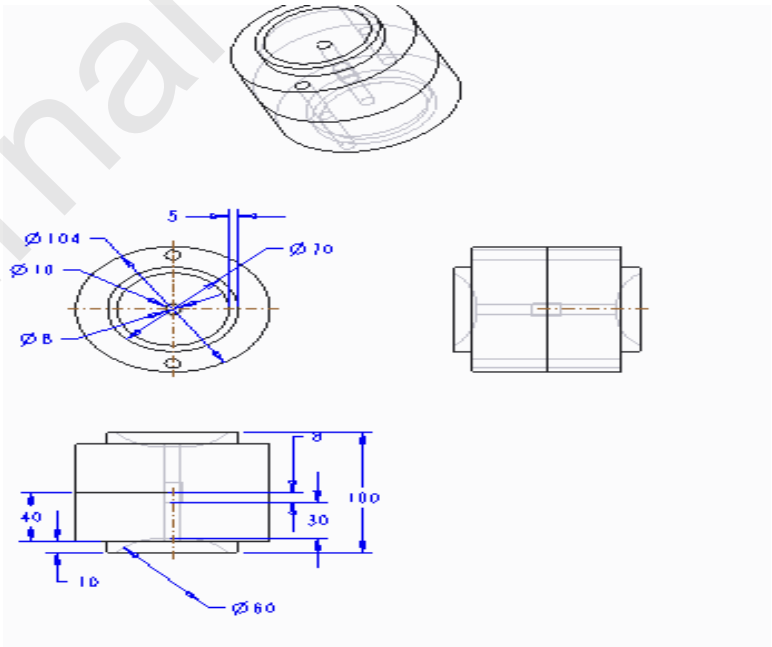
1(a) Figure shows Two way AFM components

1(b) Figure shows Two way AFM set up during operation along with fixture

1(c) Arrangement of Fixture



1(d) Model of Fixture used in Abrasive Flow Machining



1(e) Drawing of Fixture

Fig. 1 Two way developed AFM set up with fixture design

Journal Pre-proofs

Fig.2 Shown Aluminium work piece and CNT based media used for finishing

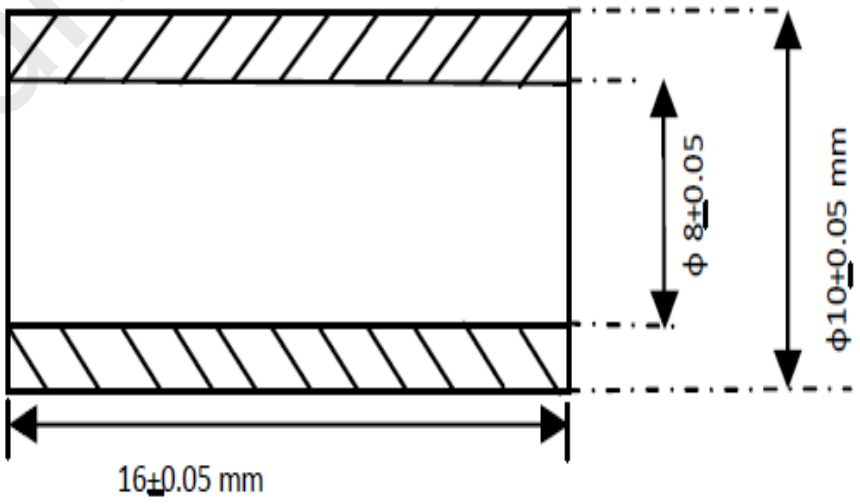


Fig. 3 Specification of work piece used for this study [35]

2.1 Material Removal

Material removal signifies the material removed from the specimen in a specified to and fro motion of the media. Material removal was calculated from the formula as follows:

MR= weight of specimen prior machining - weight of specimen after machining

2.2 (% improvement in R_a)

% improvement in R_a = (Initial Roughness - Final Roughness)/(Initial Roughness) * 100

DOE was used to identify the effects of the parameters being taken on response characteristics (material removal & % improvement in surface roughness) of AFM process. Taguchi parametric design methodology was adopted. The experiments were performed by using orthogonal array (OA) of L_9 standard with 3-levels of orthogonal array and consists 8 degree of freedom. The process parameter affecting the performance and their levels are given in the Table 1.

Table 1 Process Parameters and their values at different levels

Symbol	Process Parameters	Unit	L1	L2	L3
P	Extrusion Pressure	MPa	11	22	33
C	Amount of CNT	Gram	With-out CNT	5	10
N	No. of Cycles	Number	3	6	9
Polymer to Gel ratio: 1:1, Work piece Material- Aluminium, Abrasive type- Al_2O_3& CNT in the ratio (1:2), Al_2O_3 Mesh Size: 180, Media Flow Volume: 290 cm^3, Temperature: 32± 2 °C, Initial surface roughness: 1.1 - 2.96 microns					

The three process parameter pressure, amount of CNT, no. of cycles was selected as shown in Table 1. Experiments were performed according to the standards described by the L_9 OA. Every experiment was performed for three times. Because of 9 operations performed three times each, twenty seven test specimens were selected whose initial surface was in the range of (1.1 – 2.96 micron). For each sample operation and its replications, the material removal and % improvement in R_a was measured and has been shown in Table 2.

Table 2 Experimental Results of Material Removal as per Taguchi L_9 OA

E.N	Run order	% Improvement in R_a			S/N ratio (db)	Material Removal (MR) (mg)			S/N ratio (db)
		R1	R2	R3		R1	R2	R3	
1	1	7.23	8.32	8.76	18.09	0.9	0.4	0.3	-7.92
2	4	16.1	17.1	17.88	24.60	0.7	0.6	0.2	-9.97
3	7	15.5	16.85	15.60	24.05	0.7	0.8	0.6	-3.27
4	2	16.43	17.95	17.23	24.69	0.5	0.6	0.4	-3.28

5	5	29.04	31.86	31.90	29.78	1.2	1.1	1.3	1.52
6	8	20.84	23.74	28.30	27.51	1.5	1.9	1.8	4.64
7	3	9.55	10.44	10.85	20.20	0.6	0.5	0.6	-5.03
8	6	13.94	14.87	13.31	22.92	0.9	0.7	0.8	-2.08
9	9	14.51	14.25	15.66	23.39	0.7	0.9	0.8	-2.07
Total		143.14	155.38	159.49		7.7	7.5	6.8	
		T ΔR_a = Overall mean of ΔR_a=16.96%				TMR =Overall mean of MR =0.81mg			

3. Modelling of 3- dimensional media flow path in AFM process

Modelling of Abrasive flow machining was carried out to know the critical parameters (pressure, velocity of media and temperature distribution on work piece surface) during the machining to increase the efficiency of the process. Also this analysis was used further to predict the wear using the Reye–Archard–Khrushchov wear law. The equation given by the law is as follows [36]:

$$Q = \frac{K.W.L}{H}$$

Where, Q = Total volume of the wear debris

K= Dimensionless Constant

W= Total Normal Load

L= Sliding distance

H= Hardness of the softest contacting surfaces

The wear obtained from the computational analysis was compared with the experimental results.

Sachin et al. [37] performed FEM modelling to obtain the finishing forces and found experimentally that the process was capable of producing mirror like surfaces.

A 3-dimensional model of Abrasive Flow Machining with the flow domain and boundary conditions is shown in fig. 4. Assumptions made during the modelling of abrasive flow machining, are as following-

- This model is pressure based model.
- The analysis is based on the Green Gauss cell based.
- The formulation used is absolute velocity and implicit.
- The media flows in a 3- dimensional path.
- The media flows symmetric to the path of hollow surface.
- Media is homogeneous and isotropic.
- The flow of media is steady in nature.
- Flow is laminar.
- There is sliding motion between the abrasive particle and work piece surface.
- It is assumed that the area of asperities is circular in shape.
- Abrasive particles will completely pass the asperities.
- Wear occurs in hemispherical shape.
- Each asperity will not participate in the material removal.

Journal Pre-proofs

Fig. 4(d) and (e) shows the top view and cross-sectional view of the model which includes abrasive particles, CNT particles and work piece which is taken as aluminium in this analysis. Properties taken for CNT, Polyborosiloxane, alumina and aluminium work piece during analysis are given in Table 3.

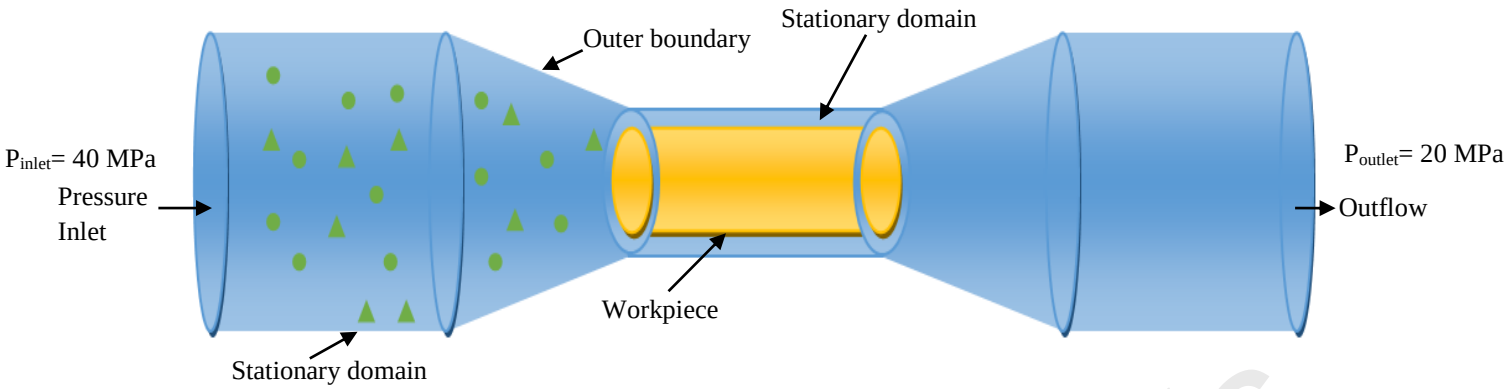
Table 3 Properties taken for Analysis

CNT density	1340 kg/m^3
CNT thermal conductivity	3000 W/m-K
CNT strength	40 GPa
Polyborosiloxane viscosity	414 kg/m-sec
Density of Polyborosiloxane media	1219 kg/m^3
Polyborosiloxane thermal conductivity	0.22 W/m-K
Density of Alumina	3890 kg/m^3
Specific heat of Alumina	451 J/kg-K
Thermal conductivity	12 W/m-K
Density of Aluminium	2700 kg/m^3
Elastic Modulus of Aluminium	69 GPa
Specific heat of Aluminium	900 J/kg-K
Thermal conductivity of Aluminium	205 W/m-K

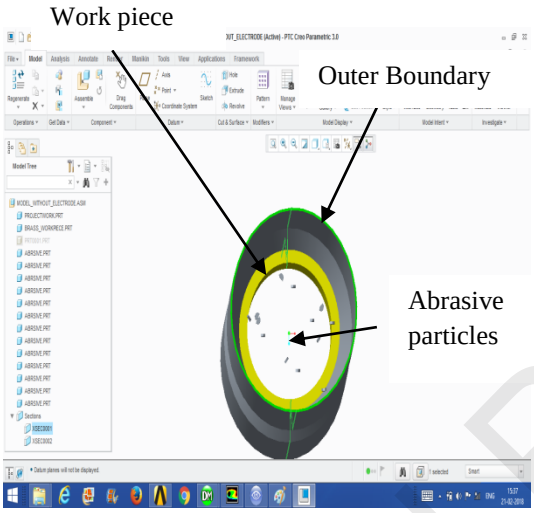
To analyze the variables in AFM, Fluent software® was also used for final analysis of results. Figure 4 (c) shows the flow domain and boundary conditions used in the analysis. Fluid inlet and outlet was selected and then meshing was done in ANSYS® 15 as shown in figure 4 (f). After meshing, it was imported to fluent software® for further analysis. For analyzing the model, the media was taken as the mixture of polyborosiloxane and alumina abrasive along with CNT particles. The applied pressure was 40 MPa on the inlet side and at the outlet side pressure was 20 MPa because the media flow in the abrasive flow machining was due to pressure difference between the upper and lower cylinder. For the media movement from the upper to lower chamber, upper cylinder was maintained at high pressure and lower cylinder was maintained at low pressure.

4(a)

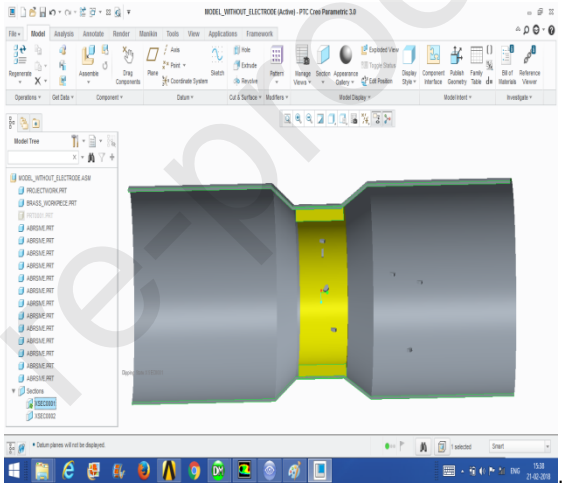
4(b)



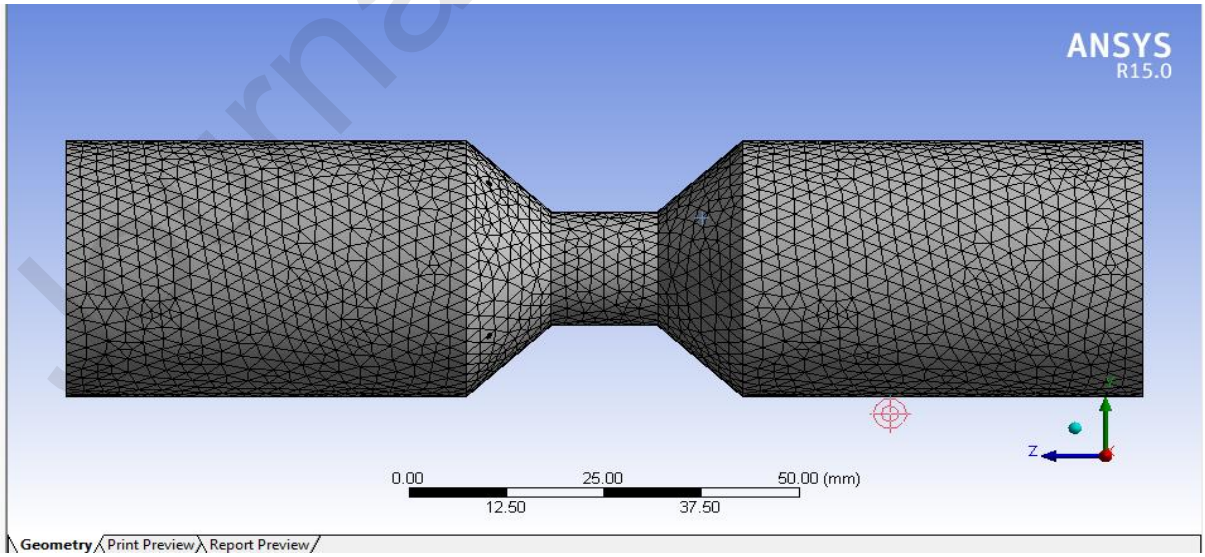
4(c)



4 (d)



4(e)



4(f)

Figure 4 (a), (b) Shown media flow path inside the fixture (c) Flow domain and boundary conditions (d), (e) 3D Model of Abrasive Flow Machining (f) Meshing of Abrasive Flow machining model

Journal Pre-proofs

4. Calculation of Total Wear debris

If it is assumed that the asperities are circular in shape then area of asperity will be = $\pi.a^2$

Where a = radius of circular asperity

Acting Load W = P. $\pi.a^2$. N_a

Where P = Pressure exerted on the work piece surface during the media flow

N_a = dynamic number of abrasive particles participating in the material removal process

Sliding length for one asperity L = 2a

Number of abrasive particle in full volume of media (N), [8] = $\frac{V_a}{V_o}$

Where V_a = Volume of abrasive grain in full volume of media

V_o = Volume of one abrasive particle

= $\frac{4}{3}.\pi.(\frac{d}{2})^3$

$V_a = \frac{C.V_m.d_m}{d_a}$

Where C = Concentration of abrasive in the media

d_m = density of media

V_m = Volume of media

d_a = Density of abrasive particle

Total number of abrasive particle in contact with the work surface (N_a), [8] = $\frac{4.V_a(D.d-d^2)}{V_o.D^2}$

Fraction of abrasive particle participating in Material removal = $\frac{N_a}{N}$

5. Synthesis of Carbon Nanotubes

This section deals with the growth mechanism and production of carbon nanotubes using chemical vapor deposition (CVD) technique. In CVD technique, the carbon source in the gas phase and an energy source as heated coil for transferring the energy to the gaseous carbon molecule were used. The CVD technique has an advantage of low power consumption, better purity, low range of temperature. The CVD method can be used for producing both single wall and multi wall CNT depending upon the range of temperature. A carbon nano tube structure consists a long and thin cylinder of carbon atoms. The carbon nano tubes are two types; single wall and multi-wall. The difference between the single wall and multi-wall is due to atomic layer of CNT. CNT single walled is formed by a single atomic layer whereas in multi-walled CNT, the wall is formed with more than one atomic layer. Single wall has better distributed carbon atoms structure while multi-wall has defects in its. The tube diameter of single wall ranges between 0.5-4.8 nm. Multi-wall nanotube has its inner diameter in the range of 1.4-14 nm and outer diameter in the range of 2.4-48 nm.

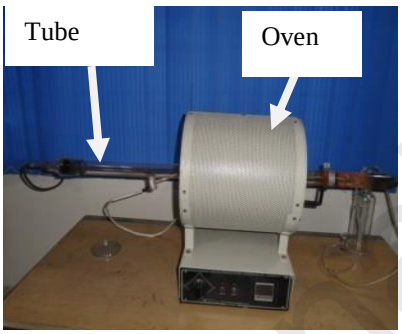
To synthesis CNT, in the CVD method, a hydrocarbon vapor is passed through the tubular reactor which consist a catalyst material having temperature in the range of 600-1200 ^oC to decompose the hydrocarbon. CNTs grow on the catalyst in the reactor on cooling it at the room temperature. For synthesizing CNTs some metal particles are required which can decompose hydrocarbon at a lower temperature. Normally used metals are Fe, Co, Ni and so on [38]. The reason for using such metals is high solubility of carbon in these metals at a very high temperature and carbon diffuse at a faster rate in these metals. Catalyst acts in a different manner in different materials. In CVD method normally substrates used are silicon, silicon carbide, alumina etc. For better growth of CNT,

Journal Pre-proofs

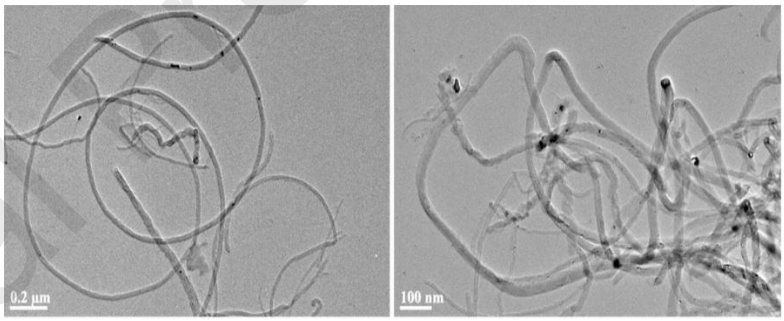
there should be proper interaction between the catalyst and the substrate. The interaction between the catalyst and the substrate will form a chemical bonding.

The main apparatus used to make CNT, is as shown in Fig. 5(a). CNT particles were synthesized by the chemical vapour deposition (CVD) method. Major parts of CNT equipments were oven, tube, beaker, cylinders of argon and hydrogen for shielding. A quartz tube of the dimension as inner diameter of 3.5 cm and length of 60 cm was used to generate CNT vapour. The machine had two zones preheating zone and heating zone. The temperature of the preheating zone varied from 100-150 ^{o}C and its length was 10 cm. The heating zone was 15 cm long and its temperature ranges from 700-825 ^{o}C. 50 ml solution of ferrocene / toluene with the ratio of ferrocene as 0.02 gram in 1 ml of toluene was sonicated and then it was dropped into the preheating zone for evaporation. Argon gas was supplied to carry out the vapours of ferrocene/ toluene to the reaction zone. As the evaporation process completed, then system was allowed to cool. CNT particles were scratched off from the inner wall of the quartz tube [38].

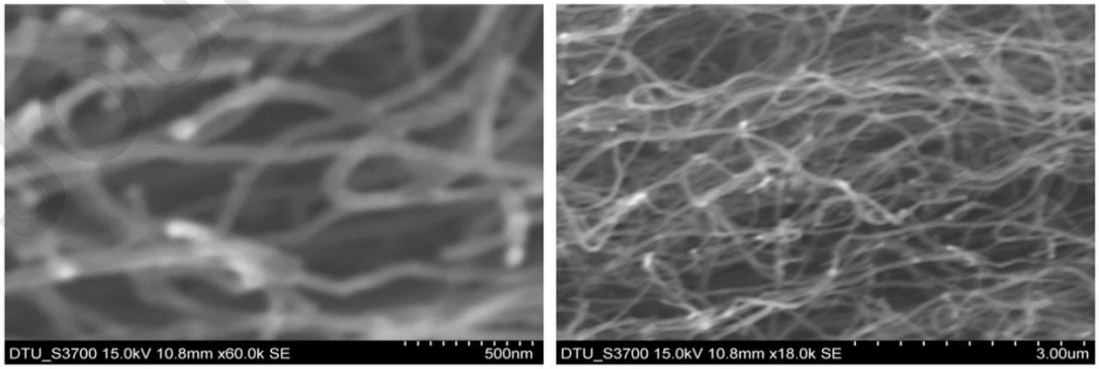
To overcome the major limitation of AFM that is low material removal, iron based carbon nano tubes particles were used along with aluminium oxide (AI_2O_3), because CNT particles are harder than the aluminium oxide. The TEM and SEM images shown in Fig. 5(b) and 5(c) describe iron filled CNT at various locations. The image also shows that the CNT growth has lesser number of carbon impurities which improves its strength and the sharp edges of CNT will not deteriorate easily during the abrasion process.



5(a)



5(b)



5(c)

Figure 5 (a) The apparatus used for the production of CNT particles (b) TEM image of CNT particles (c) SEM image of CNT particles

6. Results and discussion

The final analysis of modelling was performed in Fluent® software to identify the critical conditions of the selected parameters (Pressure, Velocity of media and temperature distribution on work piece surface). The effects were taken on the plane (0,0,0) & (1,0,0) in the geometry. Fig. 6 shows the media velocity along Z axis, where the media was entering and continuously decreasing as it was reported similarly in the literature [39]. As shown in Fig. 7, the velocity of the media flow was maximum at the centre after striking on the work piece surface [40]. When abrasive particles in the media impacts over the work surface, it increases the velocity of the media due to gain of energy after impact [40]. The figure also shows that the media velocity decreases near the wall [41]. This is due to continuously energy loss caused by the frictional forces developed between the fluid particles as already documented [39]. The result also showed that as the media started flowing, continuously from one media cylinder to another (Fig. 8), there would be increase in the media temperature, which was reported previously [6]. The influence of media temperature on the work efficiency is very critical. Figure 8 also shows hot spots on the surface. These hot spots on the surface are observed due to media temperature rise. The rise in the media temperature occurs due to frictional resistance offered by the media particle and outer boundary. Media temperature increases continuously with the increase in number of cycles [42]. This may be the reason of increasing hot spots over the surface. It increases the media temperature that will reduce the viscosity of the media which corresponds to less abrasive holding capacity. The wear rate of the particles increases initially but more increment in the temperature of media results in reduction of the wear rate. Fig. 9 shows the variation of extrusion pressure from inlet to outlet of media fixture chamber in abrasive flow machining setup. The total pressure in the flow is the combination of static and dynamic pressure. Static pressure is due to the random movement of molecules in the fluid and dynamic pressure is due to the motion of the fluid. In AFM process media flows between the media cylinders due to the pressure difference created between both the hydraulic cylinders. Figure 9 shows that as the media flow start from the higher pressurized region to the lower pressurized region, pressure reaches its maximum value at the centre, where minimum standoff distance occurs between the flexible tool and work piece surface [40]. After increasing pressure again it starts decreasing gradually till its outlet. The reason of the pressure reduction is the friction losses in media flow.

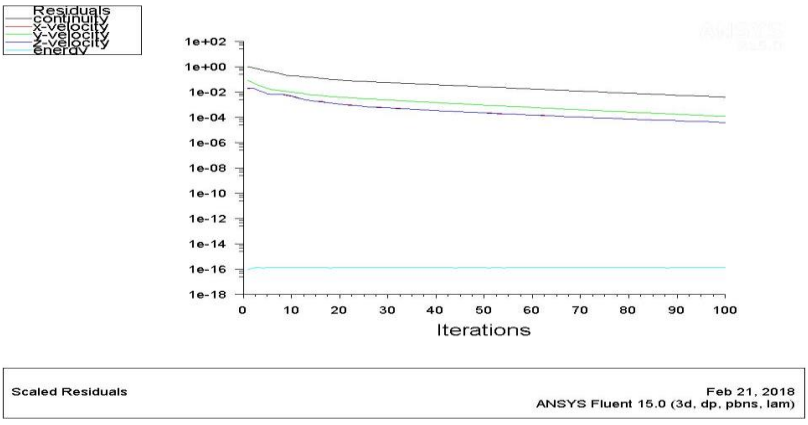


Fig. 6 Effect on the velocity and energy of the media during flow

Journal Pre-proofs

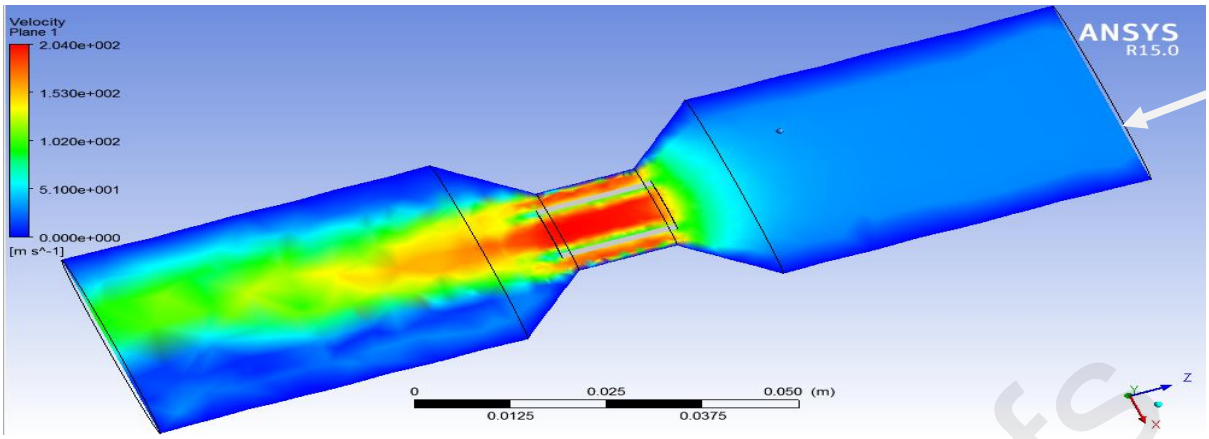


Fig.7 Effect on velocity of the media during the flow on the plane (0,0,0) & (1,0,0)

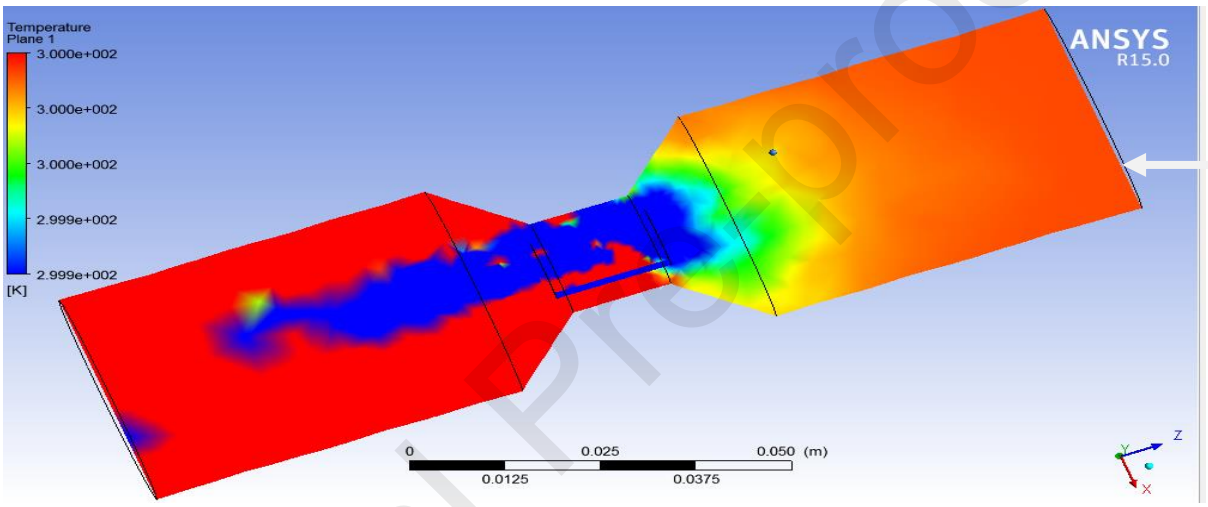


Fig.8 Effect on temperature of media at various points during flow on the plane (0,0,0) & (1,0,0)

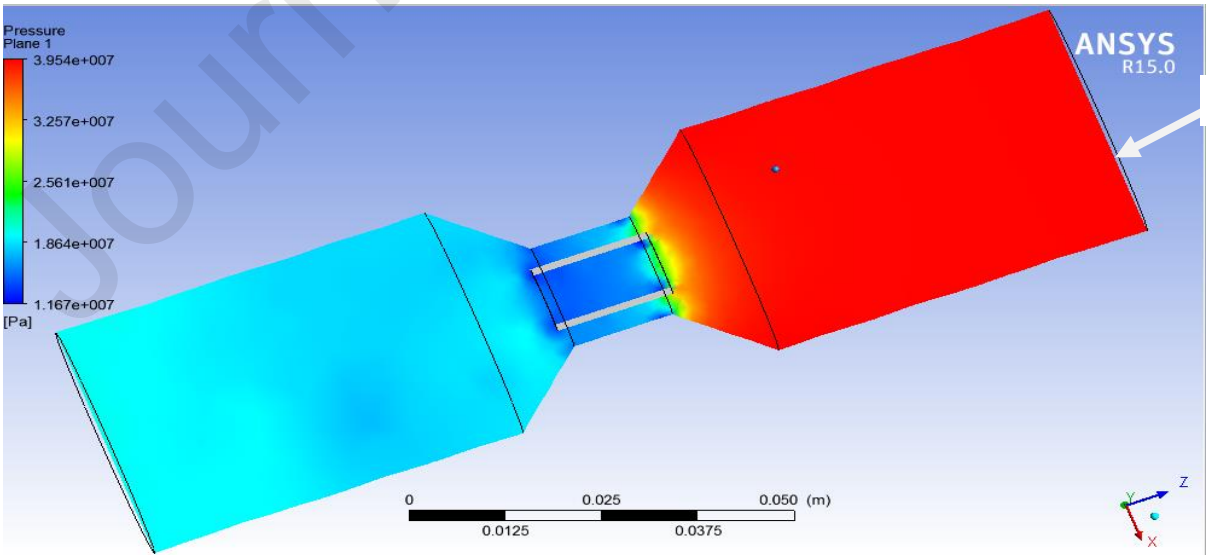


Fig.9 Effect on pressure of the media in flow path on the plane (0,0,0) & (1,0,0)

Journal Pre-proofs

6.1 Material removal calculation using CFD model

Hardness of Aluminum work piece = 1049 MPa

Pressure exerted by the abrasive particle over the surface = 2.997 X 10^7 Pa

a = 16.5 micron

C= 0.6, d_m = 1219 kg/m^3, V_m = 290 X 10-6 m^3,

d_a = 1340 kg/m^3

On putting the values,

V_a = 159.37 X 10^{-6} m^3,

V_o = 0.188 X 10^{-13} m^3,

Number of abrasive particle in full volume of media = N = $\frac{V_a}{V_o}$ = 847.71 X 10^7,

N_a = 1.392 X 10^8,

K = $\frac{N_a}{N} = 0.0164$,

W = 3.565X10^6 N,

Total wear debris = 1.84 mg

6.2 Experimental observations

The experiments were performed as per L_9 OA table. For every trial conditions three work pieces were used for finishing at random run order according to Table 2. The response parameters were MR and % improved R_a.

Weight of the work piece was obtained by precision weighing machine (SHIMADZU AUX220) having 0.1 mg resolution. Surface roughness was taken at five different surfaces of the work piece and average of these values was taken for the initial and final value of surface roughness.

% improvement in R_a

= (Initial Roughness- Final Roughness)/(Initial Roughness) * 100

Surface roughness was measured by Taylor Hobson precision machine having resolution of 0.01 micrometer. The % improved R_a value for each trial condition is shown in Table 2. Higher the better quality characteristic was taken for the responses and for calculating the signal to noise ratio (Table 2) as given in equation 1 [43, 44].

$$(S/N)_{HB} = -10 \log[\frac{1}{n} \sum_{j=1}^{R} \frac{1}{Y_j^2}] \quad (1)$$

It can be seen from the Fig.10 that up to level 2 of extrusion pressure, material removal increases and after level 2 the material removal decreases. It is due to fact that on increasing the pressure cutting force increases and due to large cutting force, abrasives cut the peaks of the surface; therefore material removal is more [42]. However, when the pressure increases further, abrasives in the media cannot roll properly and so media may be drifted as a block with less number of cutting edges with contact of the work surface that slows down the rate of material removal.

From Fig.11, it is found that on increasing amount of CNT particles, material removal also increased. Using conventional abrasives like alumina, silicon carbide has major limitation that during the finishing operation its sharp edges go blunt easily and breaks into small pieces because of low strength and brittle in nature, which reduces the material removal. But CNT particles have a greater strength in comparison to the other conventional abrasives and maintain its sharp edges for a long time. It also does not break easily during the impact on the work surface. This greater strength makes it perfect to use as an abrasive and will reduce the energy requirement during the finishing process.

From Fig. 12, it was found that as the number of cycle were increased initially, the MR value decreased but after sometime as the number of

Journal Pre-proofs

cycle further increased from 6 cycle to 9 cycle, material removal also increased. It may be due to, up to some number of cycles media is very viscous and leads to number of active abrasive particles decrease. As soon as number of cycle's increases, more abrasive particles can contact the work piece due to some decrement in the viscosity and can provide more indentation.

Sankar et al. [45] reported that as the no. of cycle increased, the total no. of indentations by abrasive also increased, so material removal increased. But when the number of cycles was further increased material removal would decrease. Liang et al. [46] investigated that as the no. of cycle increased the temperature continuously increased that caused less material removal. It was reported in the literature, number of cycles was the most significant parameter among various parameters towards the responses such as material removal and improvement in surface finish [47].

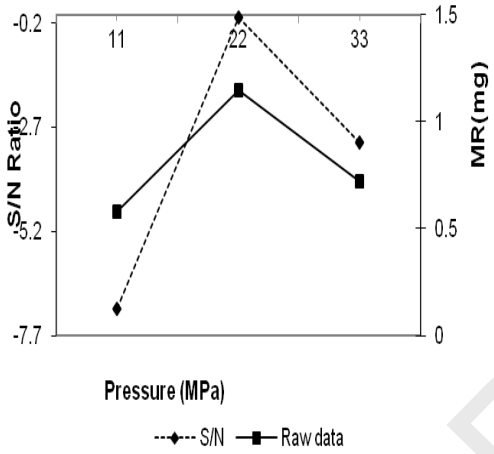


Fig. 10 Effect of extrusion pressure (MPa) on the S/N Ratio and MR

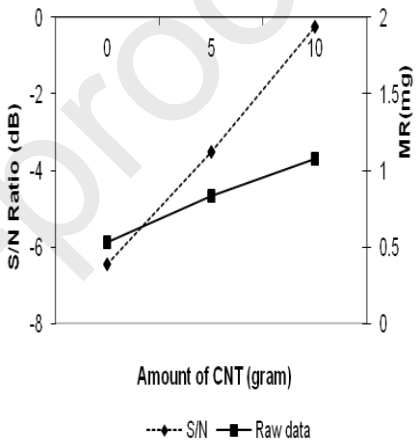


Fig.11 Effect of Amount of CNT (gram) on the S/N Ratio and MR

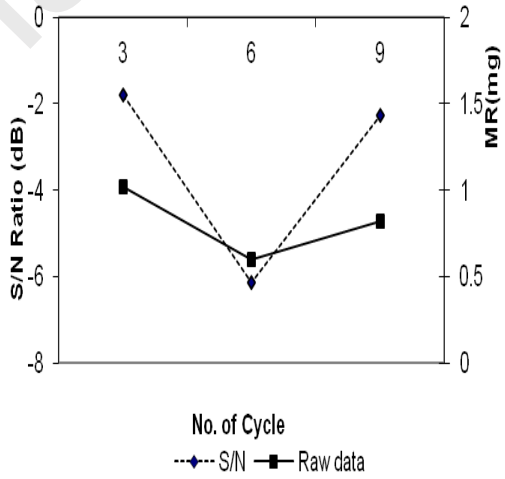


Fig.12 Effect of No. of Cycle on the S/N Ratio and MR

Journal Pre-proofs

Table 4 Average values and Main effects: Material Removal, MR (in mg)

Process Parameter	Level	Pressure(P)		Amount of CNT(C)		No. of Cycle (N)	
Type of Data		Raw Data	S/N Ratio	Raw Data	S/N Ratio	Raw Data	S/N Ratio
Average Values		L1	0.58	-7.06	0.53	-6.44	1.02
		L2	1.14	-0.07	0.83	-3.50	0.6
		L3	0.72	-3.06	1.08	-0.24	0.82
Main Effects		L2-L1	0.56	6.99	0.3	2.94	-0.42
		L3-L2	-0.42	-2.99	0.25	3.26	0.22
Difference (L3-L2)–(L2-L1)		-0.98	-9.98	-0.05	-0.32	0.64	8.23

L1, L2, L3 represent levels 1, 2 and 3 respectively of parameters. (L2-L1) is the average main effect when the corresponding parameter changes from Level 1 to Level 2. (L3-L2) is the main effect when the corresponding parameter changes from Level 2 to Level 3

To analyze the effects of parameters on the response character, which is material removal in this study, analysis of variance was used. Table 5 and Table 6 respectively represent the ANOVA of raw and S/N data. It is denoted from the tables that the parameters extrusion pressure, amount of CNT and No. of Cycle [35] affects the mean and the variability in the material removal values.

Table 5 Shown the ANOVA for Raw Data (MR)

Source	SS	DOF	V	F-Ratio	SS’	P%
Pressure	1.56	2	0.78	17.90*	1.47	32.22
Amount of CNT	1.34	2	0.67	15.35*	1.25	27.36
No. of Cycle	0.80	2	0.40	9.21*	0.72	15.65
e	0.87	20	0.04	-----	1.13	24.78
Total (T)	4.57	26	*	-----	4.57	100

*Significant at 95% confidence level, F critical =3.4928(Tabular Value)

SS-Sum of Squares, DOF-Degree of Freedom, V-Variance, SS’-Pure sum of Squares

Journal Pre-proofs

Table 6 Shown the ANOVA for S/N ratio (MR)

Source	SS	DOF	V	F-Ratio	SS'	P%
Pressure	73.74	2	36.87	41.40*	71.96	42.93
Amount of CNT	57.85	2	28.92	32.48*	56.07	33.44
No. of Cycle	34.27	2	17.13	19.24*	32.48	19.38
e	1.78	2	0.89	----	7.12	4.25
Total (T)	167.64	8	----	----	167.64	100
*Significant at 95% confidence level, F critical =19(Tabular Value)						

The optimal material removal was determined and related confidence intervals were predicted. The mean at the optimal is estimated as

$$\mu = \bar{P}_2 + \bar{C}_3 + \bar{N}_1 - 2\bar{T} \quad (2)$$

$\overline{T}$= total mean of the response = 0.81 mg [Table 2]

$\bar{P}_2$ = Average material removal at the 2nd level of Pressure = 1.14 mg [Table 4]

$\bar{C}_3$ = Average material removal at the 3rd level of CNT= 1.08 mg [Table 4]

$\bar{N}_1$ = Average material removal at the 1st level of no. of cycle = 1.02 mg [Table 4]

On putting the values, Mean MR = 1.62 mg

$$CI_{COE}\ [48,50] = \sqrt{F_a(1, f_e) V_e \left[\frac{1}{n_{eff}} + \frac{1}{R}\right]} \quad (3)$$

$$CI_{POPU} = \sqrt{\frac{[F_a(1, f_e)\, V_e]}{n_{eff}}} \quad (4)$$

Where CI_{COE} = Confirmation experiments confidence interval

CI_{POPU} = confidence interval of population

f_e = Error degree of freedom = 20

N = Entire number of experiments

=27 (experiment =9, Iteration =3)

R = Confirmation experiments size = 3

V_e = 0.04 (Table 5)

n_{eff} [44, 49]=

$$\frac{N}{1+[Degree\ of\ freedom\ for\ the\ estimated\ mean\ response]} \quad (5)$$

= 3.86

So, CI_{COE} = ± 0.29

and CI_{POPU} = ± 0.19

The optimal range for the confirmation runs of the repetition at 95% confirmation interval:

Average MR – CI_{COE}< Material Removal < Material Removal + CI_{COE}

1.33< Material Removal <1.91

Predicted mean for 95% confirmation interval is:

Average Material Removal – CI_{POPU}< Material Removal < Material Removal + CI_{POPU}

1.43< Material Removal <1.81

Journal Pre-proofs

Journal Pre-proofs

The mean values of % improved R_a and S/N ratio for every input variable at all the levels were calculated and given in Table 7.

Table 7 Average values and Main Effects: Percentage improvement in surface roughness

Process Parameter	Level	Pressure(P)		Amount of CNT(C)		No. of Cycle (N)	
Type of Data		Raw Data	S/N Ratio	Raw Data	S/N Ratio	Raw Data	S/N Ratio
Average Values(%R_a)	L1	13.70	22.24	11.86	20.99	15.48	22.84
	L2	24.14	27.33	20.67	25.77	16.34	24.23
	L3	13.04	22.17	18.36	24.98	19.07	24.68
Main Effects(%R_a)	L2-L1	10.44	5.09	8.81	4.78	0.86	1.39
	L3-L2	-11.10	-5.16	-2.31	-0.79	2.73	0.45
Difference (L3-L2)–(L2-L1)		-21.54	-10.25	-11.12	-5.57	1.87	-0.94

It can be clearly noticed from the Fig.13 that as the pressure increased, the surface finishing of the work piece increased initially but after level 2 (i.e 22 MPa) of pressure, the value of improvement in surface finishing decreased. Because of more material removal, initially greater surface finishing was achieved but after level 2, due to higher pressure the abrasive in the media could be drifted and less number of abrasive could contact to the work surface that reduces the material removal. Therefore it can lower improvement in surface finishing. It can also be seen from the Fig. 14 that as the amount of CNT increased, percentage increase in R_a also increased but after level 2, percentage improvement in R_a started decreasing at a slow rate. Carbon nano tubes are strong and elastic in nature. When the pressure is being applied at the tip of the nano tube it starts bending without being damaged. Again when the pressure is relieved, it comes to the original position if the pressure applied is within the limit. That is the reason it may remove more material from the surface and provides better surface quality initially.

Fig. 14 shows that on increasing the amount of CNT, more material removal caused better surface finish. But after 5 gram of CNT less surface finishing was obtained. The reason may be that on further increasing CNT amount, the material removal would deteriorate the level of finishing and the rate of % increment in R_a would slow down. From Fig. 15, it is defined whenever the number of cycle is increased; the percentage increase in surface finish is increased. It may be that on increasing the number of cycle, the contact area of the abrasive and work surface can be more which corresponds to the better surface finishing.

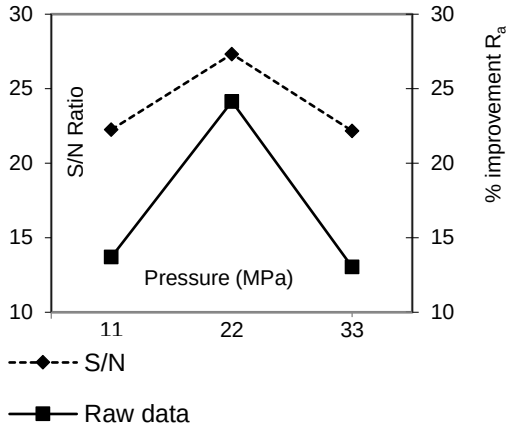


Fig 13 Shown effect of pressure on S/N ratio and % improvement in R_a

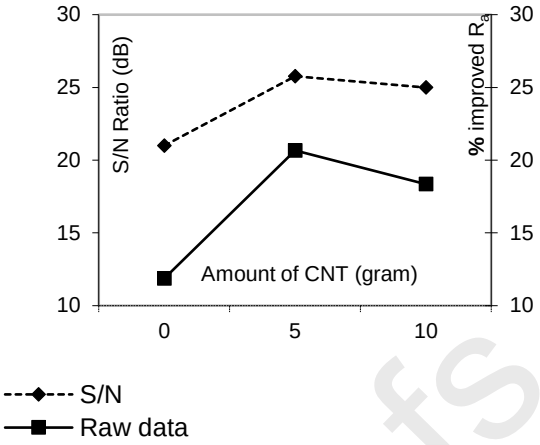


Fig 14 Shown effect of Amount of CNT on S/N ratio and % improvement in R_a

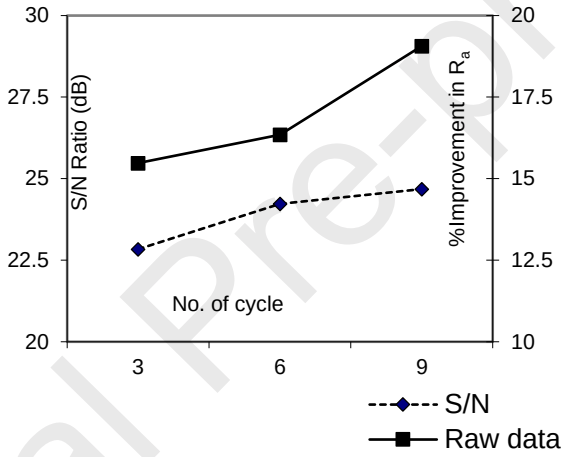


Fig 15 Shown effect of No. of Cycle on S/N Ratio and % improvement in R_a

To analyze the significance of the parameters against the % improved R_a, (ANOVA) was used. The pooled versions of Analysis of variance for both raw and S/N data against % improved R_a are provided in Table 8 and 9.

Table 8 Pooled ANOVA (Raw Data)(ΔR_a)

Source	SS	DOF	V	F-Ratio	SS'	P%
Pressure	697.78	2	348.89	107.95*	691.32	57.58
Amount of CNT	375.21	2	187.60	58.047*	368.75	30.71
No. of Cycle	63.08	2	31.54	9.76*	56.61	4.72
e	64.64	20	3.23	------	84.03	6.99
Total (T)	1200.70	26	*	-------	1200.70	100
*Significant at 95% confidence level, F critical =3.4928(Tabulated Value)						

Journal Pre-proofs

Table 9 Pooled ANOVA (S/N Data)(ΔR_a)

Source	SS	DOF	V	F-Ratio	SS'	P%
Pressure	52.43	2	26.22	279.77*	52.24	53.61
Amount of CNT	39.30	2	19.65	209.70*	39.11	40.14
No. of Cycle	5.53	2	2.76	29.50*	5.34	5.48
e	0.19	2	0.09	------	0.75	0.77
Total (T)	97.45	8	--------	-------	97.45	100
*Significant at 95% confidence level, F critical =19(Tabulated Value)						

The optimum values of % improved R_a is estimated as:

$\mu = \bar{P}_2 + \bar{C}_2 + \bar{N}_3 - 2\bar{T}$ (6)

$\bar{T}$ = Total mean = 16.96 %

$\bar{P}_2$ =Average % age improvement in R_a at the 2nd level of extrusion Pressure = 24.14%

$\bar{C}_2$ =Average % age improvement in R_a at the 2nd level of CNT = 20.67 %

$\bar{N}_3$ = Average % age improvement in R_a at the 3rd level of No. of Cycle =19.07%

Putting the given values, % improvement in R_a = 29.96 %.

So, $CI_{COE} = \pm 2.58$

and $CI_{POPU} = \pm 1.70$

f_e =3.49 (Tabulated Value)

f_e =error Degree of freedom = 20

N= Entire number of experiments =27 (experiment =9, Iteration =3)

R = Confirmation experiments size = 3

V_e = Error variance =3.23

By using equation no. (5), one may achieve

n_{eff} = 3.86

The optimal range for the confirmation runs of the iteration at 95% confirmation interval:

27.38< % improved ΔR_a<32.54

Predicted mean for 95% confirmation interval is:

28.26<% improved ΔR_a <31.66

From the confirmation experiments, the optimal levels of the process parameter for material removal are $P_2C_3N_1$.

Whereas for the maximum % improved R_a, the optimal parameters settings are $P_2C_2N_3$.

P_2 = Extrusion Pressure at second level =22 MPa

C_2 = Amount of CNT at second level =5 gram

C_3 = Amount of CNT at Third level =10 gram

N_1 =No. of Cycle at first level =3

N_3 = No. of Cycle at third level =9

The results are presented in Table 10 for material removal and % improved R_a both.

Journal Pre-proofs

The values of material removal and % improved R_a obtained through the confirmation experiments are inside the specified range.

Table 10 Predicted Optimal Values, Confidence Intervals and Results of Confirmation Experiments

Response Characteristic	Optimal Process Parameters	Predicted Optimal Value	Confidence Intervals 95%	Actual Value
MR	$P_2C_3N_1$	1.62 mg	CI_{CE} 1.33<MR<1.91 CI_{POP}:1.43<MR<1.81	1.73 mg
%Improvement ΔR_a	$P_2C_2N_3$	29.96 %	CI_{CE}:27.38<%ΔR_a<32.54 CI_{POP}:28.26<%ΔR_a<31.66	30.93%

7. Conclusions

Using carbon nano tubes along with alumina as abrasives in AFM for which the results are reported in this study can be one way to get highly precision components. Following conclusions can be made:

1. The simulation and experimental results were quite similar for the material removal; minor discrepancies might be minimized through the adjustments in the boundary conditions.
2. On comparing the material removal results obtained from experimental and modeling, it can be observed that material removal for the model is slightly higher (6.35 %) in comparison to experimental results. This reduction in material removal for experimental results may be due to losses occurring during the media flow.
3. The ANSYS® modelling shows that the velocity of the media is maximum at the centre when the abrasive particle is impacting over the surface and is least at the outlet point. It can also be observed that velocity of media continuously decreases after the impact on work surface. From modelling it can also be concluded that temperature of media starts increasing with the flow and is maximum at the outlet, while pressure continuously decreases between the initial and final point.
4. The addition of carbon nano tubes as abrasive along with the alumina enhances the Material Removal and R_a of the work piece.
5. The experimental results show that Pressure is contributing a maximum as 32.22%, amount of CNT is contributing 27.36% and number of cycle is contributing 15.65% in the material removal.
6. The experimental result shows that on using CNT as abrasive particle, 0.6 mg more material removal is obtained at 10 gram of CNT and also gives 10 % more percentage improvement in surface roughness at 5 gram amount of CNT compared to the conventional abrasives.
7. In terms of percentage improvement of surface roughness during finishing process pressure is contributing 57.78%, amount of CNT contributes 30.71% and number of cycle contributes 4.72 % in surface finishing.
8. The predicted optimal range for material removal is 1.33 < Material Removal (mg) < 1.91 and that for % increase in ΔR_a is 27.38<%ΔR_a< 32.54.
9. The 95% confidence interval of the material removal is 1.43 < Material Removal (mg) < 1.81

Journal Pre-proofs

and that for % increase in ΔR_a is 28.26< $\%\Delta R_a$< 31.66.

10. The result shows that the optimum process parameter for the material removal is second level of pressure i.e 22 MPa, third level of amount of CNT i.e 10 gram and first level of number of cycle i.e 3, which gives the optimum value 1.73 mg for the material removal.

11. The result shows that the optimum process parameter for the % improvement in surface roughness is second level of pressure i.e 22 MPa, second level of amount of CNT i.e 5 gram and third level of number of cycle i.e 9 and gives the optimum value 30.93% for the % increment in surface roughness.

Acknowledgement

The authors acknowledge Precision Manufacturing Lab, Delhi Technological University for necessary supports for this work.

8. References

[1] M.E. Merchant, Newer methods for the precision working of metals research and present status, Proc. Int. Con. Production Res. ASME, (1962) 93-107.

[2] S. Jaikishan, M. H. Singh, Recent developments in abrasive flow finishing process: A review of current research and future prospects, Proc IMechE Part B: J Engineering Manufacture, (2017) 1–12.

[3] V.K. Jain, C. Ranganatha, K. Muralidhar, Evaluation of rheological properties of medium for AFM process, Machining science and Technology 5(2) (2001) 151–170.

[4] M. Ravi Sankar, J. Ramkumar; V.K. Jain, Abrasive flow machining (AFM): An Overview. INDO-US WORKSHOP on Smart Machine Tools, Intelligent Machining Systems and Multi-scale Manufacturing (2008).

[5] Jose. Cherian, Dr. Jeoju, M. Issac, Effect of process variables in Abrasive Flow Machining, International Journal of Emerging Technology and advanced engineering 3 (2) (2013) 554-557.

[6] R. S. Walia, Development and Investigations in centrifugal force assisted Abrasive Flow machining Process, IIT Roorkee, Phd. Thesis (2006).

[7] R. S. Walia, H.S.Shan, P.K. Kumar, Parametric optimization of centrifugal force assisted abrasive flow machining process by the Taguchi method, Materials and Manufacturing Processes 21(4) (2006) 375- 382.

[8] R.S. Walia, H.S Shan, P.K. Kumar, Determining dynamically active abrasive particles in the media used in centrifugal force assisted abrasive flow machining process, International Journal of Advanced Manufacturing Technology 38 (2008) 1157-1164.

[9] R.S. Walia, H.S. Shan, P.K. Kumar, Morphology and integrity of surface finished by centrifugal force assisted abrasive flow machining, International Journal of Advanced Manufacturing Technology 39(2) (2008) 1171-1179.

[10] R.S.Walia, H.S.Shan, P.K. Kumar, Enhancing AFM process productivity through improved fixturing, International Journal of Advanced Manufacturing Technology 44 (2009) 700-709.

[11] S. Mittal, V. Kumar, H. Kumar, Experimental investigation and optimization of process parameters of Al/SiC MMCs finished by abrasive flow machining, Materials and Manufacturing Processes 30 (2015) 902-911.

[12] A. C. Wang, K. C.Cheng.; K. Y. Chen, Y. C. Lin, Enhancing the surface precision for the helical passageways in Abrasive flow machining, Materials and Manufacturing Processes 29 (2014) 153-159.

[13] S. Jai kishan, H.S Mali, V. Garg, Experimental investigation on unidirectional Abrasive Flow Machining of trim die work piece, Materials and Manufacturing Processes (2017) doi. 10.1080/10426914.2017.1364847.

Journal Pre-proofs

[14] G. Venkatesh, A.K. Sharma, P. Kumar, On ultrasonic assisted abrasive flow finishing of bevel gears, International Journal of Machine Tools & Manufacture 89 (2015) 29–38.

[15] B. S. Brar, R. S. Walia, V. P. Singh, Electrochemical-aided abrasive flow machining (ECA2FM) process: a hybrid machining process, International Journal of Advanced Manufacturing Technology 79 (2015) 329–342.

[16] M.A. Marzban, S.J. Hemmati, Modeling of abrasive flow rotary machining process by artificial neural network, International Journal of Advanced Manufacturing Technology 89 (2017) 125–132.

[17] S. Wan, Y. J. Ang, T. Sato; G. C. Lim, Process modeling and CFD simulation of two-way abrasive flow machining, International Journal of Advanced Manufacturing Technology 71 (2014) 1077–1086.

[18] J. Kim, K. Kim, Deburring of burrs in spring collets by Abrasive Flow Machining, International Journal of Advanced Manufacturing Technology 24 (2004) 469–473.

[19] H. Tzeng, B. Yan.; R. Hsu, H. Chow, Finishing effect of abrasive flow machining on micro slit fabricated by wire-EDM, International Journal of Advanced Manufacturing Technology 34 (2007) 649–656.

[20] H.S. Mali, A. Manna, Optimum selection of abrasive flow machining conditions during fine finishing of Al/15 wt% SiC-MMC using Taguchi method, International Journal of Advanced Manufacturing Technology 50 (2010) 1013–1024.

[21] A.A. Mollah, D. K. Pratihar, Modeling of TIG welding and abrasive flow machining processes using radial basis function networks, International Journal of Advanced Manufacturing Technology 37 (2008) 937–952.

[22] Y. Seifu, S. Kumar S., S. S. Hiremath, Modeling and Simulation: Machining of Mild Steel using Indigenously Developed Abrasive Flow Machine, Procedia Technology 25 (2016)1312 – 1319.

[23] G. Venkatesh, A. K. Sharma, N. Singh, Simulation of media behaviour in vibration assisted abrasive flow machining, Simulation Modelling Practice and Theory 51 (2015) 1–13.

[24] K. Chen, K. Cheng, A study of helical passageways applied to polygon holes in abrasive flow machining, International Journal of Advanced Manufacturing Technology 74 (2014) 781–790.

[25] Y. Fu, X. Wang, H. Gao, H. Wei, S. Li. Blade surface uniformity of blisk finished by abrasive flow machining, International Journal of Advanced Manufacturing Technology 84 (2016) 1725–1735.

[26] H. Wei, H. Gao, X. Wang, Development of novel Guar gum hydrogel based media for Abrasive Flow Machining: Shear thickening behavior and finishing performance, International Journal of Mechanical Sciences (157-158) (2019) 758-772.

[27] C. Bouland, V. Urlea, K. Beaubier, M. Samoilenko, V. Brailovski, Abrasive Flow Machining of laser powder bed fused parts: Numerical modeling and experimental validation, Journal of Materials Processing technology 273 (2019) 116262.

[28] S. Han, F. Salvatore, J. Rech, Residual stress profiles induced by abrasive flow machining in 15-5PH stainless steel internal channel surfaces, Journal of Materials Processing technology, DOI: https://doi.org/10.1016/j.jmatprotec.2018.12.024.

[29] Z. Lv, R. Hou, C. Huang, H. Zhu, H. Qi, Meshfree analysis on dynamic behavior of hard brittle material in Abrasive Flow Machining, International Journal of Advanced Manufacturing Technology 100 (2019) 2021–2030.

[30] X.C Wang, C.C Wang, C.Y Wang, F.H Sun, Approach for polishing diamond coated

Journal Pre-proofs

complicated cutting tool: Abrasive Flow Machining (AFM), Chinese Journal of Mechanical Engineering 31(97) (2018), DOI: https://doi.org/10.1186/s10033-018-0296-4.

[31] Y. Shao, K. Cheng, Integrated modelling and analysis of micro cutting mechanics with the precision surface generation in Abrasive Flow Machining, International Journal of Advanced Manufacturing Technology, DOI: https://doi.org/10.1007/s00170-019-03595-4.

[32] L.J. Rhoades, T.A. Kohut, Reversible Unidirectional AFM. US patent number 5,070,652, Dec 10th (1991).

[33] R.S. Walia, H.S. Shan, P.K. Kumar, Finite element analysis of media used in centrifugal force assisted Abrasive Flow Machining Process, Proceeding the institute of mechanical engineering, Part B, Journal of engineering manufacture 222 (2006) 1775-1785.

[34] R.S. Walia, H.S. Shan, P.K. Kumar, Abrasive flow machining with additionally centrifugal force applied to media, Machining Science Technology 10 (2006) 341-354.

[35] T. Kohut, Surface finishing with Abrasive flow machining, SME Technical Paper Proc. 4th International Aluminium extraction Technology seminar ,Washington D.C, (1988) 35-42.

[36] https://en.wikipedia.org/wiki/Archard_equation.

[37] S. Singh, M. Ravi Sankar, V.K. Jain, Simulation and experimental investigations into abrasive flow nanofinishing of surgical stainless steel tubes, Machining science and Technology (2017) doi: 10.1080/10910344.2017.1365897.

[38] R. Kumari, A. Singh, R. Kumar, L. Krishnia, V. Kumar, N.K Puri, P.K Tyagi, Synthesis of Ni filled multi walled carbon nanotubes and study of magnetic behaviour, Advanced Materials Letters 7(3) (2016) 197-200.

[39] H. Tzeng, B. Yan, R. Hsu, Y. Lin, Self-modulating abrasive medium and its application to abrasive flow machining for finishing micro channel surfaces, International Journal of Advanced Manufacturing Technology 32 (2007) 1163–1169.

[40] T. Wang, D. Chen, W. Zhang, L. An, Study on key parameters of a new Abrasive Flow Machining process for surface finishing, International Journal of Advanced Manufacturing Technology 101(1-4) (2019) 39-54.

[41] M.A Manjunath, A. Kumar, V. Prakash, N. Balashnmugam, Simulation of force generated and material removal in Abrasive flow finishing for Aluminium material, Proceedings of 10th International conference on Precision, Meso, Micro and nano engineering, IIT Madras (2017).

[42] S. Singh, M. Ravi Sankar, V.K Jain, J. Ramkumar, Modelling of finishing forces and surface roughness in Abrasive flow finishing process using rheological properties, 5th International & 26th All India Manufacturing Technology Design and Research Conference (AIMTDR) (2014) December 12th–14th, IIT Guwahati, Assam, India.

[43] T. Goyal, R.S. Walia, T. S. Sidhu, Study of Coating Thickness of Cold Spray Process Using Taguchi Method, Materials and Manufacturing Processes 27 (2012) 185–192.

[44] T. Goyal, R. S. Walia, T. S. Sidhu, Effect of Parameters on Coating Density for Cold Spray Process, Materials and Manufacturing Processes 27 (2012) 193–200.

[45] M. Ravi Sankar, J. Ramkumar, V.K. Jain, Experimental Investigation and mechanism of material removal in nano finishing of MMCs using Abrasive Flow Finishing Process, Wear 266 (2009) 688-698.

[46] L. Fan, J. Zhao, K. Sun, D. Zheng, D. Ma, Temperature as sensitive monitor for efficiency of

Journal Pre-proofs

work in Abrasive flow Machining, Wear 266 (2009) 678-687.

[47] R.S. Walia, R. Singh, Study the effect of centrifugal force on abrasive flow machining process, International Journal of research in mechanical engineering and technology 2(1) (2012) 34-39.

[48] R. S. Walia, H. S. Shan, P. Kumar, Multi-Response Optimization of CFAAFM Process through Taguchi Method and Utility Concept, Materials and Manufacturing Processes 21(2006) 907–914.

[49] T. Goyal, R. S. Walia, T. S. Sidhu, Multi-response optimization of low-pressure cold-sprayed coatings through Taguchi method and utility concept, International Journal of Advanced Manufacturing Technology 64 (2013) 903–914.

[50] K. Sudhir, S. Rajender, Optimization of process parameters of Metal Inert gas welding with preheating on AISI 1018 mild steel using grey based Taguchi method, Measurement (2019), DOI: https://doi.org/10.1016/j.measurement.2019.106924.

Journal Pre-proofs

Highlights

- Three dimensional modelling of the media flow path to visualize in AFM
- Developed abrasive media with CNT and alumina to enhance material removal in AFM
- AFM process parameters; pressure, CNT amount, number of cycle were analysed using Taguchi L9 OA.
- Reye–Archard–Khrushchov law through computational approach was used to predict the wear

Journal Pre-proofs

Conflict of interest

The authors declare that there is no conflict of interest of whatsoever as far as the present work is concerned for publishing in Brazilian Journal of Microbiology.

R S Walia
(Corresponding author)
waliaravinder@yahoo.com

Journal Pre-proofs

Scholastic Seed Inc.

www.gjeis.com

global Journal of ENTERPRISE INFORMATION SYSTEM

Exploring Start-up Ecosystem and its Structural Impact : A Review

– **Surabhi Kapoor***
Research Scholar, Department of Humanities, Delhi Technological University, Delhi
surabhikapoor1@gmail.com https://orcid.org/0000-0003-3831-709X

– **Seema Singh**
Professor in Economics, Department of Humanities, Delhi Technological University, Delhi
seemasinghdtu@gmail.com

ARTICLE HISTORY

Paper Nomenclature: Review of literature (RoL)

Paper Code: V11N2AJ2019ROL3

Originality Test Ratio: 4%

Submission Online: 6-April-2019

Manuscript Acknowledged: 10-May-2019

Originality Check: 14-May-2019

Peer Reviewers Comment: 24-June-2019

Blind Reviewers Remarks 20-July-2019

Author Revert: 21-Aug-2019

Camera-Ready-Copy: 28-Aug-2019

Editorial Board Citation: 05-Sep-2019

Published Online First: 20-Sep-2019

EDITORIAL BOARD EXCERPT Initially at the Time of Submission (ToS) submitted paper had a 4% plagiarism which is an accepted percentage for publication. The editorial board is of an observation that paper had a successive close watch by the blind reviewers which at a later stages had been rectified and amended by authors (Surabhi & Seema) in various phases as and when required to do so. The comments related to this manuscript are noticeable related to **Start-up Ecosystem and its Structural impact** both subject-wise and research-wise. As we all know Start-ups are effective in promoting innovation and Setting up start-ups is a driving force for the development of worldwide economy. The study has made a deliberate attempt in identify the emerging nature of start-ups and to discover factors helping in the development of start-up ecosystems. By and large all the editorial and reviewer's comments had been incorporated in paper at the end and further the manuscript had been earmarked and decided under "**Review of Literature**" category as the paper critically analyze the research studies describing the stages of start-ups, attributes of start-up ecosystem and their inter-relationship.

ABSTRACT

Purpose: Present research focus has primarily been around the management and development about the startup ecosystem framework but research on relative significance of start-up ecosystem factors are yet to be done. The purpose of the research is to identify the emerging nature of start-ups and to identify factors helping in the development of start up ecosystems

Design/Methodology Approach-Review based approach findings- The paper critically analyze the research studies describing the stages of start-ups, attributes of start-up ecosystem and their inter-relationship

Originality/Value-The study has made a comprehensive literature review of the area to understand the evolution of the field, development of startup through various stages and growth of effective start-up ecosystem. The study summarizes briefly each section, integrate the knowledge and suggest entrepreneurial and academic significance, contribution in the area of the startup ecosystem.

KEYWORDS Start-ups | Start-up Ecosystem | Entrepreneurship | Entrepreneurial Environment

***Corresponding Author** (Kapoor - Et. Al)

https://doi.org/10.18311/gjeis/2019
Volume-11 | Issue-2 | Apr-Jun, 2019 | Online ISSN : 0975-1432 | Print ISSN : 0975-153X
Frequency : Quarterly, Published Since : 2009

©2019 GJEIS Published by Scholastic Seed Inc. and Karam Society, New Delhi, India. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Introduction

Setting up start-ups is a driving force for the development of worldwide economy (Lederman et al., 2013)[1]. Start-ups are effective in promoting innovation. In present Era, countries' potential to transform innovations, to successful ventures and creating high returns with multiple jobs is of high importance for resource utilization and to develop a quality life for its citizens (Europe 2012)[2].

Entrepreneurship is series of action which gives scope for novel products or services to be discovered, analyzed and exploited (Shane and Venkatamaran,2000)[3]. Start-up establishments are formed by a group of two, three people having burning ideas and zeal to form new business and by the help of a small number of employees, pilot runs different possibilities to produce viable product or service and float in the market within few months of idea formation (Blank and Dorf 2012)[4].

There is an escalation in entrepreneurship around the globe as well "outburst start-up ecosystem all over the world" (Start-up Genome, 2012)[5]. Start-ups contribute to the economic growth of a nation through an increase in employment, exports as well as the GDP. Entrepreneurship drives the economy and society of almost all nations (Brock and Evans, 1989[6]; Acs,1992[7]; Carree and Thurik, 2002[8]).

Through time memorial the USA has been the cradle of high-growth start-ups primarily through startup ecosystems. The startup ecosystems power the high-growth start-ups through a combination of multiple factors like support to creative, marketable ideas, market potential, policy support as well as open market environment. Start-ups help in bringing structural change in an economy by initiating knowledge base products, promotes innovation, steers productive growth and employment (OECD, 2013)[9].

While start-ups have the potential to generate jobs and contribute to regional and national economic growth, they have been identified with a high rate of failure, and therefore, have a very limited rate of survival and growth (Robinson, 1998[10] ; Song et al., 2008[11]). More often, they suffer from scarce resources, skeptical customers, inaccessible apt human resources, inadequate or inaccessible or prohibitively high cost, support services (Bala Subrahmanya, 2010)[12]. Start-ups have certain important predictors or influential factors – internal and external (Gilbert, et al., 2006[13]; Folta, et al., 2006[14]). Given the internal set of factors, start-ups are based on local environment for "diverse resources" to a huge extent to maintain their operations (Romanelli and Schoonhoven, 2001)[15].

The present government has taken major steps to refine ease of doing business. In fact, the '*Doing Business Report 2008*' (published by the World Bank-IFC) states that "India can jump 55 places from its current rank of 120th if some of the local best practices are adopted nationwide."

A start-up is a new venture developed around a creative idea for commercial purpose. It is an innovative concept for changing the existing system. India has recently seen an increase in start-ups. For some every new business is a start-ups , while for some the technology driven and innovative businesess are startups.

Startup Definition

Start-ups may belong to any sector, manufacturing or services or both. The start-up may be using technology extensively or trying to promote technologically based products or services. A start-up may explore an inventive business model which may churn existing organizations or markets. The startup entrepreneur would have gained the knowledge about his main product/ service through a high level of self-motivation, creativity, learning capabilities and desire to succeed.

More about start up and start up ecosystem can be learned through literature review in the next section.

Author	Definition
Blank and Dorf 2012 [4]	"A startup is a temporary organization in search of a scalable, repeatable, profitable business model."
Certo, 2003[16]	"A start-up, in general, is defined as a new venture with no previous history of operations. They suffer from the liability of newness since they are unfamiliar and without any precedence."
OECD, 2013[8]	"Start-ups focus on performance, using terms such as high-impact, high growth, and gazelles. The OECD defines high-growth enterprises as those with average annualized growth greater than 20% regarding growth in the number of employees or growth in sales volume over a three-year period, with ten or more employees at the beginning of the observation period. Gazelles are the youngest of these enterprises, typically in existence for less than five years)."
Ministry of Commerce and Industry, (2019)[17]	"An entity shall be considered as a 'startup'- a) Up to ten from the date of its incorporation/registration, b) If its turnover for any of the financial years has not exceeded Rupees 100 crore, and c) If it is working towards innovation, development or improvement of products or processes or services, or if it is a scalable business model with a high potential of employment generation or wealth creation. "
Blank (2010)[18]	"An organization formed to search for a repeatable and scalable business model."
Graham (2012)[19]	"A startup is a company designed to grow fast. The only essential thing is growth. Everything else we associate with start-ups follows from growth."
Stokes (1995)[20]	"A start-up means creating a new business, which stands alone and is not tied to other organizations, except in the normal course of trading. It does not mean that the idea is necessarily new, it is only the vehicle which is set up to exploit it."

Stages of Start-up Development

Understanding of different stages of start-up development is essential, as every stage pose a challenge for the organization survival. Many research works have been done earlier to understand the framework and model of the new firm. Each framework, the model provides insight and to draw a conclusion for it. The stages of development by different authors share several common traits but have a deviation about event, durations, etc. The author shares their distinct perspective to start-ups growth stages.

Greiner's (1972)[21]	Churchill and Lewis (1983)[22]	Aldrich, H.E. (1999)[23]	(Blank and Dorf 2012)[4]	(Marmer et al 2011)[24]
Organization age –Attitude change with age of firm is required	**existence stage-** broadening initial customer, the initial money	**Variation-** reviewed the formulation of various organization structures	**Customer Discovery-** identify consumer need, extent of product or service capability to reach consumer needs form a business model, test the hypothesis through feedback, validate it, and to develop solution fit.	**Discovery- the** minimum viable product is generated. Initial funds are traced
Organization size- with expansion of business, more inter relation is required	**survival stage-** break even stage and increase the business size	**Adaptation-** the traits and techniques adopted by entrepreneurs for survival of organization	**Customer Validation-** scalability of business model	**Validation-** acceptance for product, product refinement, metrics to trace user growth is implemented
Evolution stage- smooth growth of firms	**success stage-**business expansion with strategic planning	**Selection-** the factors that led to success of organization	**Customer Creation-** increases its investment and sales	**Efficiency-** refining values proposition, revamping users' experience, discovering scalable routes
Revolution stage- sudden change due to change in technology, expansion,management	**take off stage-** maintain cash debt ratio, expansion with an adequate delegation of duty	**Retention-** the way the path of success tend to be replicated by other entrepreneurs	**Company Building-** startup transforms itself to the company and makes structural changes to make large establishment.	**Scale-** pitch for larger funds, expand the administrative and infrastructure base
Growth stage- brings creativity, direction, decentralization, coordination and collaboration	**resource maturity-** amalgamation and control of financial resources, the flexibility of response as in small firms.			**Profit Maximisation-** Huge funding round takes place with further expansion of production and operations.
				Renewal/Decline Success leads to IPO; Failiure leads to business decline

Startup Ecosystem in General

A start-up ecosystem, also commonly named as "entrepreneurial ecosystem" **(Mason & Brown, 2014)[25]**, "entrepreneurial system" **(Spilling, 1996[26]; Neck *et al.*, 2004[27])**, "entrepreneurial environment" **(Gnyawali & Fogel, 1994[28]; Fogel, 2001[29])**, or "local entrepreneurial climate" **(Roxas *et al.*, 2007)[30],** consists of factors intermingled to promote entrepreneurship spirit, contribute in startup process and development of entrepreneurship **(Gnyawali & Fogel, 1994)[28].**

Isenberg (2010)[31] defines an entrepreneurial ecosystem (he uses the term "entrepreneurship ecosystem)" as a "set of individual elements such as leadership, culture, financial and human capital, which are intertwined in a complex manner".

Such framework, denominated of "ecosystem", was first coined by James Moore, who claimed that "successful business can't evolve in vacuum, necessitating to attract resources of all types, such as capital, partners, suppliers and customers to create cooperative networks, where companies can work jointly and competitively to support new products, satisfy customer needs, and eventually incorporate the next round of innovations" **(Moore, 1993)[32].**

In general, a typical entrepreneurship ecosystem has innumerable features. The following organizations and activities are the most active ones: Universities, Venture Capitalists, Angel Investors, Incubation Center, Accelerators, Crowd funding, Meetup groups, Services such as legal, financial and consultancy, Competition of start-ups and other supporting organizations.

The research studies are mainly divided into two outlooks to study startup ecosystem: **Attribute Centred Research and Framework Centred Research**.

Start-up Ecosystem Attributes Centered Research

Many researchers look at the perspective to assure specific city; the region is enriched with requisite features to have **start-up clusters, funding support, university or educational institute of excellence (Whittington et al. 2009[33]; Cooke 2002[34])**. Silicon Valley has 6,000 innovative companies. Unique eleven characteristics, values, goals for a healthy ecosystem of Silicon Valley are regarded as Mecca of start-ups. The attributes are **high ranked university(Stanford University); Highly motivated entrepreneurs, investors, and academicians; Ethical cycle of Venture capital industry where entrepreneurs who have gained success invest their exit earnings in new ventures; welfare ethics and quality life with non-hierarchical communication models; entrepreneurial risks embraced; well established infrastructure; savvy entrepreneurs; strong network with free flow of ideas; established patent industry; fearless, risk-taking, idea-driven, resilient people; encouragement for immigrants having skills to shift their base to urban areas (Piscione 2013)[35]** The Silicon Valley successful culture is based not only because of **skilled labour, infrastructure, and technology** but people have **distinctive behaviour** in Valley **(Hwang and Horowitt 2012)[36].**

The advantageous seven attributes of New York City contributes towards its successful start-up ecosystem which includes **easy access to capital, in proximity to world financial centre; Tie with Cornell University and Technion and ambitious Major Bloomberg program to convert New York into biggest innovation hub; tolerance involved for high risk and failure; high networking and love for community; New York Tech Meetup having more than 36,000 members; a big investor group spirit to make good to society; bigger community with approach to global business (Cometto and Piol 2013)[37].** Leading factors endowed in Washington D.C. start-ups ecosystem are its **nascent entrepreneurs, support from society, investors support, universities encouraging entrepreneurial culture; start-up support services (Feldman 2001)[38].** Local Governments in the USA assisted in investments to start-ups to enhance start-up ecosystem culture with flexible policies comprising guidance, infrastructure, business investments and entrepreneurship education. However, the magic of highly established entrepreneurial ecosystems can't be established. Very few internationally successful start-ups have emerged in the USA **(Lerner2009)[39].**

(Feld 2012)[40] in his Boulder thesis and book 'Startup Communities' found four presuppositions necessary for vibrant start-up community: (1). **Entrepreneurs driven start-up community** instead of decisions by policymakers, professors etc. 2) **Penetrative boundary** in society especially among serial entrepreneurs and nascent founders 3) **Organized events** which involve people in the entrepreneurial state of affairs. 4). **The commitment of entrepreneurial community trendsetters** for minimum twenty years. The university of Colorado Boulder provided only indirect support to start-ups by providing space for meetings and events. Ecosystem capacity to enhance innovation drives from a talent pool of human capital, belief beyond the societal barrier, inspiration beyond momentary logic, experiential learning among people and culture that promote to work together **(Hwang and Horowitt 2012)[36]. A helpful culture** like tolerance of risk and mistakes, entrepreneur personality, entrepreneur income; **policies and leadership** through government incentives, government support for start-up etc; **finance** support from angels investors, venture capital, microloans, public capital markets; **human capital** in form of talent pool, training programs; **Strong network among stakeholders**; **Infrastructure support** comprise basis of successful start-up ecosystem **(Isenberg (2011)[41].**

Metropolitan cluster and Sector Specific for flourishing requisite **skills such as IT, Education** is major factors of start-up ecosystem (**Reynolds and Storey 1994)[42]**. However **(Kakati 2003)[43]** found that **Entrepreneur traits, resource-utilisation potential, competitive strategy** form key attributes of success for start-up growth while analysing twenty-seven venture capitalists according to their experience rated one successful and one unsuccessful venture in six areas i.e. entrepreneur traits, resource-utilisation potential, competitive strategy, product features, market traits and financial return. Products compared to competition bring success rather than a unique approach that firm tries to bring. **Accommodation of knowledge from all sectors and regions** has a positive influence on start-up survival and growth as a rapid change in technologies can either drive start-up success or lead to closure **(Folta 2006)[14]**. For start-up survival funding at the right time and amount plays a vital role. Funds from **Venture Capitalist, Angel Investors, and support from Incubators and Accelerators** form 90% of aggregate funds invested in start-ups (**Joshi, K. & Satyanarayana, K. 2014)[44].** The rise in **funding** has a productive transformation on the development of start-ups. **Policy changes** which increase **institutional credit** would help in the creation of new start-ups and increase exports **(Kasturi, S.V. and Bala Subrahmanya, M.H. 2014)[45]** The rise of **New Technology Ventures (NTV)** as well as the development of start-up ecosystem improves

employment contribution and growth of the economy **(Bala Subrahmanya, M H 2015)[46].** Development in the **talent pool** of human resources increases the number of start-ups in the region. Start-up firms which concentrate on intermediate goods and services fuel the growth of the regional economy. The start-ups with a high rate of returns would encourage other entrepreneurial start-ups **(Gries & Naudé, 2009)[47]**.

The above Attributes-centred research attempts to find the presence of factors across regions, cities that attributes essential for growth and survival of start-ups, be physical infrastructures such as university of excellence promoting entrepreneurial culture **(Feldman 1994)[48]**, creative workforce **(Florida 2002)[49]**, or favorable region features **(Saxenian, A. 1994)[50].** The key attempt of such research is to determine such attributes. These attempts will draw the focus of region or city to attain and develop attributes to flare entrepreneurship activities for swift economic growth.

For Attribute-centered research, following attributes are necessary to for startup clusters **(Feldman 1994[48]; Saxenian1994[50]; Cooke 2002[34]; Kakati 2003[43]; Isenberg 2011[41]; Feld 2012[40]; Florida 2002[49]; Braunerhjelm 2006[51]; Whittington et al 2009[33]; Smilor et al 1990[52]; Piscicone 2013[35]; Hwang and Horowitt 2012[36]; Cometto and Piol 2013[37];Feldman2001[38])**

1. Entrepreneurial Culture
2. Skilled Human Personnel
3. Favourable Policies
4. Beneficial City Features
5. Education Excellence Research University
6. Market
7. Funding
8. Rich network connection

Attributes of Startup Ecosystems

Funding

Funding is the lifeline of startup clusters. Financial access is the biggest hurdle to achieving entrepreneurial growth in India (NKC Entrepreneurship Report). Avenues of seed capital are unorganized and mainly government is initial seed capital provider in India (World Bank 2015).

Venture Capital and private equity fund have shown a notable increase in India. However, Investment Trends by venture capital and private equity are basically on established firms having high potential and involves high risk (Nahata, 2008)[53]. Venture Capitalist helps new ventures to overcome problems related to new organization formation and protect them from failure through expertise and in investment (Stinchcombe, 1965)[54]. They provide strategic guidance, open new avenues, help them to acquire legal status, play a major role in startups decisions and are a board of directors of the company (Berglund 2007[55]; Elfring & Hulsink 2003[56]). In India, Venture Funds are regulated by SEBI, 1992 and 1996. In India, there are approximately 200 VC registered with SEBI. According to SEBI regulations, Venture capital fund means fund instituted in the form of a company, trust which lend money as a loan, donation, and securities for investment with regulated guidelines. According to Newscorp VCEdge, a risk capital data monitoring service, Venture capital deals comparison of first six months of year 2014-2016 found that Venture deals were167 of $863million on first six months of 2014, which rose to 278 deals of value $2.8 billion in first six months of 2015 but declined to 183 deals of value $2.8 billion. The fall is due to the focus on unit economics and cash crunch for later stage companies (Chanchani, 2016)[57].

Angel Investor

Angel Investor provides initial capital to early stage ventures. Angel investor invests money in startups in debt or equity form from funds possessed by him. Investment by angel helps to make money and participates in new learning, acts as motivator for startups and community (Shane 2000)[3]. In India startups has raised Rs 113.6 crores through angel investment with 69 deals in year 2016. The total number of deals on angel investments were172 in 2014 which has increased to 368 deals in 2016.

Beneficial city features

Beneficial city features play as a catalyst for firms growth and give an edge to firms from the one located in other regions (Casper 2007[58], Saxenian 2001[59], Feld 2012[40], Piore et al 1994[60]). This factor has been recognized as one of the important features and consistently recognized by researchers to make its mark in regional economic success and decline at the times of economic setback and progress.(Safford 2009[61], McDermott 2007[62]). The local cultural embeddedness of entrepreneurship sparks innovation in the region and encourages policy makers and people associated with the ecosystem to enhance growth through consistently reforming the system (Herriegel 1993[63], Storper 1997[64], Uzzi 2005[65]). Urban agglomeration and discipline expertise are important characteristics of startup region (Reynold and Storey 1993)[42]. External knowledge gets easily accommodated in some of the entrepreneurial region (Cooper and Folta 2000)[66].

Favourable policies

Favourable policies by government play an important role in startup growth. Government actions in form relaxation

of taxes, policies initiative, and encouragement through entrepreneurship program plays key role in firms' growth and survival (Gurley et al.2008)[67]. A higher tax on self employment demotes entrepreneurship. Indian Government has launched action plan on 16th January 2016 to promote startup. Startup friendly policies by present government are: Single window mechanism for Startups, no Capital Gains tax for startups, special incentives for women, bringing incubators investing in particular sectors, startup profits not taxable for first three years, patent fees for startups reduced to eighty percent and Rs 10,000 cr fund for startups @ 2500 crore each year for the first 4 years. (Startup Action Plan 2016)

Entrepreneurial culture

Entrepreneurial culture is necessary for startup growth. An entrepreneur capacity to take risk, drives to achieve goals, accepts ambiguity and thrives to give full effort and time to make organization a successful venture (Schere 1982[68], Mc Clelland 1961[69], Aldrich 1999[23]). Firms' ability to learn, strategy for growth, is essential features to encourage entrepreneurial growth (Sandberg 1986[70], Geroski 1995[71])

Skilled personnel

The skilled personnel of the age between 18-34 years provide resourceful human capital to technology based company. (Cortright and Colleta 2006)[72]. Skilled personnel in different domains provide rich, diverse experiences and helps in generating high-quality life (Saxenian 2001[51], Breshnahan 2001). Rich human capital provides an edge over to confront barriers related to finance as educated people know how to find more avenues to obtain credit, are skillful in forming network and able to find domains to market their products (Kim et al.2006)[74].

Network

Clusters help in technology transfer and encourage learning (Cooke 2002[34], Lester and Piore 1994[64], Breznitz 2005[75]). With increased competition, there is a consistent flow of information and enhance learning and innovation (Asheim and Conenon 2006)[76]. The cluster development helps to build social capital which has multiple fold effect in the growth of startup ecosystem (Lencher and Leyronas 2012)[77]. Multiple ties with a different organization with the entrepreneurial region help in rapidly transmitting collective information and enable startup ecosystem in achieving common goals (Putnam1993)[78].

Incubator

India has 120 incubators and science parks and has funded 2700 companies till 2014 (Fuelling Entrepreneurship 2014). NSTEDB has established 87 Technology based Incubator (TBI) with universities and research institute collaboration and has invested 100 crores. Incubated companies has earned 595 crores till 2014 (Fuelling Entrepreneurship 2014)[79].

Accelerator

Accelerator program designed for 3-6 months to take scalable startup to next level.. Besides incubating the startups, its structured program helps the startups to maintain success. The accelerator besides mentoring the startups help in accomplishing other goals too. The accelerator program runs in batches. The help startups in each stage, from idea development to finding new business opportunities, providing network to gain higher opportunities to scale. It lays hands on team development and host investors meet for start up to pitch and scale. (DST NIDHI)

(Sharma, Joshi, & Shukla, 2013)[80] found that the mortality of the start-ups attending the Accelerator programme is reduced to 20%. Fund raising chances of start-ups through accelerators increase to 78%. Accelerated start-ups have higher chances of getting series A, which is the next round funding. Chances of Accelerated start-up getting acquired are high.

Universities

Universities and educational institutes having entrepreneurship culture helps in the spin-off and provide skilled labor (Breznitz and Alan 2008[81], Huggins et al 2012[82]). Research institutes advance innovation and capabilities. Theses institutes provide skilled personnel and enable to develop clusters (Florida 2002[49], Hervas et al. 2012[83])

Conclusion and Recomendation

Entrepreneurial ecosystem approach cannot be copied from other established system without understanding its evolution under local context (Isenberg 2010)[31]. The situational consideration of ecosystem helps to develop insight on understanding functioning or dis-functioning of the system. The Entrepreneurial or start-up ecosystem is path dependent, and existing structures create barriers in developing new path (Henning et al. 2013)[85].

Various researchers tried to scale entrepreneurial ecosystem. Feld (2012)[40] described cities (Boulder City), (Saxenian 1994)[50] consider region (Silicon Valley) and studied other regions, (Hamel 1999)[86] perceived companies and (Senor and Singer 2009)[84] studied countries (Israel's ecosystem. All outlooks for measuring ecosystem are relevant, and can be encapsulated in cities and shared among the nation. The entrepreneurial ecosystem acts like a magnet. Europe shares international boundaries with many countries and attracts entrepreneurial activities from all interlinked nations (Holmes 2014)[87]. Corporate entrepreneurial ecosystem can reach to many communities in cities and villages. City-

based entrepreneurial ecosystem also stretches its reach to regional, national and international level. The policy level changes may have a conscious and unconscious effect on all levels (Bosma and Stam 2012)[88]. Entrepreneurial ecosystem flourishes through interaction and connections which could be between the ecosystem builders or between communities, regions, nation and international ecosystem.

References

- **Lederman, Daniel, Julidn Messina, Samuel Pienknagura, and Jamele Rigolini (2013)** "Latin American Entrepreneurs: Many Firms but Little Innovation". World Bank Publications.
- **Europe, United Nations Economic Commission for. 2012**. Fostering Innovative Entrepreneurship - Challenges and Policy Options, UNECEF-UN.
- **Shane, S., & Venkataraman, S. (2000).** The promise of entrepreneurship as a field of research. Academy of Management Review, 25(1), 217-226
- **Blank, Steve Gary, and Bob Dorf. (2012**). *The Startup Owner's Manual: The Step-by-Step Guide for Building a Great Company*. K&S Ranch, Incorporated.
- **Start-up Genome (2012):** *Start-up Ecosystem Report 2012*, USA.
- **Brock, W. A. and Evans, D.S. (1989).** Small Business Economics. *Small business economics,* 1,7-20.
- **Acs, R. L. (1992).** Towards a ecological of systems concepts. Management Science, 17(11), 661–671.
- **OECD 2013** Start-up Latin America. Promoting Innovation in the Region. **2013.** Available from: <http://www.oecd.org/innovation/startup-latin-america.htm>.
- **Carree, M. A. and Thurik A.R. (2000).** The impact of entrepreneurship on economic growth. In Audretsch, D.B. and Acs, Z.J. (eds.) *Handbook of Entrepreneurship*, Kluwer Academic Publishers, Boston.
- **Robinson, K C (1998)** An examination of the influence of industry structure on eight alternative measures of new venture performance for high potential independent new ventures, *Journal of Business Venturing*, 16, 165-187.
- **Song, M, K Podoynitsyna, H Bij and J I M Halman (2008)** Success Factors in New Ventures: A Meta-analysis, *The Journal of Product Innovation Management*, 25 (1), 7–27.
- **Bala Subrahmanya, M H (2010).** Internet New Ventures in the USA: Winning Venture Capital Support and *Going Public (1995-2006)*, Report prepared as part of the Fulbright-Nehru Senior Research Fellowship at University of California, Davis, USA.
- **Gilbert, B A, P P McDougall and D B Audretsch (2006)** New Venture Growth: A Review and Extension, *Journal of Management*, 32(6), 926–50.
- **Folta, T.B., Cooper, A.C. and Baik, Y.S. (2006)** Geographic cluster size and firm performance, Journal of Business Venturing, 21, 217-242.
- **Romanelli, E and C B Schoonhoven (2001)** The local origins of new firms, In C B Schoonhoven and E. .Romanelli (Eds) *The entrepreneurship dynamic*, Stanford University Press, Stanford, 40-67.
- **Certo, S T (2003)** Influencing Initial Public Offering Investors with Prestige: Signalling with Board Structures, *Academy of Management Review*, 28(3), 432–46
- **Ministry of Commerce and Industry. (2019).** G.S.R. 127(E) (pp. 6-7). Start- up Definition New Delhi: Department for Promotion of Industry and Internal Trade. www.startupindia.gov.in/content/dam/invest-india/Templates/public/198117.pdf
- **Blank, S. (2010).** What's A Startup? First Principles. Retrieved November 15, 2016, from http://steveblank.com/2010/01/25/whats-a-startup-first-principles/
- **Graham, P. (2012).** Startup = growth. Retrieved November 23, 2016, from http://www.paulgraham.com/growth.html
- **Stokes, D (1995):** Small Business Management, London: DP Publications Ltd.
- **Greiner, L.E. (1972).** Evolution and revolution as organizations grow. Harvard Business Review, 50, 4, 37-46.
- **Churchill, N.C., and Lewis, V.L. 1983**. Five stages of small business growth. Harvard Business Review 83(3), 32-51.
- **Aldrich, H.E. (1999)** Organizations evolving, Sage Publications. London
- **Marmer, M., Herrmann, B. L., Dogrultan, E., Berman, R., Eesley, C., & Blank, S. (2011).** Startup Genome Report 01 -A new framework for understanding why startups succeed, 1–52.
- **Mason, C., & Brown, R. (2012).** Entrepreneurial ecosystems and growth oriented entrepreneurship. Paper prepared for the workshop organised by the OECD LEED Programme and the Dutch Ministry of Economic Affairs, http://www.oecd.org/cfe/leed/ Entrepreneurial-ecosystems.pdf
- **Spilling, O. R. (1996).** The entrepreneurial system: On entrepreneurship in the context of a mega-event. *Journal of Business Research*, *36*(1), 91–103.
- **Neck, H.M., G.D. Meyer, B. Cohen, and A. Corbett. 2004.** An entrepreneurial system view of new venture creation. Journal of Small Business Management 42, 190–208
- **Gnyawali, D. R., & Fogel, D. S. (1994).** Environments for entrepreneurship development: key dimensions and research implications. *Entrepreneurship Theory and Practice*, *18*, 43.
- **Fogel, G. (2001).** An analysis of entrepreneurial environment and enterprise development in Hungary. *Journal of Small Business Management*, *39*(1), 103–109.
- **Roxas, H., Lindsay, V., Ashill, N., & Victorio, A. (2007).** An institutional view of local entrepreneurial climate. *Journal of Asia Entrepreneurship and Sustainability*, *3*(1), 1–28.
- **Isenberg, D. J. (2010).** How to start an entrepreneurial revolution, Harvard Business Review 88(6), 40-50
- **Moore, J. F. (1993).** Predators and prey: a new ecology of competition. *Harvard Business Review*, *71*(3), 75–83.
- **Whittington, Bunker Kjersten, Jason Owen-Smith, and W. Walter Powell. 2009**. "Networks, Propinquity, and Innovation

in Knowledge-intensive Industries." Administrative Science Quarterly 54 (1), 90–122.

- **Cooke, Philip. 2002**. "Biotechnology Clusters as Regional, Sectoral Innovation Systems." International Regional Science Review 2002 (1), 8–37.
- **Piscione, D. P. (2013).** *Secrets of Silicon Valley: What everyone else can learn from the innovation capital of the world*. Palgrave Macmillan, New York
- **Hwang, V W, and G Horowitt. 2012.** The Rainforest: The Secret to Building the Next Silicon Valley. CreateSpace Independent Publishing Platform
- **Cometto, M. T. and Piol, A. 2013.** Tech and the City: The Making of New York's Startup Community, Mirandola Press
- **Feldman, M. 2001.** The entrepreneurial event revisited: Firm formation in a regional context. Industrial and Corporate Change, 10,861–891.
- **Lerner, Josh. 2009**. Boulevard of Broken Dreams: Why Public Efforts to Boost Entrepreneurship and Venture Capital Have Failed – And What to Do About It. Princeton, Princeton University Press, NJ.
- **Feld, B. 2012.** Startup communities: Building an entrepreneurial ecosystem in your city, Hoboken, John Wiley & Sons NJ.
- **Isenberg, Daniel (2011)** "Intelligent Evolution: New Ecosystems for Entrepreneurs," Babson Entrepreneurship Ecosystem Project web page, May 25, 2011, http://entrepreneurialrevolution.com/2011/05/intelligent-evolution.
- **Reynolds, P. D., Storey, D. J., and Westhead, P. (1994).** Cross-National Comparisons of the Variation in New Firm Formation Rates. Regional Studies, 28, 443-456.
- **Kakati, M. (2003)**. "Success Criteria in High-Tech New Ventures." *Technovation* 23 (5), 447–457. doi:10.1016/S0166-4972(02)00014-7. http://linkinghub.elsevier.com/retrieve/pii/S0166497202000147
- **Joshi, K. & Satyanarayana, K. (2014).** What Ecosystem Factors Impact the Growth of High-Tech Start-ups in India?. *Asian Journal Of Innovation And Policy*, *3*(2), 216-244. http://dx.doi.org/10.7545/ajip.2014.3.2.216
- **Kasturi, S.V. and Bala Subrahmanya, M.H. (2014)** 'Start-ups and small scale industry growth in India: do institutional credit and start-ups make a difference?', *Int. J. Entrepreneurial Venturing*, 6 (3), 277–298.
- **BalaSubrahmanya, M H (2015)** New Generation Start-ups in India What lessons can we learn from the past, Economic & Political Weekly, L (12), 56-63.
- **Gries, T. & Naudé, W. (2009).** Entrepreneurship and regional economic growth: towards a general theory of start-ups. *Innovation: The European Journal Of Social Science Research*, *22*(3), 309-328. http://dx.doi.org/10.1080/13511610903354877
- **Feldman, M. P. 1994**, "The University and Economic Development: The Case of Johns Hopkins University and Baltimore," *Econ. Dev. Q.*, 8 (1), 67–76
- **Florida, Richard L. 2002**. The Rise of the Creative Class. Basic Books, New York.
- **Saxenian, A. (1994).** Regional Advantage: Culture and Competition in Silicon Valley and Route 128. Cambridge, Massachusetts, US: Harvard University Press.
- **Braunerhjelm, Pontus, and Maryann P. Feldman, eds. 2006**. Cluster Genesis: Technology-Based Industrial Development,Oxford University Press, Oxford
- **Smilor, W.R, Gibson, V.D, Dietrich, B.G (1990)** University spin-out companies: Technology start-ups from UT-Austin Journal of Business Venturing 5(1):63-76 DOI: 10.1016/0883-9026(90)90027-Q
- **Nahata, R. (2008),** "Venture capital reputation and investment performance", Journal of Financial Economics, 90 (2), 127-151.
- **Stinchcombe, A. (1965),** "Social structure and organizations", in March, J.G. (Ed.), The Handbook of Organizations, Rand-McNally, Chicago, IL, 142-193.
- **Berglund, H., Hellström, T. and Sjölander, S. (2007),** Entrepreneurial learning and the role of venture capitalists, *Venture Capital*, 9(3), 165-181.
- **Elfring, T. and Hulsink, W. (2003),** "Networks in entrepreneurship: the case of high-technology firms", Small Business Economics, 21 (4), 409-420.
- **Chanchani, M. (2016).** Venture capital investments in startups have shrunk both in value and volume: Report Read more at: http://economictimes.indiatimes.com/articleshow/52999586.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cpps. *ET Bureau*. Retrieved from http://economictimes.indiatimes.com/small-biz/startups/venture-capital-investments-in-startups-have-shrunk-both-in-value-and-volume-report/articleshow/52999586.cms
- **Casper, Steven. 2007.** "How Do Technology Clusters Emerge and Become Sustainable? Social Network Formation and Inter-Firm Mobility Within the San Diego Biotechnology Cluster." Research Policy 36, 438–455.
- **Saxenian, A., & Hsu, J. Y. (2001)**. The Silicon Valley-hsinchu connection: Technical communities and industrial upgrading. Industrial and Corporate Change, 10, 893-920. doi:10.1093/icc/10.4.893
- **Piore, Michael J., Richard K. Lester, Fred M. Kofman, and Kamal M. Malek. 1994**. "The Organization of Product Development." Industrial and Corporate Change 3 (2),405–434.
- **Safford, Sean. 2009.** Why the Garden Club Couldn't Save Youngstown: The Transformation of the Rust Belt. Cambridge, MA: Harvard University Press.
- **McDermott, Gerald A. 2007.** "The Politics of Institutional Renovation and Economic Upgrading: Recombining the Vines that Bind in Argentina." Politics and Society 35 (1), 103–143.
- **Herrigel, Gary B. 1993**. "Power and the Redefinition of Industrial Districts: The Case of Baden- Wurttemberg." In The Embedded Firm: On the Socioeconomics of Industrial Networks, edited by G. Grabher, 227–251. London: Routledge
- **Storper, Michael. 1997**. The Regional World: Territorial Development in a Global Economy. New York: Guilford Press

- **Uzzi, Brian, and Jarrett Spiro. 2005.** "Collaboration and Creativity: The Small World Problem." American Journal of Sociology 11 (2), 447–504.
- **Cooper, A. and Folta, T. (2000)** Entrepreneurship and High-Technology Clusters, Blackwell Handbook of Entrepreneurship, 348-367
- **Gurley-Calvez, T. and Bruce, D. (2008)** Do tax cuts promote entrepreneurial longevity? National Tax Journal, 61(2), 225-250
- **Schere, J, 1982,** Tolerance of ambiguity as a discriminating variable between entrepreneurs and managers, Proceedings of the Academy of Management Conference, New York, NY, pp. 404-408
- **McClelland, D. C. (1998)** Identifying competencies with behavior event interviews. Psychological Science, 9, 331–339.
- **Sandberg, W.R. (1986)** New Venture Performance: The Role of Strategy and Industry Strurture, Lexington, MA: Lexington Books
- **Geroski, P.A. (1995)** What do we know about entry? International Journal of Industrial Organization, 13, 421-440
- **Cortright, Joseph, and Carol Coletta. 2006.** The Young and the Restless: How Atlanta Competes for Talent, Impresa Consulting, Atlanta, GA.
- **Bresnahan, Timothy, Alfonso Gambardella, and Annalee Saxenian. 2001**. "'Old Economy' Inputs for 'New Economy' Outcomes: Cluster Formation in the New Silicon Valleys." Industrial and Corporate Change 10 (4), 835–860.
- **Kim, P.H., Aldrich, H.E. and Keister, L.A. (2006)** The impact of financial, human and cultural capital on entrepreneurial entry in the United States, Small Business Economics, 27, 5-22
- **Breznitz, Dan. 2005** Collaborative Public Space in a National Innovation System: A Case Study of the Israeli Military's Impact on the Software Industry,Industry and Innovation, 12 (1), 31–64.
- **Asheim, Bjørn T., and Lars Coenen. 2006**. Systems in a Globalising Learning Economy: On Knowledge Bases and Institutional Frameworks. Journal of Technology Transfer, 31, 163–173.
- **Lechner, Christian, and Chrisophe Leyronas. 2012.** "The Competitive Advantage of Cluster Firms: The Priority of Regional Network Position Over Extra-Regional Networks – A Study of a French High-Tech Cluster." Entrepreneurship & Regional Development 24 (5–6), 457–473.
- **Putnam, Robert D. 2001**. Bowling Alone: The Collapse and Revival of American Community. Simon and Schuster New York.
- **Fuelling Enterpreneurship (2014)**, National Science and Technology Entrepreneurship Development Board (NSTEDB)
- **Sharma, A., Joshi, M., & Shukla, B. (2014)** Is Accelerator an Option? Impact of Accelerator in Start-up Eco-System!. *SSRN Electronic Journal*. http://dx.doi.org/10.2139/ssrn.2438846
- **Breznitz, S. M., Rory P. O'Shea, and Thomas J. Alan. 2008**. "University Commercialization Strategies in the Development of Regional Bioclusters." Journal of Product Innovation Management 25 (3), 129–142.
- **Huggins, Robert, Andrew Johnston, and Chris Stride. 2012**. "Knowledge Networks and Universities: Locational and Organisational Aspects of Knowledge Transfer Interactions." Entrepreneurship & Regional Development 24 (7–8), 475–502
- **Hervas-Oliver, Jose-Luis, Jose Albors-Garrigos, Blanca de-Miguel, and Antonio Hidalgo. 2012.** "The Role of a Firm's Absorptive Capacity and the Technology Transfer Process in Clusters: How Effective are Technology Centres in Low-Tech Clusters?" Entrepreneurship & Regional Development 24 (7–8), 523–559.
- **Senor, D., & Singer, S. (2009).** Start-Up Nation: The Story of Israel's Economic Miracle. Random House LLC.
- **Henning, M., Stam, E. & Wenting, R. (2013)** Path dependence research in regional economic development: Cacophony or knowledge accumulation? Regional Studies 47 (8), 1348-1362.
- **Hamel, G. (1999).** Bringing Silicon Valley Inside. Harvard Business Review, Sep-Oct, 71-84.
- **Holmes, B. (2014).** A Different Tech Model In Europe. Wall Street Journal
- **Bosma, N., Hessels, J., Schutjens, V., Van Praag, M. and Verheul, I. (2012)** Entrepreneurship and role models. Journal of Economic Psychology 33(2), 410-424.

GJEIS Prevent Plagiarism in Publication
The Editorial Board had used the turnitin plagiarism [http://www.turnitin.com] tool to check the originality and further affixed the similarity index which is 10% in this case (See Annexure-I). Thus the reviewers and editors are of view to find it suitable to publish in this Volume-11, Issue-2, Apr-June, 2019

Annexure 1

Submission Date	Submission Id	Word Count	Character Count
6-April-2019	1172230268 (turnitin)	5847	35924

ORIGINALITY REPORT

4%	2%	1%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	apjor.com Internet Source	1%
2	Submitted to Carnegie Mellon University Student Paper	<1%
3	www.jmveindia.com Internet Source	<1%
4	Submitted to Deakin University Student Paper	<1%
5	Neeta Baporikar. "Framework for Social Change through Startups in India", International Journal of Civic Engagement and Social Change, 2015 Publication	<1%
6	www.researchgate.net Internet Source	<1%
7	Submitted to University of Witwatersrand Student Paper	<1%
8	eadtu.eu Internet Source	<1%
9	www.emeraldinsight.com Internet Source	<1%
10	Submitted to Lappeenrannan teknillinen yliopisto Student Paper	<1%
11	"Entrepreneurship in Transition Economies", Springer Nature, 2017 Publication	<1%

Reviewers Comments

Reviewer's comment 1:
The paper is crafted in a structured manner. The introduction portion is quite impressive.

Reviewer's comment 2:
The paper has tried to cover almost all the attributes of start-up ecosystem. And every attribute has been explained through supportive studies.

Reviewer's comment 3:
Studies on management and development about the startup ecosystem framework are the need of the hour. And author has tried to satisfy this need by the manuscript which critically analyze the research studies describing the stages of start-ups, attributes of start-up ecosystem and their inter-relationship.

Citation

Surabhi Kapoor and Seema Singh
"Exploring Start-up Ecosystem and its Structural Impact : A Review"
Volume-11, Issue-2, Apr-June, 2019. (www.gjeis.com)

https://doi.org/10.18311/gjeis/2019
Volume-11, Issue-2, Apr-June, 2019
Online ISSN : 0975-1432, **Print ISSN** : 0975-153X
Frequency : Quarterly, Published Since : 2009

Google Citations: Since 2009
H-Index = 96
i10-Index: 964

Source: https://scholar.google.co.in/citations?user=S47TtNkAAAAJ&hl=en

Conflict of Interest: Author of a Paper had no conflict neither financially nor academically.

Scholastic Seed Inc.

https://orcid.org/0000-0002-7952-096X

Engineering Research Express

PAPER

Fabrication and optimization of AA7075 matrix surface composites using Taguchi technique via friction stir processing (FSP)

To cite this article: Ravi Butola *et al* 2019 *Eng. Res. Express* **1** 025015

View the article online for updates and enhancements.

This content was downloaded from IP address 14.139.251.100 on 24/10/2019 at 12:47

Engineering Research Express

PAPER

Fabrication and optimization of AA7075 matrix surface composites using Taguchi technique via friction stir processing (FSP)

RECEIVED
19 August 2019

REVISED
23 September 2019

ACCEPTED FOR PUBLICATION
4 October 2019

PUBLISHED
17 October 2019

Ravi Butola , Ranganath M S and Qasim Murtaza
Mechanical Engineering Department, Delhi Technological University, New Delhi, India
E-mail: ranganathdce@gmail.com

Keywords: Friction stir processing, Rice Husk Ash, silicon carbide, boron carbide

Abstract

In the present research, work is carried out for the parametric effects like tool rotation speed, tool profile and reinforcement in friction stir processing (FSP) of AA7075 on the microhardness and microstructural properties. The experiment design was carried out using Taguchi L_9 orthogonal array with three factors at three levels. Three different tool pin profiles viz. circular, triangular and square were used to experiment. The AA7075 matrix material was reinforced with SiC, RHA, and B_4C. The tool rotational speed was chosen as 700, 1000 and 1300 rpm. Using the Taguchi method, the optimum level of each of these parameters was obtained. It was found that rotational speed is the most influential parameter followed by the type of reinforcement and tool profile respectively. The average microhardness of AA7075/B_4C composite was found to be increased by 1.5–1.6 times of the base indicating that B_4C is better compared to SiC and RHA in improving the microhardness value of the composites.

1. Introduction

Materials having a good ratio of strength to weight are exhibiting tremendous importance in modern-day applications especially in the field of automobile and aerospace engineering where importance is placed on improved efficiency and reduced fuel consumption [1]. The fields of automotive, aerospace, defense, marine and recreation require materials with enhanced characteristics of wear resistance, fatigue resistance, temperature withstanding ability, and fracture toughness with cost considerations [2, 3]. Modern infrastructure, equipment, and machinery require a complex mix of material properties to meet the demands [4]. Metal Matrix Composites are an entirely new generation of composites that can satisfy the broad spectrum of demands of the modern-day applications with their characteristic isotropic properties, and try to develop HMMC using ceramics particles and natural fibres like SiC,RHA and, GSA as a reinforced material [2, 5, 6]. Al-based MMC have aluminum as matrix metal that has properties which are well concerned with the automobile industry [3]. The highlighting properties of aluminum and magnesium matrix composites are lower density, better wear resistance, better corrosion resistance, and high formability, especially under condition, which promote creep for which these can be employed for automotive, aerospace and various other technologies [3, 6]. Aluminum matrix composites can be reinforced with silicon carbide, boron carbide, aluminum oxide, titanium carbide, titanium oxide, etc there is an enhancement in the properties of hybrid composites so formed on addition of these reinforcing materials to aluminum composites [7].

A schematic of the Friction stir process is shown in figure 2(a) and FSW setup shown in figure 2(b)illustrates the basic working principle of FSP. Friction stir processing has been developed as a non- consumable rotating tool was inserted into the base metal that developed severe plastic deformation of the materials due to combined action of tool stirring and friction heat for materials processing and modification. Stirring was created through the tool pin and heat was generated by the axial force of the tool. After achieving the direction from advancing to retreating side of the base metal, a non- consumable tool, rotating at high rpm and traverse along the length of the base metal and to achieve the desired surface composite [8, 9]. Around 80 million tons of RHA is produced

© 2019 IOP Publishing Ltd

annually worldwide which poses as a major environmental threat to the areas in which it is dumped. Thus, the use of RHA for making composites is beneficial to the environment.

There are various ways to synthesize reinforced AMCs, some of the techniques used come from liquid metallurgy processing routes and solid route processing. There are many liquid metallurgy processing techniques, which include the process of stir casting, compos casting, liquid infiltration and many more [1]. Solid route processing includes the method of powder metallurgy and friction stir processing can be used on aluminum alloy [10]. Without disregarding the technical proficiency of any of the processing routes available, friction stir processing is the technique used for synthesizing AMCs. Friction stir processing is a short route, solid-state processing technique with one-step processing that can achieve microstructural homogeneity and refinement densification [11]. Friction Stir Processing (FSP) is one of the most effective techniques for fabrication of MMCs. FSP can be considered as a derivative of Friction stir welding, which was used to join materials with low melting temperatures such as aluminum alloys [12]. Surface Composites are being widely used nowadays, owing to their improved surface properties, and FSP provides an effective way for their fabrication [13]. This process makes up for the shortcomings of liquid-phase processing techniques, employment of a process for composite fabrication which can be conducted below melting points of given matrix material can go a long way to improve and consequently optimize the MMCs design and fabrication issue [14]. A non-consumable tool having specially designed pin as well as shoulder undergoes rotational motion during the FSP [15]. During FSP, softening and plastic deformation of the base materials occur owing to frictional heat generation between the tool rotation and base materials [16]. As the tool traverses, the material which comes in contact with the front face of the pin undergoes plastic deformation as a consequence of the friction between the tool and the workpiece. This FSP region results in a stirred zone (SZ) with a fully recrystallized, equiaxed and very fine-grain microstructure and refinement of constituent particles [17, 18].

Friction stir processing has also been used in various types of particle reinforced aluminium composite materials and among them, SiC particle reinforced aluminium alloys have been the ones which are studied the most. Wang *et al* [19] concentrated their work on the feasibility of making silicon carbide (SiC) reinforced aluminium metal matrix composites. Good interface bonding was observed between the particles and the base metal. The volume percentage of silicon carbide particles is kept at 1.5% in the reinforced region. The microhardness value of the reinforced metal matrix composites was found to be nearly about 10% higher as compared to the base metal Al-6Mg-Mn. Shafiei-Zarghani et al [20] had used Friction stir processing (FSP) to incorporate the nano-sized Al_2O_3 particles into AA6082 aluminium alloy in order to form a particulate composite surface layer. The Al_2O_3 particles had an average particle size of approximately 50 nm. The microhardness value of the fabricated composite surface had improved by nearly three times when compared to that of the non-processed aluminum alloy. After this work, a lot of work on surface composite fabrication is reported. Initially, FSP was used to modify aluminium alloys but over time, FSP gains shining role in the modification of other alloys like magnesium alloys, copper alloys, titanium alloys and even steel [21]. Optimum combinations of process parameters were determined for higher microhardness and ultimate tensile strength respectively using signal/noise ratio graphs, response tables, and analysis of variance method [22]. AA6xxx series aluminium alloys are heat treatable whose properties are affected with the exposure of heat and experience annealing effect during FSP. With the reduction of heat exposure time, i.e. reduction of grain growth time or accelerating the cooling effect, larger grain refinement occurs which results in enhanced mechanical characteristics of SCs [23]. Analysis of microstructures and distribution of SiC particles in the fabricated Surface composite were studied by optical microscope and FESEM, respectively and found in good corroboration with the mechanical properties [24]. Some researchers have been reported that different reinforcement like SiC, TiC, ZrO2, and B4C particle are investigated and also to study the effect of different pin profile on the distribution of reinforcement and it concluded that TiC particle exhibit higher hardness value compared to others [25]. In another researcher investigated that the fabrication of AA6061/TiB2 composite and the effects of process parameter like tool rotation, tool traverse, axial load and reinforcement on mechanical and microstructure were studied. Using ANOVA results it indicates that tool rotation speed was found to be the most influencing process parameter followed by the reinforcement, axial load and traverse speed [26].

The objective of this research is to establish the effects of Friction stir processing (FSP) parameters viz. tool profile, reinforcements and tool rotation speed on the microhardness of AA7075 surface modification and to the utilization of ceramics and agro-waste particles to influence the properties of the base material.

2. Experimental method

The experiment was conducted in two phases. In the first phase, capping pass was performed with a pin-less tool on each groove to cover the reinforcement and maximum load during the capping pass was 590 N. In the second

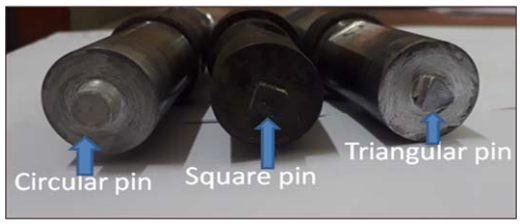


Figure 1. Different tool pin profile used for experimental work.

Table 1. Composition of AA7075.

Element	Al	Zn	Mg	Cu	Si, Fe, Mn, Ti, Cr, and other
%	90	5.6	2.5	1.6	0.23

Table 2. Taguchi's L_9 Orthogonal array.

Exp. no	Rotation speed (rpm)	Tool profile	Reinforcement
1	700	Triangular	SiC
2	700	Circular	RHA
3	700	Square	B_4C
4	1000	Triangular	RHA
5	1000	Circular	B_4C
6	1000	Square	SiC
7	1300	Triangular	B_4C
8	1300	Circular	SiC
9	1300	Square	RHA

phase, the tool profile and rpm of the tool were adjusted per the design of the experiment prepared in table 2, and the maximum load during processing is around 850 N. For both the phases, traverse speed was kept constant at 40 mm min^{-1}. The tilt angle for capping pass was kept at 0° while for the second phase it was set at 2°. The process parameter were fixed at constant value given in table 3. In figure 1, different pin profile FSP tool was made of H13 tool steel is hot worked chromium-molybdenum steel that is prominently used in hot work as well as cold work tooling applications. The hot hardness of this tool steel resists thermal fatigue cracking which occurs due to the repeated cycles of heating and cooling while performing hot working.

2.1. Matrix material

The base material used here is AA7075, which has zinc as its primary alloying element. The size of workpiece used is (200 × 80 × 8) mm as per the requirements of Friction stir processing setup. Grooves were made in the workpiece lengthwise with the help of a milling slitting cutter. Milling cutter had a bore (inner diameter) of 1.25 inch and outer diameter of 4 inch. The grooves made, were of 3.5 mm depth and 2 mm width. The alloy composition of AA7075 is given in table 1.

2.2. Reinforcements

Several agro-waste and industrial materials have been deeply investigated to use them as reinforcement but the prominent ones include coconut shell ash(CSA), bagasse ash and rice husk ash (RHA) [27, 28]. Several experiments have been carried out by the researchers and several other researchers on the utilization of ground-nut shell ash (GSA) as reinforcing materials [1–3]. These composites provide improved physical properties such as low density and low thermal coefficient of expansion. Moreover, their cost is reduced too. Mechanical properties and features of the composite has been investigated to establish the viability of RHA, SiC and B_4C in development of the composites.

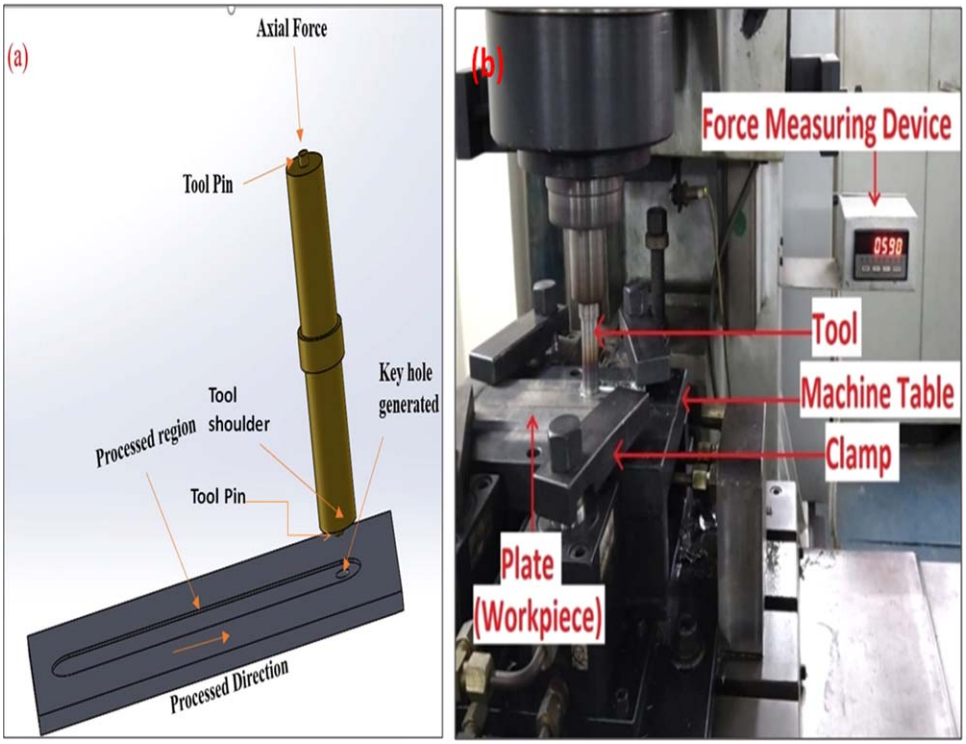


Figure 2. (a). Schematic of Friction stir processing, 2(b) FSW Setup.

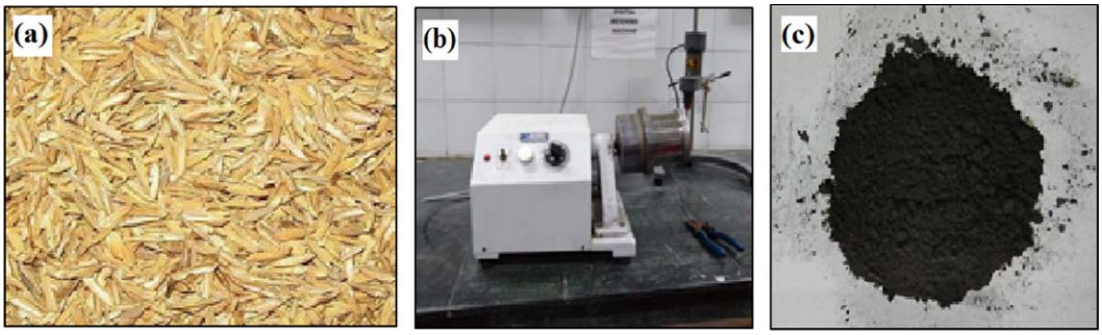


Figure 3. (a) Rice Husk Ash (b) Ball milling machine (c) Fine Rice husk ash powder obtained after sieving.

Table 3. Process parameters.

Process paramters	Value taken
Tool shoulder profile	Flat
Tool shoulder dia	19.95 mm
Tool pin profile	Triangular, Circular, Square
Tool pin length	3.5 mm
Tool plunge depth	0.50 mm
Tool tilt angle	2° constant
Type of reinforcement	SiC, RHA, and B_4C

2.2.1. Rice Husk Ash (RHA)

Rice Husk was obtained from rice mill near Hisaar shown in figure 3(a). It was burned overnight and the ash was collected in a plastic container. To reduce the size of the RHA particles shown in figure 3(c), ball milling shown in figure 3(b)were done in sand testing lab.

Ball milling works on the principle of impact and size reduction of the particles as steel balls jiggle in a cylindrical shell. Ball milling was performed for 8 h using 16 steel balls. Sieve testing of the obtained ash particles

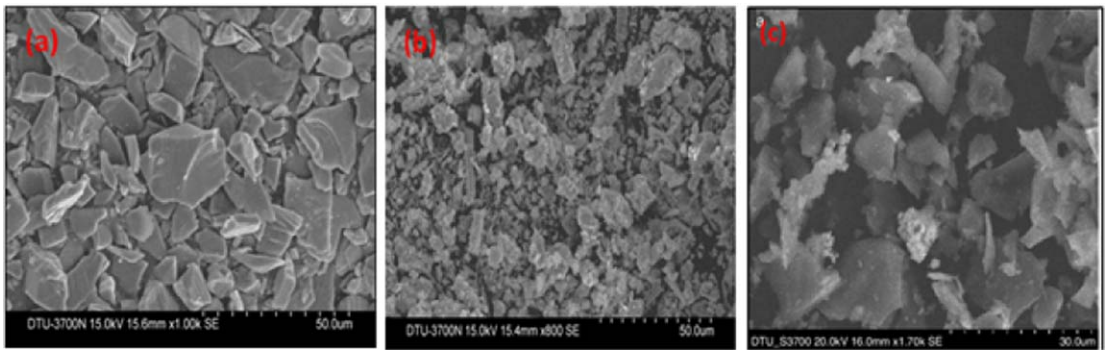


Figure 4. (a) SEM image of SiC at 50 um (b) SEM image of RHA at 50 um (c) SEM image of B_4C at 30 um.

was performed to segregate particles of size 53 or less microns. For this purpose, a sieve of aperture size ASTM 270 was used. SiC, RHA, B_4C SEM images are shown in figures 4(a)–(c), were used as a reinforcement.

2.2.2. Silicon carbide (SiC)

Silicon carbide is also known by the name of carborundum, and contains silicon and carbon. The main use of SiC powder is it being used as an abrasive. The mesh size of SiC used is 400, which corresponds to the average particles size of 37 microns.

2.2.3. Boron carbide (B_4C)

Boron carbide is known to be the third hardest material and is a boron carbon ceramic. It is used in bullet-proof vests, tank armor and engine sabotage powders. For the experiment, B_4C of average particles size less than 30 nm was used.

3. Experimental design using Taguchi methods

DOE is an experiment technique in which we analyze the effect of process parameters by varying them simultaneously at different levels. In this manner we are able to capture how different factors interact with each other. Three three-level parameters (rotation speed, tool profile, type of reinforcement) are selected for this study using L_9 orthogonal array design. It should be mentioned that only the main factor effects are taken into consideration and not the interactions. A total of nine runs of the experiment need to be conducted using the combination of levels for each process parameters. The full factorial experimental design would have required $3^*3^*3 = 27$ runs. Thus using Taguchi experiment design, we have eliminated 18 FSP runs. This helps to save a lot of time and resources. The experimental design carried out with the help of MINITAB 17 software.

3.1. Filling of reinforcements in the groove

Before filling the reinforcements, the plates were cleaned by keeping them dipped in 500 ml of acetone overnight. The filling of reinforcements has been done considering volume fraction in case of RHA and SiC while for B_4C the same was done considering molar solution.

$$Volume\ Fraction = \frac{w^*\mathrm{d}}{t^*\mathrm{d}_p} \tag{1}$$

where,

w denotes width of slot cut
d is depth of slot
t is pin length
d_p is effective diameter of pin

Based on experimental work, volume fraction of 20% were chosen for filling RHA and SiC in the slots whereas in case of B_4C, one molar solution with acetone was prepared and 5 ml solution was poured in each of the three slots.

3.2. Samples preparation for microhardness and microstructure

Specimens of dimensions 10 mm × 10 mm × 4 mm were cut from the processed region through wire EDM cutting. Dry polishing of the specimens were done using emery paper of grit sizes 220, 320, 400, 600, 800, 1000,

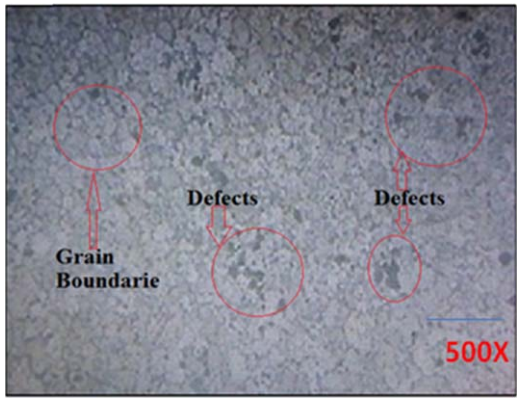


Figure 5. Microstructure of base metal at 500× resolution.

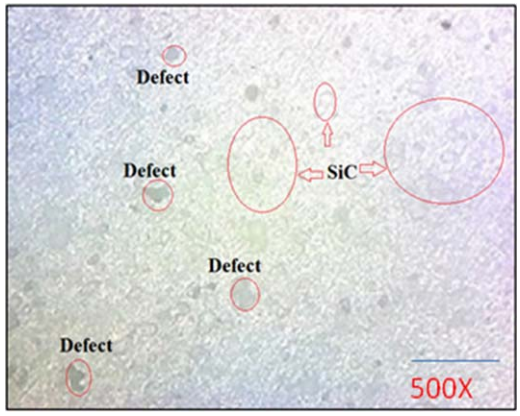


Figure 6. Microstructure of specimen 1 at 500× resolution.

1200, 1500 and 2000. After dry polishing, wet polishing is done on abrasive wheel to obtain smooth surface of specimens.

4. Results and discussion

4.1. Microstructure

The square cross-section samples were prepared especially for their microstructure analysis. Samples for the microstructure analysis were cut perpendicular to the processed line at the center of the plate. After the wet polished specimens were then etched in an acidic medium comprised of Hydrochloric acid and Nitric acid mixed in an optimal molar ratio. The process of etching helps to identify the presence of grain boundaries, which is shown in figure 5. Viewing the microstructure was important because one of the most significant effects of FSP was to be the resultant refinement of the grain structure and reinforcement particles like SiC,RHA and B_4C, which is shown in figure 6, figure 7 and figure 8. This would be translated into superior mechanical properties imparted to the composite material. For microstructure observations, we used an Olympus GX41 inverted metallurgical optical microscope. The images were taken at 500× for all the square cross-section samples.

4.2. Microhardness

The Vickers microhardness testing of the specimens was performed using Fischer scope HM2000 S microhardness tester. The indenter makes an impression on the polished specimen surface under a designated load for a specified period. Microhardness test has been carried out at 0.300 N or 0.03 kg and the dwell time was 20 s. Repeated measurements of the hardness value were taken at different positions on the specimen surface and subsequently, the arithmetic mean was used for further calculations. The average hardness was calculated using 3 readings on the same load. The average hardness value of FSP processed sample is found to be more than that of base metal alloy. The reason behind this result is the fine dispersion of B_4C hard particle in the aluminium

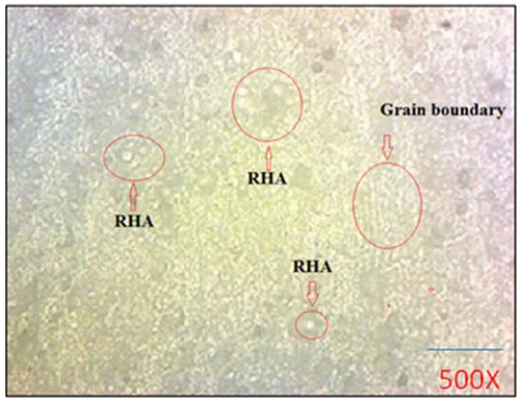


Figure 7. Microstructure of specimen 2 at 500× resolution.

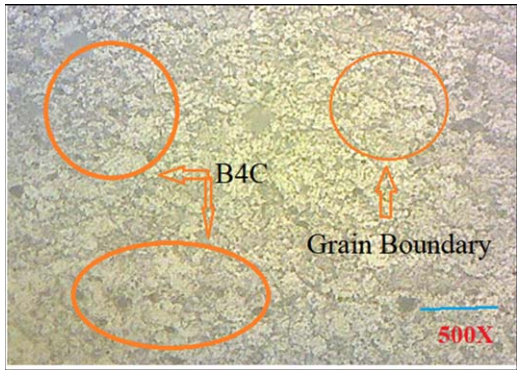


Figure 8. Microstructure of specimen 3 at 500× resolution.

substrate. Increase in hardness value of FSPed sample is also due to the Hall-Petch strengthening or grain boundary strengthening. In this, the material strengthens through the changing grain size. In FSP due to friction between base metal and tool probe grain at the surface of the matrix, become fine. The hardness of the parent metal was 68.7 Hv. During the microhardness test, an average of 3-hardness value was taken in the entire sample, see figure 11. Show the hardness bar graph of the entire processed and unprocessed sample with micro and nanoparticle. This observation in micro and Nanoparticle size reinforcement is investigated by literature review in various other material systems. Asadi *et al* [29] have been reported that the hardness of SiC nanocomposite to be increased due to the uniform distribution of SiC particle and refinement of grain size. Dolatkhah *et al* [30] have achieved a maximum hardness of 116 Hv in the surface composite of Al 5052 with 50 nm SiC particle size due to grain reduction of the particle. Shamsipur *et al* [31] reported that the microhardness value of Ti/SiC nano surface composite is three-time greater than the base metal due to the uniform distribution of nano-size reinforcement particles size. Due to the presence of nanoparticles, size reinforcement contributes to producing ultra-fine grain size. The microhardness of the surface composite using B4C nanoparticle reinforcement is higher comparison with SiC or RHA microparticles.

For optimization purpose, Taguchi method utilizes signal to noise (S/N) ratio. The signal to noise ratio signify that the effects of noises on the average responses are computed by the influence on the deviations from the average responses, which will help to gauge the responsiveness of the output of experiment relative to the involved noise factors [32]. In this study, microhardness value of the friction stir processed (FSPed) material is measured. The S/N ratio of the performance characteristic is analyzed according to the criterion higher-the-better which can be expressed as:

$$\frac{S}{N} = -10 \times log\left(\frac{1}{n}\sum_{n}^{i=1}\frac{1}{Y_i^2}\right). \tag{4.2.1}$$

where, n = number of repetitions

Y_i = value of performance characteristic for the test in ith run

Table 4. Experimental results and calculate mean and S/N ratio.

S. No.	Rotation speed (rpm)	Tool Profile	Reinforcement	HV (mean)	S/N Ratio
1	700	Triangular	SiC	75.3	37.536
2	700	Circular	RHA	73.8	37.361
3	700	Square	B_4C	87.6	38.850
4	1000	Triangular	RHA	79.2	37.974
5	1000	Circular	B_4C	122.8	41.784
6	1000	Square	SiC	136.1	42.677
7	1300	Triangular	B_4C	110.1	40.856
8	1300	Circular	SiC	74.9	37.489
9	1300	Square	RHA	76.2	37.639

Table 5. Main effects response table of HV (mean).

Level ↓	Rotation speed	Tool profile	Reinforcement
1	78.90	88.20	95.43
2	112.70	90.50	76.40
3	87.07	99.97	106.83
Delta	33.80	11.77	30.43
Rank	1	3	2

Table 6. Main effects response table of HV (S/N ratio).

Level ↓	Rotation speed	Tool profile	Reinforcement
1	37.92	38.78	39.23
2	40.81	38.88	37.66
3	38.65	39.72	40.49
Delta	2.90	0.94	2.83
Rank	1	3	2

Using the above expression S/N ratio for HV value of each specimen is calculated. The result obtained from micro-hardness testing of the samples is provided in table 4.

Minitab 17 software is used to carry out analysis of this Taguchi experiment design. Table 5 represents the response table for mean HV value. The response table clearly explains that rotation speed is most influential factor in the determination of HV value followed by type of reinforcement and tool profile.

Table 6 is the response table for the S/N ratio. The S/N ratio is used as a performance measurement characteristic to develop process which is insensitive to the variance factors. For this experiment the process parameter with high S/N ratio gives the optimum setting with least variance. The response table clearly indicates that rotation speed is the most influential factor followed by type of reinforcement and tool profile.

The main effects plot for means (figure 9) shows the variation of response variable (HV value) with process parameters i.e. rotation speed, tool profile and type of reinforcement. The horizontal dashed line represents the overall mean value of the response variable. From the plot it is evident that HV value is maximum at second level of rotation speed (1000 rpm), third level of tool profile (square profile) and third level of reinforcement type (B_4C).

Figure 10 show the main effects plot for S/N ratios. From the plot it is evident that S/N ratio is maximum at second level of rotation speed (1000 rpm), third level of tool profile (square profile) and third level of reinforcement type (B_4C).

Repeated measurements were also done to determine the HV value of the base metal and it was found to be 68.7. A bar graph is plotted to visualize the HV values of the nine specimens with respect to the base metal. It is clearly evident from figure 11, that the microhardness value of all FSPed samples is improved with respect to the base metal.

The improvement achieved in micro hardness value of the processed zone is very significant and evident from the bar graph shown in figure 11. The results of this study are in complete agreement with the previous works carried out by distinguished researchers.

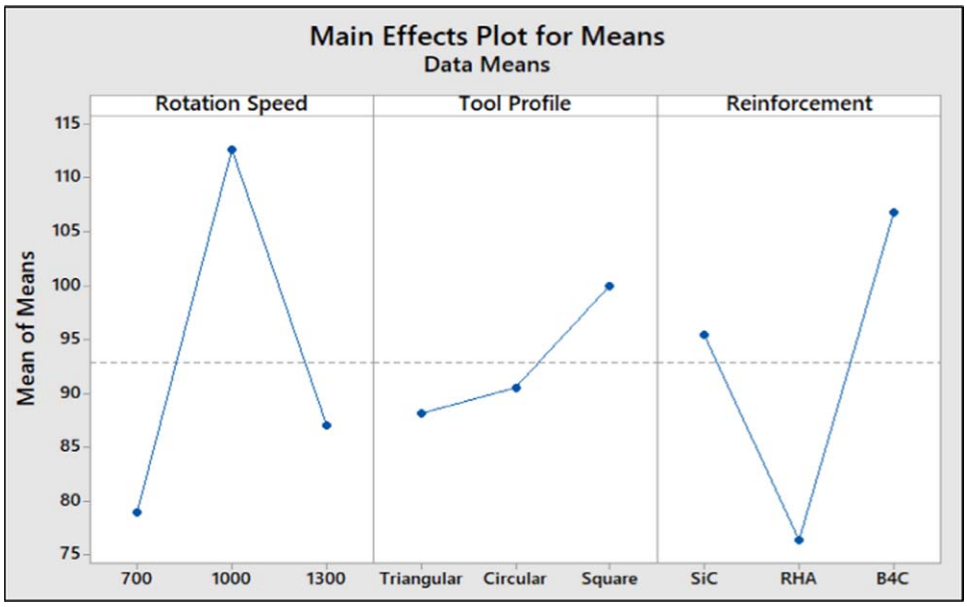


Figure 9. Main effects plot for means.

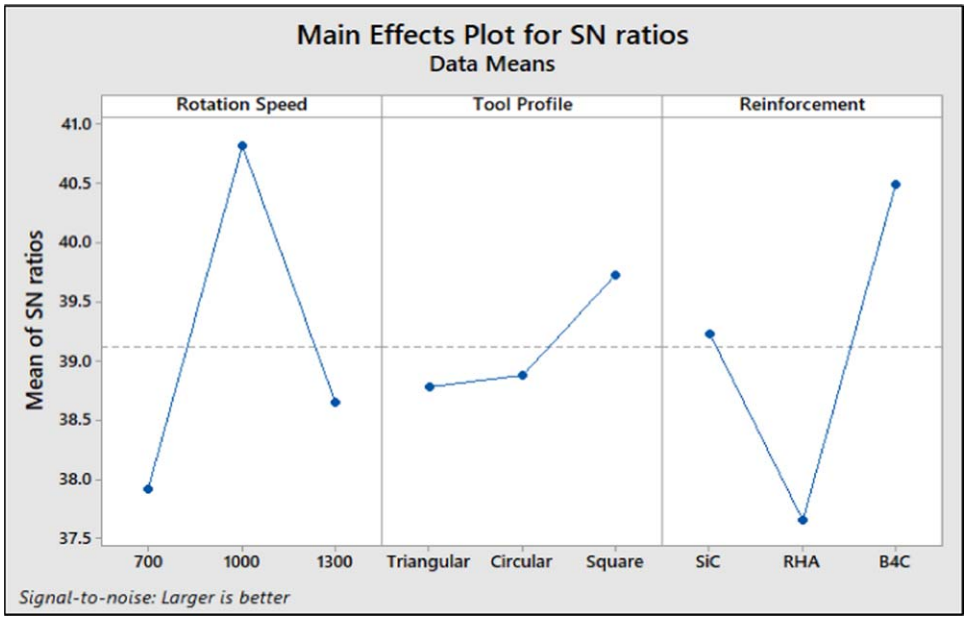


Figure 10. Main effects plot for S/N ratios.

4.3. Analysis of variance (ANOVA)

Each of the FSP process parameter responses were estimated and significant contribution using ANOVA to understand their significance. The ANOVA responses along with the related terms are summarised in table 7. Percentage contribution of individual FSP parameter for different responses was summarised in the ANOVA.

Table 7 indicates that tool rotational speed in rpm as the most influential factor for microhardness followed by the type of reinforcement and tool profile. In the previous literature survey, ANOVA was used to evaluate the process parameter and find out the most influencing parameter for micro hardness and UTS [21]. Taguchi method was used to establish the relation between input and output parameter and ANOVA was used to determine the significance of the parameters [33]. In another research, using ANOVA was performed to find the influence of reinforcement on the mechanical properties like hardness, % elongation, UTM and yield strength of the composite material [28]. It is a systematic method to determine the relationship between factors affecting a process and the output of that process. It is used to find cause-and-effect relationships. This information is

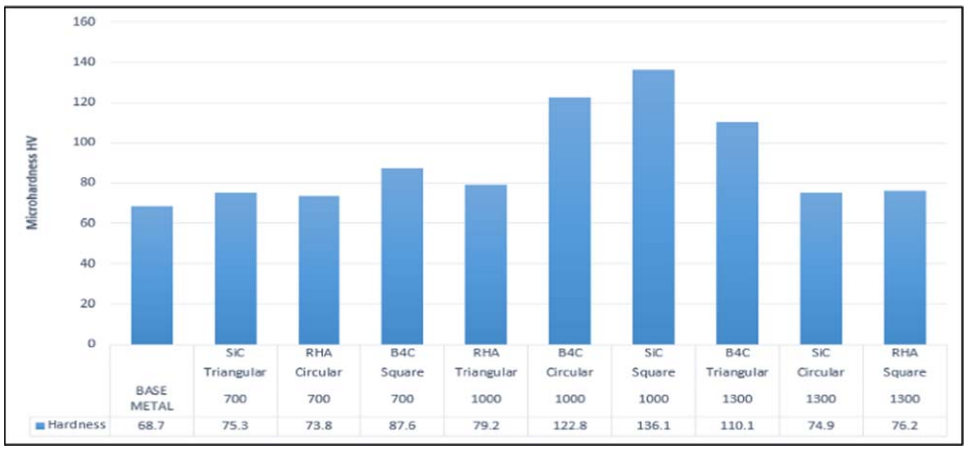


Figure 11. Bar graph of microhardness (HV value) of all specimens.

Table 7. ANOVA result for various responses.

FSP Parameter	SS	DOF	Variance	% Contribution	F-ratio
Microhardness Tool Rotation(rpm)	5122.0201	2	2561.01	35.1604	14.0425
Tool Profile	1301.5117	2	650.756	8.934 29	3.568 22
Reinforcement	4496.5639	2	2248.28	30.8669	12.3278
Error	3647.5065	20	182.375	25.0385	
Total	14 567.602	26	*	100	

needed to manage process inputs to optimize the output. DOE gives the optimized result of given experimental values or the outcome of the experiment performed [34, 35].

5. Conclusion

A detailed account of fabrication of AA7075 surface composite with different process parameters is presented. The results of the present investigation can be concluded as follows:

1. With the help of Taguchi technique, optimum condition was obtained to achieve the maximum value of microhardness. Increase in rotation speed from 700 rpm to 1000 rpm led to increase in HV value but further increase to 1300 rpm decreased the Hv value. Thus, 1000 rpm is the optimal choice.
2. Percentage contribution of the process parameters towards microhardness was achieved. Tool rotation speed is found to be the most influential factor followed by the type of reinforcement and tool profile.
3. Tool profile had much less significance on the Hv value compared to other parameters. Still square tool profile was found to be the optimal choice. This is attributed to the high pulsating action generated from square tool profile.
4. The average hardness value of AA7075/B_4C composite is found to be increased by 1.5–1.6 times of base metal, which was around 68 Hv. B_4C is found to be the most effective reinforcement compared to SiC and RHA in improving the microhardness value of the composite.
5. The average hardness value of AA7075/RHA composite was observed to be 76.4 which is about 11 % more than the base metal thus instead of dumping, RHA can be used to make aluminium composites with improved microhardness.

ORCID iDs

Ravi Butola https://orcid.org/0000-0002-6581-4131
Qasim Murtaza https://orcid.org/0000-0001-8784-2304

References

[1] Alaneme K K, Bodunrin M O and Awe A A 2018 Microstructure, mechanical and fracture properties of groundnut shell ash and silicon carbide dispersion strengthened aluminium matrix composites *Journal of King Saud University—Engineering Sciences* **30** 96–103

[2] Bodunrin M O, Alaneme K K and Chown L H 2015 Aluminium matrix hybrid composites: a review of reinforcement philosophies; mechanical, corrosion and tribological characteristics *Journal of Materials Research and Technology* **4** 434–45

[3] Butola R *et al* 2019 Experimental studies on mechanical properties of metal matrix composites reinforced with natural fibres Ashes *SAE* **2019-01-1123** 1–11 Technical Paper 2019-01-1123

[4] Surappa M K 2003 Aluminium matrix composites: challenges and opportunities *Sadhana* **28** 319–34 Parts 1 & 2

[5] Butola R, Shamrao D.M and Ranganath M.S 2018 Comparison Studies on Mechanical Properties of Hybrid Metal Matrix Composite *IJAPIE-2018-07-327* **3** 52–56

[6] Das D K, Mishra P C, Singh S and Pattanaik S 2014 Fabrication and heat treatment of ceramic—reinforced aluminium matrix composites—a review *International Journal of Mechanical and Materials Engineering* **1** 1–15

[7] Butola R, Singari R M, Bandhu A and Walia R S 2017 Characteristics and properties of different reinforcements in hybrid aluminium composites: a review *IJAPIE-SI-MM* **511** 71–80

[8] Joshi S, Singh R C and Chaudhary R 2019 Effect of rotational speed in friction stir processing on the microstructural and mechanical characteristics of cast AS21A magnesium alloy *Mater. Res. Express* **6** 056554

[9] Yadav D and Bauri R 2012 Effect of friction stir processing on microstructure and mechanical properties of aluminium *Materials Science and Engineering* A **539** 85–92

[10] Chaudhary A, Dev A K, Goel A, Butola R and Ranganath M S 2018 The mechanical properties of different alloys in friction stir processing: a review *Materials Today: Proceedings* **5** 5553–62

[11] Mahmoud T S 2008 'Effect of friction stir processing on electrical conductivity and corrosion resistance of AA6063-T6 Al alloy *Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci.* 1989-1996 **203-210** 222(7) 1117–23

[12] Rathee S, Maheshwari S, Siddiquee A N and Srivastava M 2018 A review of recent progress in solid state fabrication of composites and functionally graded systems via friction stir processing *Crit. Rev. Solid State Mater. Sci.* **43** 334–66

[13] Rathee S, Maheshwari S and Siddiquee A N 2018 Issues and strategies in composite fabrication via friction stir processing: a review *Mater. Manuf. Processes* **33** 239–261

[14] Rathee S, Maheshwari S, Siddiquee A N and Srivastava M 2018 Distribution of reinforcement particles in surface composite fabrication via friction stir processing: suitable strategy *Mater. Manuf. Processes* **33** 262–9

[15] Rathee S, Maheshwari S, Siddiquee A N and Srivastava I Investigating M Effects of groove dimensions on microstructure and mechanical properties of AA6063/SiC surface composites produced by friction stir processing *Trans. Indian Inst. Met.* (https://doi.org/10.1007/s12666-017-1060-7)

[16] Butola R, Singari R M and Murtaza Q 2019 A review on surface modification of aluminium alloy using friction stir processing' *International Journal for Research in Applied Science & Engineering Technology* **7** 2085–90

[17] Behnagh R A, Besharati Givi M K and Akbari M 2012 Mechanical properties, corrosion resistance, and microstructural changes during friction stir processing of 5083 aluminium rolled plates *Mater. Manuf. Processes* **27** 636–40

[18] Pasebani S, Charit I and Mishra R S 2015 Effect of tool rotation rate on constituent particles in a friction stir processed 2024Al alloy *Mater. Lett.* **160** 64–67

[19] Wang W, Shi Q, Liu P, Li H and Li T 2009 A novel way to produce bulk SiC_p reinforced aluminium metal matrix composites by friction stir processing *J. Mater. Process. Technol.* **209** 2099–103

[20] Shafiei-Zarghani A, Kashani-Bozorg S F and Zarei-Hanzaki A 2009 Microstructures and mechanical properties of Al/Al2O3 surface nano-composite layer produced by friction stir processing *Mater. Sci. Eng.* A (https://doi.org/10.1016/j.msea.2008.09.064)

[21] Srivastava M, Rathee S, Maheshwari S and Siddique A N 2019 Optimisation of friction stir processing parameters to fabricate AA6063/SiC surface composites using Taguchi technique *Int. J. Mater. Prod. Technol.* **58** 16–31

[22] Rathee S, Maheshwari S, Siddiquee A N and Srivastava M 2016 Process parameters optimization for enhanced microhardness of AA 6061/ SiC surface composites fabricated via friction stir processing (FSP) *Materials Today Proceedings* **3** 4151–4156 PART B

[23] Srivastava M, Rathee S, Maheshwari S and Siddiquee A N 2018 Investigating effect of multiple passes on microstructural and mechanical properties of Al-Mg/SiC surface composites using under water FSP *Material Research Express, IOP Science* **5** 066511

[24] Srivastava M, Rathee S, Maheshwari S and Siddiquee A N 2018 Investigation on underwater FSP of Al-Mg-Si alloy surface composites *Material Research Express, IOP Science* (https://doi.org/10.1088/2053-1591/aaebe9)

[25] Shojaeefard M H, Akbari M, Asadi P and Khalkhali A 2017 The effect of reinforcement type on the microstructure, mechanical properties, and wear resistance of A356 matrix composites produced by FSP *Int. J. Adv. Manuf. Technol.* **91** 1391–407

[26] Dakarapu S R and Nallu R 2017 Process parameters optimizations for producing Aa6061/Tib2 composites by friction stir processing *Journal of Mechanical Engineering—Strojnícky časopis* **67** 101–18

[27] Alaneme K K and Sanusi K O 2015 Microstructural characteristics, mechanical and wear behavior of aluminium matrix hybrid composites reinforced with alumina, rice husk ash and graphite *Engineering Science and Technology, an International Journal* **18** 416–22

[28] Butola R, Pratap C, Shukla A and Walia R S 2019 Effect on the mechanical properties of aluminum-based hybrid metal matrix composite using stir casting method *Mater. Sci. Forum* **969** 253–9

[29] Asadi P, Givi M K B, Rastgoo A, Akbari M, Zakeri V and Rasouli S 2012 Prediciting the grain size and hardness of AZ91/SiC nanocomposite by artificial neural networks *Int. Journal Adv Manuf Technol* **63** 1095–107

[30] Dolatkhah A, Golbabaei P, Besharati Givi M K and Molaiekiya F 2012 Investigating effects of process parameters on microstructural and mechanical properties of Al5052/SiC metal matrix composite fabricated via friction stir processing *Mater. Des.* **37** 458–64

[31] Shamsipur A, Kashani-Bozorg S F and Zarei-Hanzaki A 2011 The effects of friction-stir process parameters on the fabrication of Ti/SiC nano-composite surface layer *Surf. Coat. Technol.* **206** 1372–81

[32] Ahmadkhaniha D, Heydarzadeh Sohi M, Zarei-Hanzaki A, Bayazid S M and Saba M 2015 Taguchi optimization of process parameters in friction stir processing of pure Mg *Journal of Magnesium and Alloys* **3** 168–72

[33] Sharma V and Butola R 2017 Optimization of machining parameters in CNC turning of hybrid metal matrix composites using different techniques: a review *International Journal of Advance Research and Innovation* **5** 78–82

[34] Butola R, Jain R, Bhangadia P, Bandhu A, Walia RS and Murtaza Q 2018 'Optimization to the parameters of abrasive flow machining by Taguchi method' *Materials Today: Proceedings* **5** 4720–9

[35] Butola R, Murtaza Q, Walia R S and Kumar P 2017 Two start and three start helical abrasive flow machining for brittle materials *Materials Today: Proceedings* **4** 3685–93

Materials Research Express

PAPER

Graphene as charge transport layers in lead free perovskite solar cell

To cite this article: Gagandeep *et al* 2019 *Mater. Res. Express* **6** 115611

View the article online for updates and enhancements.

IOP ebooks™

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

This content was downloaded from IP address 14.139.251.100 on 24/10/2019 at 12:48

Materials Research Express

PAPER

Graphene as charge transport layers in lead free perovskite solar cell

RECEIVED
27 August 2019

REVISED
18 September 2019

ACCEPTED FOR PUBLICATION
4 October 2019

PUBLISHED
16 October 2019

Gagandeep[1,4], **Mukhtiyar Singh**[2], **Ramesh Kumar**[1,4] **and Vinamrita Singh**[3]

[1] Department of Physics, Guru Jambheshwar University of Science & Technology, Hisar, India
[2] Department of Applied Physics, Delhi Technological University, Delhi, India
[3] Department of Applied Sciences and Humanities, Ambedkar Institute of Advanced Communication Technologies & Research, Delhi, India
[4] Authors to whom any correspondence should be addressed.

E-mail: rameshkumar.bibiyan@gmail.com **and** gagankundu22@gmail.com

Keywords: AFORS-HET software, perovskite solar cell, graphene, efficiency, MATLAB

Abstract

Tin-based perovskite material is one of the promising candidates used in lead-free perovskite solar cells. In the present study, the use of graphene material as charge transport layers is suggested and the structure n-Graphene/$CH_3NH_3SnI_3$/p-Graphene is simulated in AFORS-HET (Automat for Simulation of Heterostructures) software. The effect of layer properties on the performance of the proposed structure is analyzed. After optimizing the layer parameters, the best possible structure gives a $V_{oc} = 687.1$ mV, $J_{sc} = 19.93$ mA cm^{-2}, FF = 77.92% and efficiency = 10.67%. The textured nature of front surface enhances the efficiency of the cell from 10.67% to 13.28%. The effect of operating temperature is observed on the performance of the cell. The proposed structure is also modelled using one diode model in MATLAB software for comparison with AFORS-HET results. The results of the simulated and modelled structure are in good agreement with each other. To the best of our knowledge, the proposed structure and the corresponding results reported herein are new for the research community.

1. Introduction

Perovskite solar cells (PSCs) based on $CH_3NH_3PbX_3$ material as the absorber layer are becoming the most studied photovoltaic technology as compared to other solar cells due to the broad absorption spectra, fast exciton separation, band gap tuning, and ambipolar charge transport nature of the material [1–4]. The chemical formula for perovskite compounds is ABX_3 (here A is an organic cation, B is a divalent metal ion, and X is a halide). The PSCs are manufactured by simple solution processing [5–7] or scalable vapor phase deposition method [8, 9]. The power conversion efficiency (PCE) of $CH_3NH_3PbX_3$ based PSC has been significantly improved from 3.8% [9] to 22.1% [10].

However, the toxicity of lead in $CH_3NH_3PbX_3$ material is a major concern. The lead material can be replaced by another same group element such as Sn or Ge. The band gap of $CH_3NH_3SnI_3$ is 1.3 eV as confirmed by theoretical and experimental studies [11, 12]. This amount of energy band gap can cover the wide range of visible spectrum as compared to the band gap of lead-based PSC (1.5 eV). The lead-free, tin-based PSC (TPSC) has been fabricated using $CH_3NH_3SnI_3$ with planar heterojunction architecture [13–15]. Although, the major problem with the TPSC is its chemical instability due to oxidation of Sn^{+2} to Sn^{+4} states, the advancement in fabrication techniques has improved the stability of these cells. The incorporation of SnF_2 in the system decreases the Sn^{+4} caused by the oxidation of Sn^{+2} [16]. Therefore, TPSCs provide a good platform for developing non-toxic PSCs [16, 17]. Noel *et al* [18] studied TPSC on a mesoporous TiO_2 scaffold, and obtained an efficiency of 6% under 1 Sun illumination. Koh *et al* [19] analyzed the formamidinium tin based perovskite ($FASnI_3$) material as absorbing layer in PSC with 1.41 eV energy gap and a PCE of 2.10% was obtained with the incorporation of SnF_2. Lee *et al* [20] fabricated the $FASnI_3$ PSC with high reproducibility and achieved a PCE of 4.8%. The cell presented a stable performance for 100 days and retained 98% of its starting efficiency. Shao *et al* [21] studied PSC with planar p-i-n device structure for $FASnI_3$ and achieved a PCE of 9.0%. Hao *et al* [22] reported the TPSC with

© 2019 IOP Publishing Ltd

Table 1. Summary of solar cell parameters with different structures based on graphene as one of the transporting layers or as an additive.

Structure	Voc (mV)	Jsc (mA cm^{-2})	FF (%)	PCE (%)	References
FTO/TiO_2/$CH_3NH_3PbI_3$/graphene/Au	710	21.3	47	7.1	[42]
Glass/FTO/ZnO/graphene-doped ZnO/$MAPbI_3$/Spiro- OMeTAD/Ag	926	19.97	56.31	10.34	[43]
FTO-glass/TiO_2:graphene/Al_2O_3:perovskite-layer/ /Spiro-OMeTAD/Au	1040	21.9	73	15.6	[44]
Glass+FTO/cTiO_2/graphene+mTiO_2/Perovskite/spiro-OMeTAD/Au	857	18.16	64.6	12.6	[45]

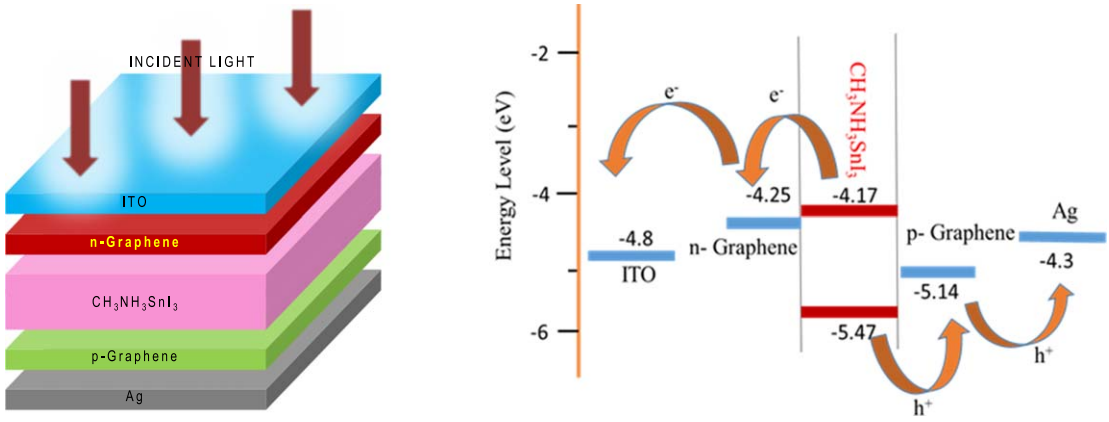


Figure 1. (a) Simulated structure and (b) energy band diagram of tin perovskite solar cell.

spiro-OMeTAD as organic hole transporting layer, giving the absorption at 950 nm, which is remarkably red shifted compared with the lead-based PSC. The band gap was tuned to overcast much of the visible spectra, thus enabling a PCE of 5.73% under simulated sunlight. Recently, Jokar *et al* [23] studied the TPSC with hybrid organic cation and attained 8.5% efficiency.

Till now, the frequently used electron transporting layer (ETL) and hole transporting layer (HTL) for PSCs are TiO_2 and spiro-OMeTAD respectively, but nowadays, to improve the charge collection and injection at the electrode; graphene is being used as HTL/ETL to enhance the PCE of the solar cell. In particular, when charge carriers are transported to the ETL or HTL from the perovskite layer, the non-radiative recombination is reduced and degradation of the device is slowed down. Currently, it is observed that the graphene material and its derivatives such as graphene oxide and graphene quantum dots are introduced as device layers to boost the performance of PSCs [4, 24–26]. This is due to the extraordinary properties of graphene such as high value of Young modulus (1.0 TPa), charge carrier mobility (2×10^5 cm^2 V^{-1} s^{-1}), optical transmittance (97.7%) and thermal conductivity (5000 Wm^{-1} K^{-1}) [27, 28]. The electronic and optical properties of graphene can be adjusted by doping a heteroatom because doping is the most suitable method to control the semiconducting properties. The boron and nitrogen atoms are the natural candidates for graphene doping because of their similar size and form [29–33]. The boron doping, regarded as hole doping, in graphene leads to the shifting of Fermi level below the Dirac point and forms p-graphene [34–36]. The nitrogen atom has one extra electron than carbon atom, and behaves as electron doping, i.e., the Fermi level shifts above the Dirac point and forms n-graphene [34–36]. The change in the Fermi level of the graphene accounts for the change in the work function of the graphene. The energy band gap can be increased from 0.13 eV to 0.54 eV for one boron atom or one nitrogen atom in 5×5 (concentration 2%) and 2×2 (concentration 12%) supercells respectively [37]. Nevis *et al* [38] experimentally confirmed the energy gap of graphene to be 0.5 eV.

The first observed Schottky junction of graphene/silicon solar cell was constructed in 2010, obtaining an efficiency of 1.65% [39]. Song *et al* reported a good hike in the efficiency of graphene/Si Schottky junction solar cell and obtained up to 15% within less than a decade [40], but Patel *et al* simulated and optimized the structure of n-graphene/p-silicon and p-graphene/n-silicon heterostructures solar cell and achieved an efficiency of 11.38% and 9.81% for 40 layers of n-graphene and 20 layers of p-graphene, respectively [41]. Sho *et al* [42] replaced the hole transporting layer, spiro-OMeTAD, by graphene layer in a PSC and obtained a PCE of 7.1% with J_{sc} of 21.3 mA cm^{-2}. Their experimental studies also indicate that the performance of the cell can be improved by modifying the work function of graphene layer and suggested doping as a mean to vary the work function of the graphene. Graphene has also been added into the ZnO electron selective contact layer, which resulted into the increase in efficiency of PSC by a factor of 47.5%, i.e., from $7.01 \pm 0.66\%$ to $10.34 \pm 0.18\%$ [43]. Wang *et al* [44] found that TiO_2 nanoparticles and graphene nanoflakes behave as an electron selective electrode at a temperature of 150 °C, i.e., the solar cell with graphene-TiO_2 nanocomposite layer yielded an efficiency of 15.6%. Agresti *et al* [45] also achieved large area PSC by dispersing the graphene flakes in mesoscopic TiO_2 and obtained an efficiency of 12.6%. The performance of PSCs utilizing graphene as one of the transporting layers or as an additive are summarized in table 1.

The above theoretical and experimental studies indicate that graphene material can act as ETL and HTL in PSC. Till date, there is no publication reported in literature for the use of graphene in TPSC. Keeping the above

literature studies and fundamental concepts in mind, the structure of TPSC with n-graphene as ETL and p-graphene as HTL was simulated as shown in figure 1. The layer parameters were optimized for obtaining the best performance of the device. The effect of operating temperature and textured front surface on the device performance is investigated. The studies demonstrated that graphene can act as an efficient charge transporting layer in TPSC and promising devices may be realized with high performance. The simulated structure was also modelled using one diode model in MATLAB software and the current density–voltage (J-V) characteristics and power-voltage (P-V) characteristics are analyzed for best optimized results. There is a considerable agreement between the simulated and modelled results. The use of doped graphene as charge transport layers in TPSC is new in this field.

2. Simulating structure

Many theoretical and experimental researches are available on the use of graphene and its derivative as efficient transport layers in PSCs [4, 26, 42–45]. Till now in literature, no research papers are published on design and simulation of the TPSC structure with graphene as charge transport materials. We designed and simulated the structure as n-Graphene /$CH_3NH_3SnI_3$/p-Graphene which is a n-i-p structure, having ITO (indium doped tin oxide) as front contact and Ag (silver) as back contact in AFORS-HET software [46–48] and the simulation is carried out under the standard AM 1.5 G spectrum. This software is designed based on the 1D semiconductor equations linked to Shockley Reed Hall statistics solver and Lambert-Beer law for optical parameters of different photovoltaic devices. The complete device structure is ITO (80 nm)/n-Graphene (0.334–13.36 nm)/Sn-Perovskite (50–700 nm)/p-Graphene (0.334 nm)/Ag (0.001 cm) as shown in figure 1(a).

The work function of graphene can be tuned depending upon the doping of heteroatoms such as Boron and Nitrogen atoms. These atoms are more suitable for graphene doping. The work function of pure graphene is 4.6 eV as verified by theoretical and experimental studies [35, 49]. The doping of nitrogen atom in graphene results into n-graphene, because it contains one extra electron that is donated to graphene. The work function of graphene increases, i.e., shifts up to 4.25 eV [50]. The boron doping in graphene results into p-graphene because of the deficiency of one electron. The work function of graphene decreases up to 5.14 eV depending upon the number of dopant atoms [35].

In the present manuscript, we considered the work function of n-graphene and p-graphene as 4.25 eV and 5.14 eV respectively. The energy band diagram for the simulated structure is given in figure 1(b). The working of the device can be explained based on this diagram. As light falls on the device from n-graphene side, most of it reaches the active layer due to the high transmittance of graphene, and generates electron-hole pairs. As seen from the energy level diagram, it is favorable for the electrons to jump from n-graphene to ITO. The down shift in energy level of 4.6 eV to 4.25 eV provides a gradual transition to the collecting electrode. Similarly, the up-shift of work function from 4.6 eV to 5.14 eV assists in hole transfer from active layer to Ag electrode by removing the barrier between graphene and Ag, and acts as an electron blocker as well. The transfer of both the charge carriers from their interface region takes place rapidly due to high charge carrier mobility of graphene transport layers. The recombination of charge carries at the interface regions reduces and thus, the device performance improves.

Further, to enhance the device performance, we optimized the device layer parameters like, energy band gap (E_g), layer thickness, electron affinity, dielectric constant, doping concentrations (N_D/N_A), effective conduction/valence band density (N_C/N_V) to observe the effect on the efficiency, fill factor (FF), open circuit voltage (V_{oc}), and current density (J_{sc}) of the TPSC. The realistic parameters that we have used in our simulating work are taken from various experimental and theoretical studies [18, 22, 23, 28–32, 41, 51]. To analyze the front surface influence on the device performance, the front surface can be considered as textured surface. The textured surface helps in light trapping and reduces the light reflection, due to which, the performance of device may increase.

3. Results and discussion

3.1. n-graphene parameter optimization

As we know that Fermi level of graphene is always either in valence band or in conduction band due to gapless nature of graphene. After nitrogen doping, the Fermi level as well as the work function of graphene are shifted. The graphene behaves as n-graphene with some energy gap introduced after doping. It has been experimentally shown that the incorporation of n-doped graphene in PSC enhances the performance of the cell and results in improved morphology and surface passivation [52]. On the basis of the parameters given in table 2, the values of performance parameters calculated for the solar cell are: efficiency = 10.65%, J_{sc} = 19.92 mA cm^{-2},

Table 2. Summary of the optimized parameters used in simulation.

Input parameter	n-Graphene	Sn - perovskite	p-Graphene
Layer thickness (nm)	0.334	700	0.334
Dielectric constant	10	6.5	7.1
Electron affinity (eV)	4.8	4.17	4.8
Band gap (eV)	0.5	1.3	0.5
Optical band gap (eV)	0.5	1.3	0.5
Valence band density (cm^{-3})	3×10^{9}	1×10^{18}	3×10^{19}
Conduction band density (cm^{-3})	3×10^{9}	1×10^{19}	3×10^{19}
Hole mobility ($cm^2 V^{-1} s^{-1}$)	10	1.6	10
Electron mobility ($cm^2 V^{-1} s^{-1}$)	1×10^{5}	1.6	1×10^{5}
Thermal velocity of electron($cm\ s^{-1}$)	1×10^{7}	1×10^{6}	1×10^{7}
Thermal velocity of hole($cm\ s^{-1}$)	1×10^{7}	1×10^{6}	1×10^{7}
Doping concentration(cm^{-3})	1×10^{26}	3.2×10^{15}	1×10^{18}
Layer density ($g\ cm^{-3}$)	2.328	2.328	2.328

$V_{oc} = 691.5$ mV and FF = 77.28%. To analyse how graphene influences the device working, the optimization of various parameters is carried out keeping other layer parameters at fixed values. The value of dielectric constant for n-graphene is in the range 10–20 [41]. As we varied the dielectric constant from 10–20, the device performance remains constant and the simulated results are shown in figure 2(a). The value of dielectric constant is therefore taken 10 for n-graphene. The value of energy band gap for n-graphene is varied from 0–0.5 eV, depending upon the doping concentration. The energy band gap (E_g) is optimized in the range of 0–0.5 eV for a single layer graphene. As the E_g varies from 0–0.5 eV, the efficiency, V_{oc}, J_{sc} and FF are increased up to 0.25 eV and remain constant beyond the value 0.25 eV. This is due to the difference in the work function of graphene and $CH_3NH_3SnI_3$ layers which creates a built-in voltage at the interface of the two layers. The V_{oc} is increased from 689.8 mV to 691.5 mV as the built-in voltage increases. The built-in voltage increases the barrier height, which separates the photo generated carriers. The J_{sc} is related to barrier voltage and hence increases from 19.17 mA cm^{-2} to 19.92 mA cm^{-2}. Therefore, the efficiency is increased from 10.21% to 10.65%. The FF slightly varies from 77.22% to 77.28%. The complete behaviour of performance parameters with E_g are presented in figure 2(b). After dielectric constant and energy band gap optimisation, the parameter N_C/N_V is varied in the range of 3×10^{9} cm^{-3} to 3×10^{22} cm^{-3}. On increasing the value of N_C/N_V of n-graphene layer, the recombination rate of photo-generated carrier is higher near the junction which probably causes the reduction in minority carrier densities. The Fermi level of n-graphene was pinned in the absence of free minority charge carriers; this makes barrier height independent of work function of graphene and trimmed down the collection of free charge carriers. Therefore, J_{sc} is reduced from 19.92 mA cm^{-2} to 17.59 mA cm^{-2} and V_{oc} sharply falls from 691.5 mV to 514.9 mV due to trapping of charge carriers. Thus, the reduced values of V_{oc} and J_{sc} correspond to drop in efficiency to 4.72% at 3×10^{22} cm^{-3} of N_C/N_V value but highest efficiency is obtained at 3×10^{9} cm^{-3} of N_C/N_V having value 10.65% and the results are shown in figure 2(c).

The donor concentration (N_D) i.e. doping of nitrogen atom in n-graphene layer is optimized, keeping other parameters at fixed value. The nitrogen atom doping reduces the work function of graphene and increases the Fermi level of graphene from the Dirac point and results in a significant increase in carrier density. The conductivity of graphene layer is increased with increase in donor concentration (N_D). As N_D is varied from 1×10^{8} cm^{-3} to 1×10^{22} cm^{-3}, V_{oc} rises from 222.6 mV to 691.5 mV, FF increases from 38.76% to 77.20% and the efficiency rises from 1.38% to 10.63%. If the donor concentration is further increased, only small changes are observed in the performance parameters of TPSC. It has been reported that both the sheet resistance and series resistance reduce due to increase in donor concentration, therefore, an increase in efficiency is observed. At $N_D = 1 \times 10^{26}$ cm^{-3}, the efficiency slightly changes and is found to be 10.65%. The variation of TPSC performance parameters with N_D are shown in figure 2(d). After this, we optimized one of the important parameters that relates with the cost of solar cell, i.e., layer thickness of n-graphene. To further enhance the performance of the device, we analysed the effect of multilayer graphene. We simulated the cell up to 40 layers of n-graphene by taking the thickness of single layer graphene as 0.334 nm, i.e., the thickness of n-graphene layer is varied from 0.334 nm to 13.36 nm. The transmittance and sheet resistance of graphene is controlled by the thickness of graphene. As we increase the n-graphene layer, the transmittance as well as sheet resistance of the graphene decreases. Both these effects are in opposition to each other. On one hand, the decrease in transmittance reduces the light reaching the active layer and consequently, a lower current is obtained. On the other hand, a reduced sheet resistance should aide in charge collection [53]. However, it is observed in the present case that the cell performance decreases with increase in thickness implying that the transmittance of

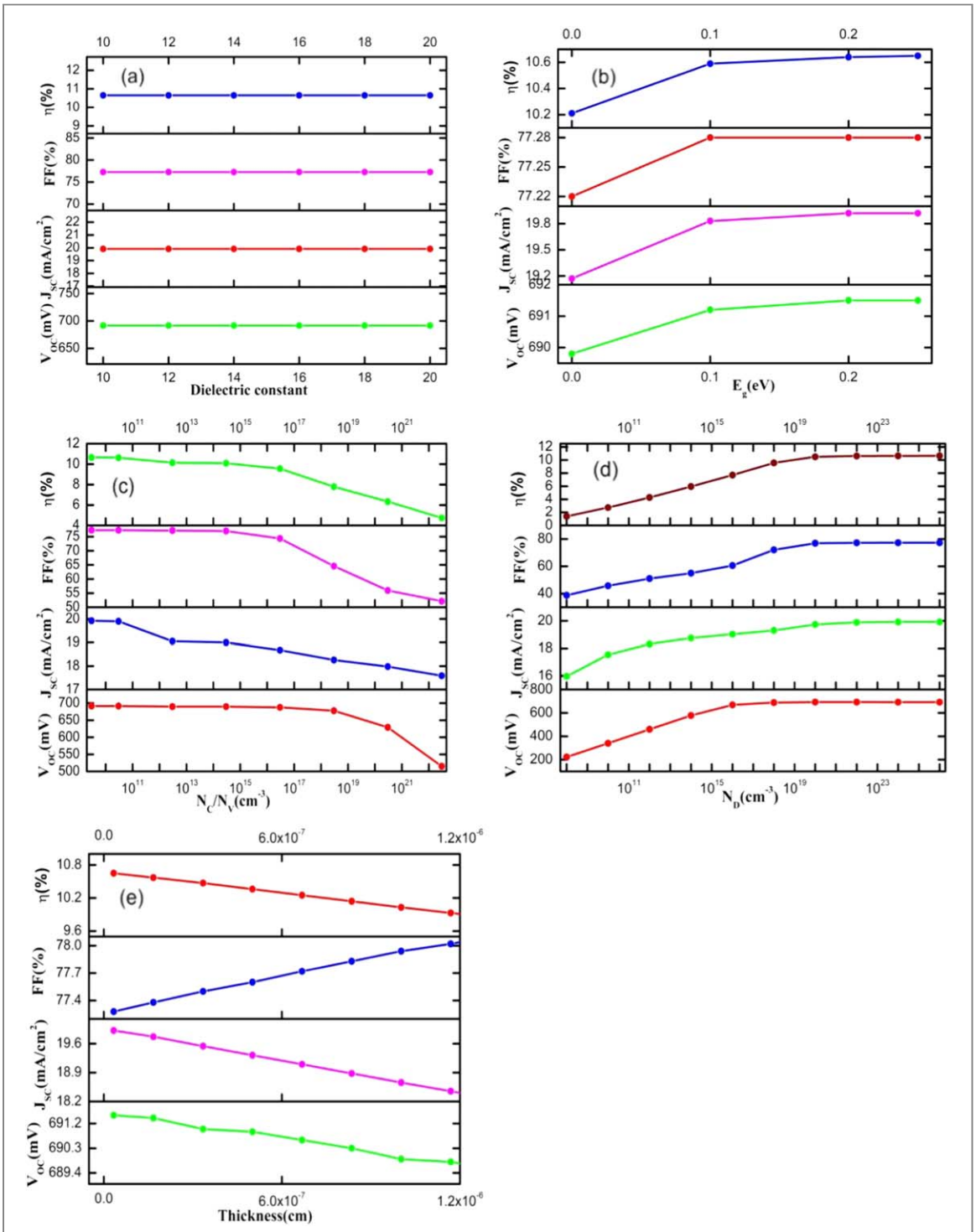


Figure 2. TPSC performance with the n- graphene parameters (a) dielectric constant, (b)energy gap E_g (eV), (c) N_C/N_V (cm^{-3}), (d) N_D (cm^{-3}), (e) thickness, keeping other layer parameters constant.

graphene dominates in effect. This may be supported by the findings that each additional layer of graphene reduces the transmission percentage by ~3% [54]. The detailed variations in cell parameters with n-graphene thickness are given in figure 2(e).

3.2. Active layer parameter optimization

Furthermore, to incorporate the impact of active layer parameters on the cell performance, we optimised the active layer parameters by keeping n-graphene layer parameter at optimized values as mentioned in table 2. Firstly, the thickness of the active layer was varied in the range from 50–700 nm. The efficiency, J_{sc} and V_{oc},

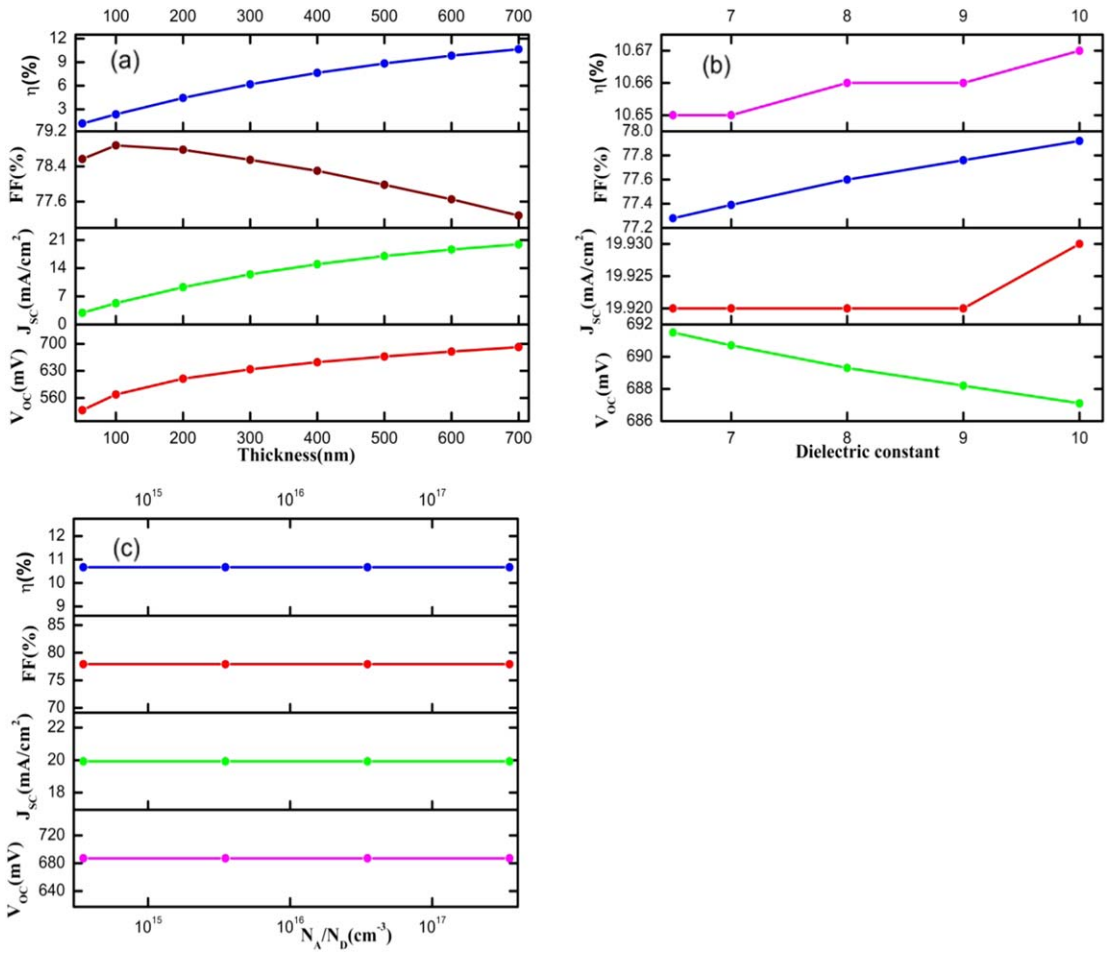


Figure 3. TPSC performance with active layer parameters optimization (a) thickness (nm) (b) Dielectric constant and (c) N_A/N_D (cm^{-3}). The optimized parameters of n- Graphene layer are kept constant.

drastically increased from 1.19% to 10.65%, 2.28 mA cm^{-2} to 19.92 mA cm^{-2} and 528.4 mV to 691.5 mV, respectively, but the FF decreased from 78.57% to 77.28% with the active layer thickness. The overall performance of solar cell parameters is mentioned in figure 3(a). The concept behind this increase in performance of device is that the absorption of light is increased with the thickness of the active layer, leading to higher carrier generation. The increase in efficiency tends to saturate around 700 nm. The effect of increased carrier generation is counter-balanced by the increase in carrier recombination since charge carriers are not able to transport through the active layer towards the electrodes. The charge diffusion length is an important factor in deciding the perovskite layer thickness. The exciton diffusion length of TPSC is found to be more than 500 nm [55, 56], which is consistent with the results obtained. Konstantakou *et al* [56] experimentally obtained the efficiency of 1.7% for $CH_3NH_3SnI_3$ absorber layer having thickness 200 nm. Noel *et al* also studied the lead-free perovskite layer of thickness 300 nm and obtained the efficiency around 6% [18], which in our case is 6.19%. Therefore, to check the influence of thickness of the absorber layer beyond 300 nm, we varied the thickness up to 700 nm and the efficiencies obtained are higher as compared to the literature studies [18, 22, 23, 56].

After the optimization of thickness, the dielectric constant of the active layer was varied in the range of 10–13.6. As we increase the dielectric constant, the J_{sc}, V_{oc}, FF and efficiency change from 19.92 mA cm^{-2} to 19.93 mA cm^{-2}, 691.5 mV to 687.1 mV, 77.28% to 77.92% and 10.65% to 10.67% respectively, and the results are shown in figure 3(b). Thereafter, the value of N_A/N_D is varied from 3.5×10^{14} cm^{-3} to 3.5×10^{17} cm^{-3} to observe its effect on the performance parameters. The variation in performance parameters with acceptor and donor densities of active layer are mentioned in figure 3(c). From the figure, we observed that the dopant concentration i.e. N_A/N_D does not affect the performance parameters.

3.3. Optimization of p-graphene layer

After optimization of ETL and active layer parameters, we optimised the HTL, i.e., p-graphene layer parameters. The acceptor concentration (N_A) that makes the graphene act as p-graphene is varied from 1×10^{14} to

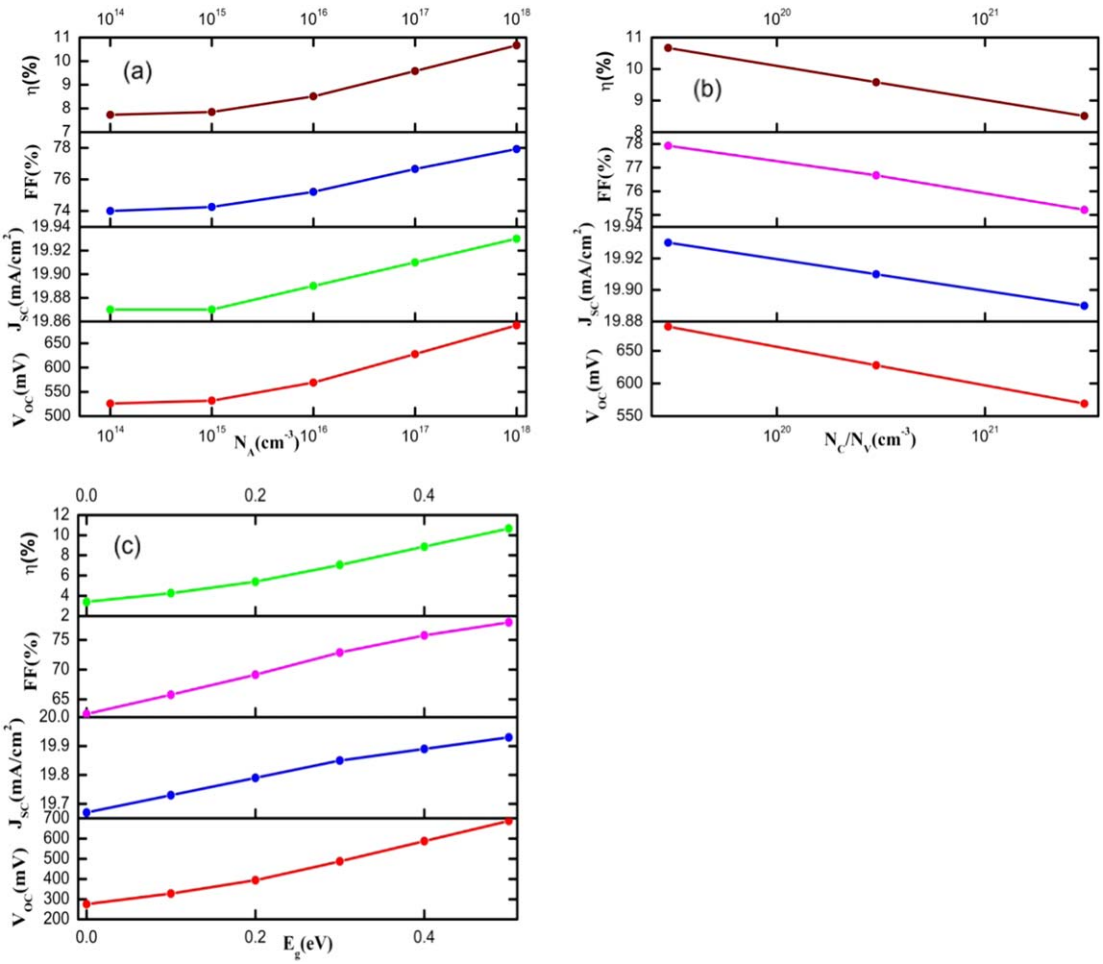


Figure 4. TPSC performance with the optimization of p-graphene parameters, (a) N_A (cm^{-3}), (b) N_C/N_V (cm^{-3}), (c) E_g (eV) and (d) keeping optimized parameter of n- graphene and active layer constant.

1×10^{18} cm^{-3}, to observe its effect on the performance parameters. The efficiency is increased from 7.73% to 10.67% with the increase in N_A. This might be due to rise in the work function of the graphene. The holes are rapidly injected by the p-graphene electrode. The sheet resistance of graphene layer is also reduced [53]. Referring to figure 1(b), p-doping of graphene leads to better alignment of energy levels between HOMO (highest occupied molecular orbital) of perovskite and Ag electrode. The value of J_{sc} is slightly changed from 19.87 mA cm^{-2} to 19.93 mA cm^{-2}. The $V_{oc,}$ and FF of the TPSC also increased from 525.9 mV to 687.1 mV and 74% to 77.92% with N_A. The detailed variations are presented in figure 4(a).

The values of N_C/N_V are changed from 3×10^{19} cm^{-3} to 3×10^{21} cm^{-3} to analyse its effect on the performance of TPSC. As from the figure 4(b), we noticed that with the increment in N_C/N_V value, the V_{oc} is reduced from value 687.1 mV to 569.1 mV, due to increase in minority carrier density. The efficiency and FF are reduced from 10.67% to 8.51% and from 77.9% to 75.21% respectively. The J_{sc} is decreased from 19.93 mA cm^{-2} to 19.89 mA cm^{-2}. This happens due to short lifetime of photo-generated charge carrier and trapping of free charge carrier because of increase in minority charge carrier density.

The effect of energy band gap of p-graphene is seen on the performance of TPSC. The energy band gap is varied from 0–0.5 eV. The variation of TPSC performance parameters with the E_g are presented in figure 4(c). From the graph we observed that the efficiency, FF, J_{sc} and V_{oc} varies from 3.38% to 10.67%, 62.53% to 77.92%, 19.67 mA cm^{-2} to 19.93 mA cm^{-2} and 275.6 mV to 687.1 mV respectively. This is due to increase in built-in voltage at the perovskite/p-graphene interface which might due to the difference in the work function at the interface. This voltage separates the electron-hole pair at the interface and holes are collected by p-graphene layer. The thickness of p-graphene layer does not influence the performance of simulated structure further. Therefore, the best optimized value of performance parameters corresponding to each layer are given as $J_{sc} = 19.93$ mA cm^{-2}, $V_{oc} = 687.1$ mV, FF = 77.92% and the efficiency is 10.67% respectively. The J-V characteristics and quantum efficiency with wavelength of the best simulated cell is given in figures 5(a) and (b) respectively.

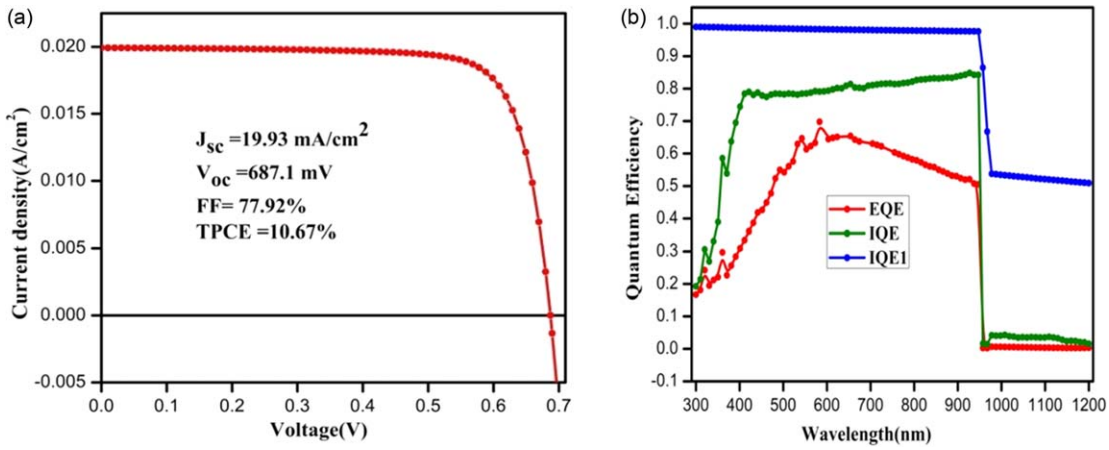


Figure 5. (a) J-V curve and (b) quantum efficiency with wavelength for best optimized parameters of the solar cell.

3.4. Operating temperature influence on the cell performance

The impact of operating temperature on the device performance is observed in the temperature range of 300–900 K. The efficiency, FF, J_{sc} and V_{oc} are significantly affected with the rise in temperature. The efficiency of the device and FF of the device are maximum i.e. 10.67% and 77.92% respectively at temperature 300 K. The efficiency and FF vary from 10.67% to 0.06%, and 77.92% to 21.49% respectively with the rise in temperature. The variation in V_{oc} with temperature is given by the following formula [57]:

$$\beta_{Voc} = \frac{10^6}{V_{oc}(25\ ^\circ\mathrm{C})} - \frac{V_{oc}(Tc - V_{oc}(25\ ^\circ\mathrm{C})}{T_C - 25\ ^\circ\mathrm{C}}$$

Here β_{Voc} is the temperature coefficient in parts per million (ppm)/°C and T_C is the temperature of the solar cell.

The V_{oc} is declining from 687.1 mV to 15.14 mV with the rise in temperature, because of increase in diode saturation current with the temperature. Recombination rate is also increased due to change in intrinsic carrier concentration which is mainly affected by the diode saturation current. The J_{sc} also decreased from 19.93 mA cm^{-2} to 18.62 mA cm^{-2} with temperature because of reduction in ideal current due to increase in energy band gap, whereas the collection fraction enhances typically with temperature. Hence, depending upon which of these conditions dominate, J_{sc} might either lower or higher with temperature. But overall report indicates that for higher temperature J_{sc} gets reduced. The detailed performance of TPSC parameters with temperature are presented in figure 6.

3.5. Effect of textured surface on the TPSC performance parameters

The nature of front contact surface also affects the performance of solar cell i.e. it may be plane or textured. The textured surface is preferred for increasing the efficiency of TPSC. The textured surface absorbed more photons from Sun light as compared to plane surface and hence reduction in light reflection from surface. This is due to the fact that when incident light falls on the surface of the cell, a certain amount of light is bounced back, but if the surface is textured then the light falls on one side of pyramid and bounces to the side of another pyramid. Hence more area is utilized to absorb the incident light and efficiency can be improved due to multiple reflections.

As shown in the figure 7, the texture angle is varied from 0° to 80° for textured surface and the performance of TPSC parameters is observed. The efficiency is increased from 10.67% for plane front surface to 13.28% for front textured surface. It means that solar cell with textured surface has higher photo-absorption rate. The J_{sc} is enhanced from a value of 19.93 mA cm^{-2} to 24.03 mA cm^{-2}.

3.6. Modelling of TPSC

We modelled the best optimised TPSC structure as one diode model in MATLAB software. The TPSC as diode can be represented through an equivalent circuit as shown in figure 8.

The value of material constant is varied from 2.6 to 4.6 for perovskite solar cell [58, 59]. The values of series resistance R_s (6.89 Ω) and shunt resistance R_p (2910.53 Ω) are obtained from the J-V characteristics of TPSC given in figure 5(a). The values of efficiency and FF are obtained as 10.73% and 78.42% respectively, from diode model using numerical programming in MATLAB. In this work, the value of material constant a = 3.5 gave

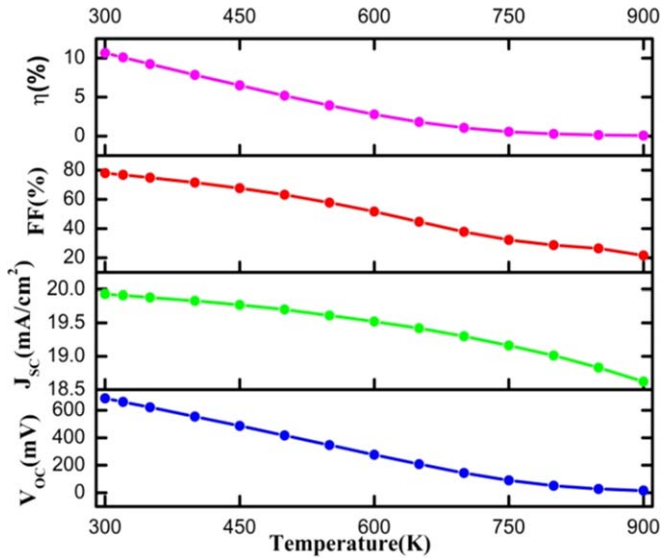


Figure 6. TPSC parameters variations with operating temperature.

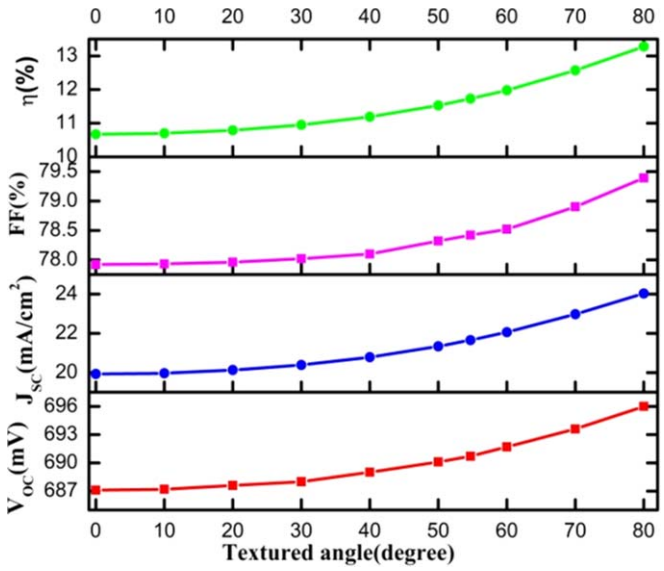


Figure 7. TPSC parameters variations with textured front surface.

better agreement with AFORS-HET simulation results. The J-V and P-V characteristics of TPSC are given in figures 9(a), (b) respectively. The above results reveal that simulation results are in good agreement with the modelling results for plane front surface.

4. Conclusion

In the present theoretical work, we studied the graphene-based charge transport layers in the tin based perovskite solar cell using AFORS-HET software. After optimizing the layer parameters, a maximum efficiency of 10.67% is obtained for plane front surface. The textured front surface increases the efficiency of the device up to 13.28%. The effect of operating temperature has been observed on the performance of the solar cell. We also presented a diode model of tin based perovskite solar cell in MATLAB software. The results of the simulation and modelling are in agreement with each other. The performance and optimization of the device produced here are useful for commercialization of lead-free perovskite solar cell.

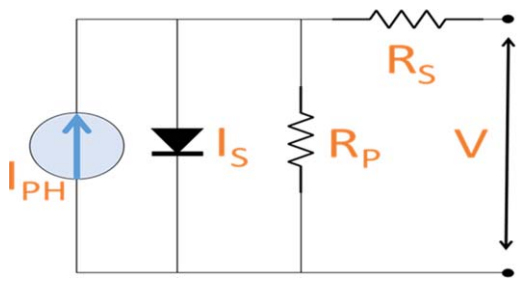


Figure 8. Equivalent circuit for tin perovskite solar cell.

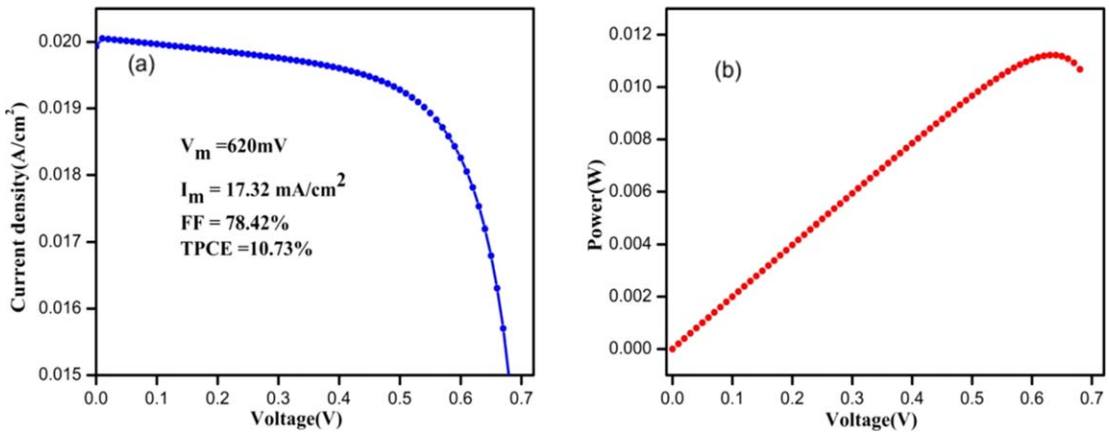


Figure 9. (a) J-V curve and (b) P-V curve obtained by modelling of TPSC using one diode model.

Acknowledgments

One of the authors (Ramesh Kumar) is grateful to the University Grants Commission, INDIA for providing the financial support through START-UP GRANT (No. F.30-364/2017 (BSR), Dated 15/12/2017). The authors gratefully acknowledge the AFORS-HET software developer for providing free access of the software.

ORCID iDs

Ramesh Kumar https://orcid.org/0000-0002-1026-6764

References

[1] Gao P, Gratzel M and Nazeeruddin M K 2014 Organohalide lead perovskites for Photovoltaic applications *Energy Environ. Sci.* **7** 2442–63
[2] Green M A, Ho-Baillie A and Snaith H J 2014 The emergence of perovskite solar cells *Nat. Photon.* **8** 506–14
[3] Park N G, Gratzel M, Miyasaka T, Zhu K and Emery K 2016 Towards stable and commercially available perovskite solar cells *Nat. Energy* **1** 16152
[4] Agresti A, Pescetelli S, Taheri B, Esa A, Rio Castillo D, Cin L, Bonaccorso F and Carlo A D 2016 Graphene–perovskite solar cells exceed 18% efficiency: a stability study *Chem. Sus. Chem* **9** 1–12
[5] Chang J A, Hyuk Im S, Lee Y H, Kim H, Lim C S, Heo J H and Seok S I 2012 Panchromatic photon-harvesting by hole-conducting materials in inorganic−organic heterojunction sensitized-solar cell through the formation of nanostructured electron channels *Nano Lett.* **12** 1863–7
[6] Lee M M, Teuscher J, Miyasaka T, Murakami T N and Snaith H J 2012 Efficient hybrid solar cells based on meso-super structured organometal halide perovskites *Science* **338** 643–7
[7] Burschka J, Pellet N, Moon S J, Humpry-Baker R, Gao P, Nazeeruddin M K and Gratzel M 2013 Sequential deposition as a route to high-performance perovskite-sensitized solar cells *Nature* **501** 395–8
[8] Malinkiewicz O, Yella A, Lee Y H, Espallargas G M, Graetzel M, Nazeeruddinand M K and Bolink H J 2014 Perovskite solar cells employing organic charge-transport layers *Nat. Photonics* **8** 128–32
[9] Kojima A, Teshima K, Shiral Y and Miyasaka T 2009 Organometal halide perovskites as visible-light sensitizers for photovoltaic cells *J. Am. Chem. Soc.* **131** 6050–1
[10] Yang W S *et al* 2017 Iodide management in formamidinium-lead-halide-based perovskite layers for efficient solar cells *Science* **356** 1376–9

[11] Stoumpos C C, Malliakas C D and Kanatzidis M G 2013 Semiconducting tin and lead iodide perovskites with organic cations: phase transitions, high mobilities, and near-infrared photoluminescent properties *Inorg. Chem.* **52** 9019–38
[12] Umari P, Mosconi E and De Angelis F 2014 Relativistic GW calculations on $CH_3NH_3PbI_3$ and $CH_3NH_3SnI_3$ perovskites for solar cell applications *Sci. Rep.* **4** 4467
[13] Hao F, Stoumpos C C, Cao D H, Chang R P H and Kanatzidis M G 2014 Lead-free solid-state organic-inorganic halide perovskite solar cells *Nat. Photon.* **8** 489–94
[14] Hao F, Stoumpos C C, Gao P, Zhou N, Marks T J, Chang R P H and Kanatzidis M G 2015 Solvent-mediated crystallization of $CH_3NH_3SnI_3$ films for heterojunction depleted perovskite solar cells *J. Am. Chem. Soc.* **137** 11445
[15] Marshall K P, Walton P I and Hatton R A 2015 Tin perovskite/fullerene planar layer photovoltaics: improving the efficiency and stability of lead-free devices *J. Mater. Chem.* A **3** 11631
[16] Kumar M H *et al* 2014 Lead-free halide perovskite solar cells with high photocurrents realized through vacancy modulation *Adv. Matter.* **26** 7122–7
[17] Lang L, Yang J H, Liu H R, Xiang J H and Gong X G 2014 First-principles study on the electronic and optical properties of cubic ABX_3 halide perovskites *Phys. Lett.* A **378** 290
[18] Noel N K *et al* 2014 Lead-free organic–inorganic tin halide perovskites for photovoltaic applications *Energy Environ. Sci.* **9** 3061
[19] Koh T M, Krishnamoorthy T, Yantara N, Shi C, Leong W L, Boix P P, Grimsdale A C, Mhaisalkar S G and Mathews N 2015 Formamidinium tin-based perovskite with low Eg for photovoltaic applications *J. Mater. Chem.* A **3** 14996
[20] Lee S J, Sik Shin S, Chan Kim Y, Kim D, Kyu Ahn T, Hong Noh J, Seo J and Seok S I 2016 Fabrication of formamidinium tin iodide perovskite solar cells through SnF_2-Pyrazine complex *J. Am. Chem. Soc.* **138** 3974–7
[21] Shao S, Liu J, Portale G, Fang H - H, Blake G R, Brink G H T, Jan Anton Koster L and Loi M A 2018 Highly reproducible Sn- based hybrid perovskite solar cells with 9% efficiency *Adv. Energy Mater.* **8** 1702019
[22] Hao F, Stoumpos C C, Hanh Cao D, Chang R P H and Kanatzidis M G 2014 Lead-free solid-state organic-inorganic halide perovskite solar cells *Nat. Photonics* **8** 489–94
[23] Jokar E, Hsun Chien C, Min Tsai C, Fathi A and Wei-Guang Diau E 2019 Robust tin-based perovskite solar cells with hybrid organic cations to attain efficiency approaching 10% *Adv. Mater.* **31** 1804835
[24] Cho A-N and Park N-G 2017 Impact of interfacial layers in perovskite solar cells *Chem. Sus. Chem* **10** 3687–704
[25] Mahmoudi T, Wang Y and Hahn Y-B 2018 Graphene and its derivatives for solar cells application *Nano Energy* **47** 51–65
[26] Lim E E, Yap C C, Hj Jumali M H, Teridi M A M and The C H 2018 A mini review: can graphene be a novel material for perovskite solar cell applications? *Nano-Micro Lett.* **10** 27
[27] Ferrari A C *et al* 2014 Science and technology roadmap for graphene, related two-dimensional crystals, and hybrid systems *Nanoscale.* **7** 4598
[28] Acik M, Lee G, Mattevi C, Chhowalla M, Cho K and Chabal Y J 2010 Unusual infrared-absorption mechanism in thermally reduced graphene oxide *Nat. Mater.* **9** 840
[29] Panchakarla L S, Subrahmanyam K S, Saha S K, Govindraj A, Krishnamurthy H R, Waghmare U V and Rao C N R 2009 Synthesis, structure and properties of boron and nitrogen doped graphene *Adv. Mater.* **21** 4726
[30] Wu T, Shen H, Sun L, Cheng B, Liua B and Shen J 2012 Nitrogen and boron doped monolayer graphene by chemical vapour deposition using polystyrene, urea and boric acid *New J. Chem.* **36** 1385
[31] Zhao L *et al* 2011 Visualizing individual nitrogen dopants in monolayer graphene *Science* **333** 999
[32] Zhang B, Hermet P and Henrard L 2010 scanning tunneling microscopy simulations of nitrogen- and boron-doped graphene and single-walled carbon nanotubes *ACS Nano* **4** 4165
[33] Gao S, Ren Z, Wan L, Zheng J, Guo P and Zhou Y 2011 Density functional theory prediction for diffusion of lithium on boron-doped graphene surface *Appl. Surf. Sci.* **257** 7443
[34] Wu M, Cao C and Jiang J Z 2010 Light non-metallic atom (B, N, O and F)-doped graphene: a first-principles study *Nanotechnology* **21** 505202
[35] Yang N, Yang D, Chen L, Liu D, Cai M and Fan X 2017 Design and adjustment of the graphene work function via size, modification, defects, and doping: a first-principle theory study *Nanoscale Res. Lett.* **12** 642
[36] Dieb T M, Hou Z and Tsuda K 2018 Structure prediction of boron-doped graphene by machine learning *J. Chem. Phys.* **148** 241716
[37] Woinska M, Milowska K and Majewski J A 2012 *Ab initio* modeling of graphene functionalized with boron and nitrogen *Acta Physica Polonica* A **122** 1087
Woinska M, Milowska K and Majewski J A 2012 Electronic structure of graphene functionalized with boron and nitrogen *Phys. Status Solidi* C **10** 1167–71
[38] Nevius M S, Conrad M, Wang F, Celis A, Nair M N, Taleb-Ibrahimi A, Tejeda A and Conrad E H 2015 Semiconducting graphene from highly ordered substrate interactions *Phys. Rev. Lett.* **115** 13680
[39] Li X, Zhu H, Wang K, Cao A, Wei J, Li C, Jia Y, Li Z, Li X and Wu D 2010 Graphene on-silicon schottky junction solar cells *Adv. Mater.* **100871** 2743–8
[40] Song Y, Li X, Mackin C, Zhang X, Fang W, Palacios T, Zhu H and Kong J 2015 Role of interfacial oxide in high-efficiency graphene–silicon schottky barrier solar cells *Nano Lett.* **15** 2104–10
[41] Patel K and Tyagi P K 2015 Multilayer graphene as a transparent conducting electrode in silicon heterojunction solar cells *AIP Advances* **5** 077165
Patel K and Tyagi P K 2017 P- type multilayer graphene as a highly efficient transparent conducting electrode in silicon heterojunction solar cells *Carbon* **116** 744–52
[42] Ishikawa R, Watanabe S, Yamazaki S, Oya T and Tsuboi N 2019 Perovskite/graphene solar cells without a hole-transport layer *ACS Appl. Energy Mater.* **2** 171–5
[43] Chandrasekhar P S and Komarala V K 2017 Graphene/ZnO nanocomposite as anelectron transport layer for perovskite solar cells; the effect of graphene concentration on photovoltaic performance *RSC Adv.* **7** 28610–5
[44] Wang J T, Ball J M, Barea E M, Abate A, Alexander-Webber J A, Huang J, Saliba M, Mora-Sero I, Bisquert J and Snaith H J 2013 Low-temperature processed electron collection layers of graphene/TiO_2 nanocomposites in thin film perovskite solar cell *Nano Lett.* **14** 724
[45] Agresti A, Pescetelli S, Palma A L, Del Rio Castillo A E, Konios D, Kakavelakis G, Razza S, Cina L, Kymakis E and Bonaccorso F 2017 Graphene interface engineering for perovskite solar modules: 12.6% power conversion efficiency over 50 cm^2 active area *ACS Energy Lett.* **2** 279
[46] Stangl R, Kriegel M and Schmidt M 2006 AFORS-HET, version 2.2, a numerical computer program for simulation of heterojunction solar cells and measurements *Proc. of the WCPEC-4, Fourth World Conf. on Photovoltaic Energy Conversion (Hawaii, USA)* pp 1350–3

[47] Stangl R, Haschke J and Leendertz C 2010 Numerical simulation of solar cells and solar cell characterization methods: the open-source on demand program AFORS-HET *Sol. Energy, Radu D Rugescu (Ed.), ISBN: 978-953-307-052-0, InTech* **432** 319–52
[48] Varache R, Leendertz C, Gueunier-Farret M E, Haschke J, Muñoz D and Korte L 2015 Investigation of selective junctions using a newly developed tunnel current model for solar cell applications *Sol. Energy Mater. Sol. Cells* **141** 14–23
[49] Yu Y-J, Zhao Y, Ryu S, Brus L E, Kim K S and Kim P 2009 Tuning the graphene work function by electric field effect *Nano Lett.* **9** 3430–4
[50] Hwang J O, Park J S, Choi D S, Kim J Y, Lee S H, Lee K E, Kim Y-H, Song M H, Yoo S and Kim S O 2012 Work function-tunable, N-doped reduced graphene transparent electrodes for high-performance polymer light-emitting diodes *ACS Nano.* **6** 159–67
[51] Du H J, Wang W C and Zhu J Z 2016 Device simulation of lead-free $CH_3NH_3SnI_3$ perovskite solar cells with high efficiency *Chin. Phys.* B **25** 108802
[52] Hadadian M *et al*, 2016 Enhancing efficiency of perovskite solar cells via N-doped graphene: crystal modification and surface passivation *Adv. Mater.* **28** 8681
[53] Lin Y *et al* 2013 Graphene/semiconductor heterojunction solar cells with modulated antireflection and graphene work function *Energy Environ. Sci.* **6** 108–15
[54] Fahad Bhopal M, Won Lee D, Rehman A and Lee S H 2017 Past and future of graphene/silicon heterojunction solar cells: a review *J. Mater. Chem.* C **5** 10701
[55] Ma L, Hao F, Stoumpos C C, Phelan B T, Wasielewski Michael R and Kanatzidis Mercouri G 2016 Carrier diffusion lengths of over 500 nm in lead-free perovskite $CH_3NH_3SnI_3$ films *J. Am. Chem. Soc.* **138** 14750–5
[56] Konstantakou M and Stergiopoulos T 2017 A critical review on tin halide perovskite solar cells *J. Mater. Chem.* A **5** 11518–49
[57] Schwenzer J A, Rakocevic L, Gehlhaar R, Abzieher T, Gharibzadeh S, Moghadamzadeh S, Quintilla A, Richards B S, Lemmer U and Paetzold U W 2018 Temperature variation induced performance decline of perovskite solar cells *ACS Appl. Mater. Interfaces* **10** 16390–9
[58] Bhatt P, Kumar M, Chandra Kant P, Panday M K and Tripathi B 2016 *Org. Electron.* **39** 258–66
[59] Kundu G, Singh M, Kumar R, Singh V and Chand F 2018 A theoretical study of perovskite material for solar cell application *AIP Conf. Proc.* **2006** (030018) 1–4

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/336196802

High Diversity Gain Super-Wideband Single Band-Notch MIMO Antenna for Multiple Wireless Applications

Article *in* IET Microwaves Antennas & Propagation · October 2019

DOI: 10.1049/iet-map.2019.0450

CITATIONS
0

READS
17

3 authors, including:

Gaurav Saxena
Delhi Technological University
2 PUBLICATIONS **2** CITATIONS
SEE PROFILE

Yogendra Awasthi
Manav Rachna International University
31 PUBLICATIONS **71** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

SUPER ULTRA WIDE BAND MIMO ANTENNA View project

All content following this page was uploaded by Gaurav Saxena on 01 October 2019.

The user has requested enhancement of the downloaded file.

Content may change prior to final publication in an issue of the journal. To cite the paper please use the doi provided on the Digital Library page.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited.
Content may change prior to final publication in an issue of the journal. To cite the paper please use the doi provided on the Digital Library page.

High Diversity Gain Super-Wideband Single Band-Notch MIMO Antenna for Multiple Wireless Applications

Gaurav Saxena[1, 2], Priyanka Jain[2*], Y K Awasthi[3]

[1]Electronics and Communication Engineering, Galgotia College of Engineering and Technology, Greater Noida, India
[2]Electronics and Communication Engineering, Delhi Technological University, Delhi, India
[3]Electronics and Communication Engineering, Manav Rachna International Institute of Research and Studies, Faridabad, India
[*]priyankajain@dtu.ac.in

Abstract: A two-element super-wideband MIMO antenna is proposed in this article. A step impedance microstrip feed line structure is introduced to achieve a super-wideband impedance bandwidth of 185%. This antenna is designed to be circularly polarized at 1.575GHz (L_1-band) and 26.0GHz (K-band) by using closed ring resonator on the radiating patch. The EBG structures are used to improve FBR in lower (1.5-5GHz) and higher band (25-35GHz). The proposed antenna is designed with compact size 55.6×50.5×1.6mm^3 on FR-4 substrate including exotic diversity performance. This MIMO antenna is suitable for defence and handheld devices covering GPS/ DCS/ PCS/ UMTS/ Wi-BRO/ ISM/ IRNSS/ LTE (M/HB)/ BLUETOOTH/ IoT/ WiMAX/ X/ Ku/ K/ Ka-band and it is also suitable for 5G (sub-6GHz/mm) applications with channel capacity of 11.34bps/Hz. Simulated results of MIMO antenna like radiated power based ECC and diversity parameters are also verified experimentally, which are in the acceptable range. Finally, this antenna is tested in a realistic environment for SAR application.

1. INTRODUCTION

Due to rapid advancement in wireless communication systems, better channel capacity with seamless high-data-rates connectivity is required. In order to fulfill the requirements, MIMO-communication systems are deployed that use multiple antennas on the transmitter and receiver end as depicted in Fig. 1. MIMO system is a state-of-art technology which improves the reliability of the communication systems by utilizing the diversity technique to mitigate the multi-path fading issues, where signals may come together contentiously at the receiver antennas. Spectral efficiency is also improved by the total transmitted power spreading over the MIMO antennas. Thus, MIMO antennas can increase channel capacities as well as the reliability of the communication system without sacrificing extra transmitted power or power spectrum. MIMO antennas have been designed to improve their characteristics in terms of impedance bandwidth, miniaturization & isolation. Several decoupling techniques are used in MIMO antennas to improve isolation and diversity performance [1-5]. For size miniaturization, a Yagi-shaped antenna is designed with better FBR [6]. Electromagnetic Band Gap (EBG) structures are also used to improve isolation in X-band MIMO antenna [7-9] which is useful for efficient wireless communication without affecting the adjacent antenna. In order to mitigate interference in UWB MIMO antenna, a WLAN notch is introduced using T-shaped open slots [10-12]. A Defected Ground Structure (DGS) mechanism has been introduced in MIMO antenna design to avoid correlation between two antenna and fast data rates [13]. Similarly, Split Ring Resonator (SRR) and Complementary Split Ring Resonator (CSRR) structures near the microstrip feeding line and ground plane were used respectively for high isolation [14-15]. The dual band-notched UWB MIMO antenna was tailored by cutting two rectangular slits on the patch and a fin-shaped decoupling structure was also introduced to get better ECC and directive gain [16]. MIMO antennas for sub 5G application (3.4-3.7GHz) is constructed without introducing any decoupling structure [17-18]. So, the advantages and disadvantages of the decoupling techniques over the proposed design are explained in Table 1 which has been used in MIMO antenna designs.

Table 1 Performance comparison of the decoupling techniques over the proposed design

Isolation Level (-dB)/Refs	Isolation Techniques	Advantages	Disadvantages
20 [3]	EBG WITH FSS, Double-layer EBG, Mushroom-shaped EBG	Improved isolation, increased diversity gain	Complex geometry
15 [10]	T-shaped	Improved isolation	Limited to UWB only
20 [11]	Defected Ground Structure (DGS)	Improved isolation at a lower frequency, compact structure	Complex structure
17 [14]	Split Ring Resonator (SRR)	Compact structure and high isolation bandwidth	Limited to a wide bandwidth
15 [15]	Complimentary Split-Ring Resonator (CSRR)	Compact structure and high isolation bandwidth	Limited to a wideband bandwidth
24 [16]	Funnel-shaped decoupling	Provided good isolation for 173% bandwidth	Large in size
20 [18]	F-shaped decoupling structure	Good isolation for UWB bandwidth	Poor isolation at low frequency (>2GHz)
16 [20]	Fractal uni-planar compact EBG	Improved isolation and FBR	Complex structure with narrow bandwidth
20 [25]	Flag-shaped stub on the ground plane	Good isolation for dual-band	Complex structure for the wideband only
13 [29]	Coupled-feed structure	Improved ECC, DG, MEG	Up to 6GHz only
20 [Proposed]	Mushroom-shaped EBG including E-shaped decoupling structure	Improved isolation for 1-40GHz band width with reduced cross-polarization	Complex structure

This article has been accepted for publication in a future issue of this journal, but has not been fully edited.
Content may change prior to final publication in an issue of the journal. To cite the paper please use the doi provided on the Digital Library page.

Specific Absorption Rate (SAR) is an important parameter for handheld devices. Therefore, it has been investigated for 1g and 10g tissue as per the Federal Communication Commission (FCC) [19]. Various types of decoupling techniques for antennas are deployed to enhance isolation which is responsible for good diversity parameters like ECC, DG, TARC, and CCL [20-25]. Furthermore, to enhance the throughput in the communication system, a cognitive radio LTE700 was demanded in wireless communication [26-30]. This type of antennas is also useful for IoT/WLAN/Wi-MAX/LTE and other multi-band applications. A technique is investigated to reduce Radar Cross Section (RCS) and enhance gain by using a polarized meta-surface with solid blocks of EBG structure in [31-32]. A semi-crescent cascaded structure is also designed for good diversity performance with circular polarization [33]. The circular polarization antennas are beneficial for moving handhold or mobile communications like GPS which needs at least half power when the object is movable so the antenna is required below 3dB axial ratio in a prescribed band [34]. The strip integrated circular polarized antenna is designed for radiofrequency energy harvesting application [35].

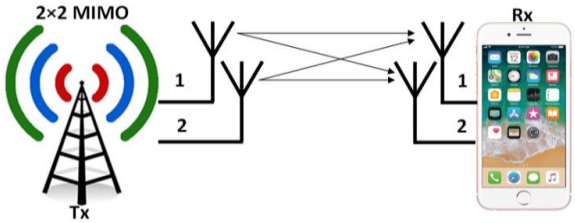


Fig.1. *A block-diagram of 2×2 SWB-MIMO antenna transceiver network*

In this paper, a compact super wideband (1.5-40GHz) dual-band circularly polarized MIMO antenna including notched-band characteristic (5.9-7.1GHz) is presented. The super-wideband (SWB) bandwidth of an antenna is attained by introducing a step impedance microstrip line to feed the radiator. Furthermore, two rectangular-shaped slots are etched on the ground plane to attain band-notched characteristic. Unequal E-shaped stubs and mushroom-shaped EBG structure is used for high isolation. Circular-shape (Closed Ring Resonator-CRR) is etched on the patch to achieve circular polarization in two bands. Section 2 explains the antenna design procedure. Section 3 demonstrates the simulated and measured results. Section 4 depicts the diversity performance of the antenna along with a comparison from literature. Section 5 concludes the results in terms of different parameters. The proposed antenna is designed using computer simulation technology-CST microwave studio simulation tool.

2. MIMO ANTENNA DESIGN PROCEDURE

The designed 2-element MIMO antenna including notched-band (5.9-7.1GHz) characteristic is fabricated on FR4 substrate of height (h) 1.6mm with relative permittivity (ε_r) of 4.4 and loss tangent (tanδ) of 0.02 having an overall dimension of 55.5mm×50.6mm as shown in Fig. 2 (a) - (b). All dimensions of the proposed antenna are given in Table 2. To achieve SWB impedance bandwidth (1.5-40GHz), the impedance feeding mechanism comprises of three feeding subsections having an impedance of 50, 65 and 81Ω respectively with different length. To improve the intended bandwidth, some other steps also discussed later. The current is accumulating along the periphery of the circular radiator which is affecting the bandwidth in a low-frequency range, Therefore, to improve the low-frequency bandwidth, a "Closed Ring Resonator-CRR" of radius 5.3mm is etched at circular patch without affecting the antenna performance.

Table 2 Proposed SWB MIMO antenna dimensions in mm

R_1	R_2	D_1	D_2	D_3	D_4	D_5	E_1	E_2	E_3
11	5.3	3.2	1.9	1.2	6.85	11.7	5.8	5.5	10.6
E_4	**E_5**	**G_1**	**G_2**	**C_1**	**C_2**	**C_3**	**C_4**	**C_5**	**C_6**
7.92	9.35	0.45	0.6	0.6	1.3	12.7	1.2	3.3	5.7
C_7	**C_8**	**C_9**	**B_1**	**B_2**	**B_3**	**B_4**	**B_5**	**B_6**	**B_7**
8.7	24.3	55.6	5.5	1.5	1.5	1.5	1.5	24	26.5

This antenna consists of a circular patch and a circular slot etched out of it to achieve circular polarization in dual-band. Further, E-shaped stubs are used with unequal arm length to achieve the intended isolation between antennas because E-shaped stubs of equal length cause poor isolation at a lower frequency. Stubs of unequal length are also helpful to extend bandwidth. The number of stubs is optimized to provide maximum decoupling at a high frequency so it cannot be improved further by increasing the number of stubs. Now, super-wideband including band-notch characteristic is obtained by using a stepped impedance feed line with a partially tailored ground plane together with rectangular slots. The diameter of the EBG slots (R_3) is 1.2mm. The fundamental resonance frequency (f_r) of the circular-shaped monopole antenna is calculated by equation (1) in Ref-2.

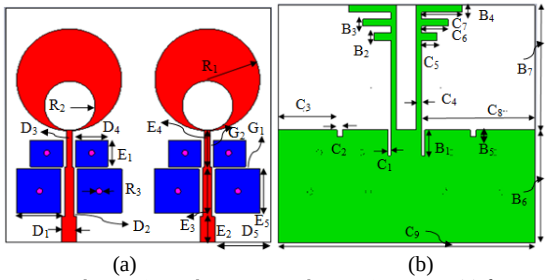


Fig.2. *Two-element SWB planar monopole MIMO antenna (a) front sight (b) bottom sight*

$$f_r(\mathrm{GHz}) = \frac{144}{C_9 + R_1 + h + \frac{A_1}{2\pi\sqrt{\varepsilon_r + 1}} + \frac{A_2}{2\pi\sqrt{\varepsilon_r + 1}}} \quad (1)$$

where A_1, A_2 is the area of anticipated radiation patch and ground plane, C_9 is the length of a ground plane, R_1 is a diameter of circular radiated patch depicted in Fig. 2 (a) - (b). The step by step design modus operandi of the proposed SWB-MIMO antenna is represented in Fig. 3 and analogous return losses and isolations are presented in Fig. 4 (a) - (b) respectively. Ant-1 provides unsatisfactory return loss and isolation in SWB bandwidth. By changing the feeding line width of Ant-1, return loss and isolation is improved as depicted in Ant-2 but return loss is not acceptable in the lower frequency range. Since charges are accumulated on the etched portion of patches, the discontinuity is added to improve the return loss with compromised isolation in the intended band. Therefore, in Ant-3, a circular slot is etched from the patches. Further, in Ant-4, mushroom-shaped EBG is designed near the feed line which improves the reflection coefficient and isolation in the intended bandwidth.

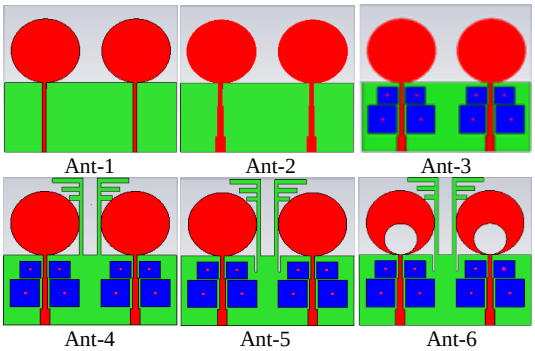


Fig.3. *Geometrical design steps of the proposed super-wideband MIMO antenna with band notch characteristics*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited.
Content may change prior to final publication in an issue of the journal. To cite the paper please use the doi provided on the Digital Library page.

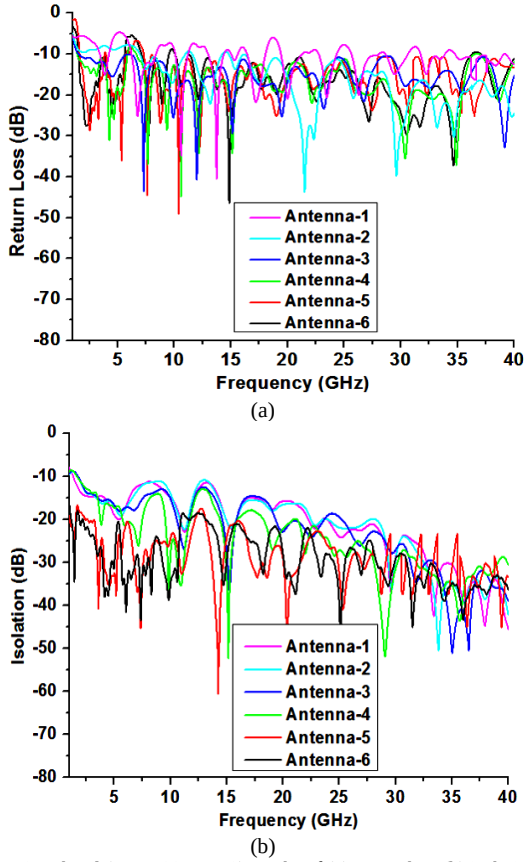


Fig.4. *Simulated (using CST MWS) results of (a) return loss (b) isolation of the antennas shown in Fig. 3*

Unequal E-shaped decoupling structure is made in the ground plane of Ant-5, to improve isolation (greater than -20dB) in the entire SWB range. Further, vertical slots are engraved in the ground plane to get better isolation in the frequency range 2.1 to 4.8GHz and it also creates a notch band characteristic for avoiding the interference from E-S satellite communication (5.9-7.1GHz) as shown in Fig.-3 Ant-6. The bandwidth of the notched band can be controlled by changing the ground slot length (B_1) at the notch frequency, $\mathbf{f_n = c/(2 \times B_1 \times \sqrt{\varepsilon_{reff}})}$, where c is the speed of the signal in a vacuum (3×10^8m/sec) and ε_{reff} is the effective relative permittivity [36] of a substrate. Center frequency of the intended notched-band is 6.5GHz at B_1.

EBG structures including a microstrip feed line and its equivalent circuit model with simulated results represented in Fig. 5 (a) - (d) and center frequency of the designed EBG is calculated by equation (2)

$$f_{EBG} = \frac{1}{2\pi\sqrt{L_{eq}(C_{gap\,eq} + C_{plate\,eq})}} \qquad (2)$$

C_{gap} is the gap capacitance between step impedance feeding line and square mushroom-shaped EBG structures having length E_1 & E_5 respectively and it is calculated by equation 3(a)-(b).

$$C_{gap1} = \frac{E_1\varepsilon_o(1+\varepsilon_r)}{\pi}\text{Cosh}^{-1}\left(\frac{E_1+G_2}{G_2}\right)F = 49.69\times10^{-15}\ F \qquad 3(a)$$

$$C_{gap3} = \frac{E_5\varepsilon_o(1+\varepsilon_r)}{\pi}\text{Cosh}^{-1}\left(\frac{E_5+G_2}{G_2}\right)F = 143.74\times10^{-15}\ F \qquad 3(b)$$

As C_{gap1} & C_{gap2} are equal and parallel therefore the total $C_{gap1,\,2}$ is $= C_{gap1} + C_{gap2} = 99.38\times10^{-15}$ F , similarly C_{gap3} & C_{gap4} are also equal and parallel so $C_{gap3,\,4}$ is $= C_{gap3} + C_{gap4} = 287.49\times10^{-15}$ F .

C_{plate} is the plate capacitance associated with EBG plate and ground plane, it is calculated by equation 4 (a)-(b) respectively

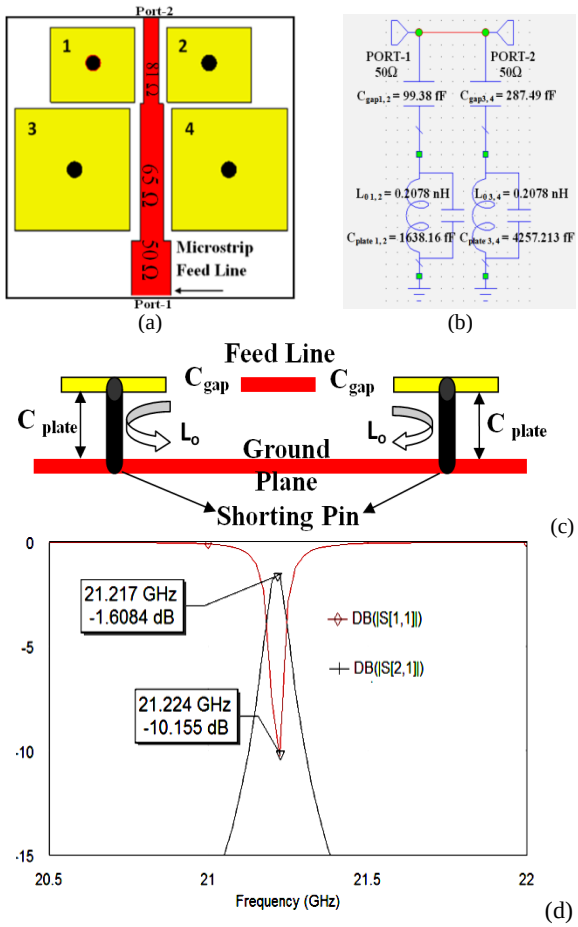


Fig.5. *Circuit analysis of Electromagnetic Band Gap structures (a) Mushroom-shaped EBG designs near to the microstrip stepped feed line (b) Equivalent circuit model of EBG (c) Front cross-sectional view of EBG (d) Simulated results of the equivalent circuit model of EBG with its resonance frequency*

$$C_{plate1} = \frac{\varepsilon_0\varepsilon_r E_1 D_4}{h}F = 819.08\times10^{-15}\ F \qquad 4(a)$$

$$C_{plate3} = \frac{\varepsilon_0\varepsilon_r E_5^2}{h}F = 2128.606\times10^{-15}\ F \qquad 4(b)$$

C_{plate1} & C_{plate2} are equal and parallel, therefore, $C_{plate1,2}$ is $= C_{plate1} + C_{plate2} = 1638.160$ fF , similarly C_{plate3} & C_{plate4} are equal and parallel so $C_{plate3,\,4}$ is $= C_{plate3} + C_{plate4} = 4257.213$ fF .

L_O is the inductive loading of centrally located shorting pin in EBG structure and it is calculated by equation (5)

$$L_O = 2\times10^{-7}\times h\left[\ln\left(\frac{2\times h}{R_3/2}\right) + 0.5\left(\frac{R_3}{h}\right) - 0.75\right]H = 0.4156\times10^{-9}H \qquad (5)$$

Four equal and parallel shorting pins associated with four mushroom-shaped EBG structures, therefore total equivalent inductance (L_{eq} = Lo/4) is 0.1039×10^{-9} H.

Hence, the calculated resonance frequency of the EBG structure is 21.224GHz by equation (2).

The proposed step quarter-wave impedance transformer technique without EBG is verified by a circuit simulation tool "Microwave Office" as shown in Fig. 6(a) and the proposed feeding mechanism is compared with existing feeding mechanism in Table 3. Stepped impedance feed line with EBG, shown in Fig. 5(a), provide permissible scattering results in the proposed antenna.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited.
Content may change prior to final publication in an issue of the journal. To cite the paper please use the doi provided on the Digital Library page.

Table 3 Comparison of proposed step impedance feeding mechanism with an existing feeding mechanism

Techniques	Return Loss (dB)	Frequency (GHz)	Features	Disadvantages
Quarter-wave impedance transformer technique (50Ω to 110Ω) [42]	25	0.6-1.6	Three transmission zero at θ = 60^0, 90^0, 120^0	Design for wideband only
Quarter-wave impedance transformer technique [43]	35	1.6,2.5,5.3	CP at a different frequency band	Design for multi-band with narrow bandwidth only
Coupled transmission line impedance transform technique (50Ω to 100Ω) [44]	20	0.4-1.7	Three transmission zero at θ =40^0, 90^0, 140^0	Design for wideband only
Step quarter-wave impedance transformer technique (50Ω to 81Ω) [Proposed]	10	1.5-40	Transmission coefficient near to 0dB for maximum power transfer	Poor return loss and only suitable for antenna

The return loss is less than -10dB throughout the SWB (1.5GHz-40GHz), which means that maximum input power appears at port-2 as shown in Fig.6 (b). In addition, E-field cross-polarization of the antenna is also improved due to the EBG structure.

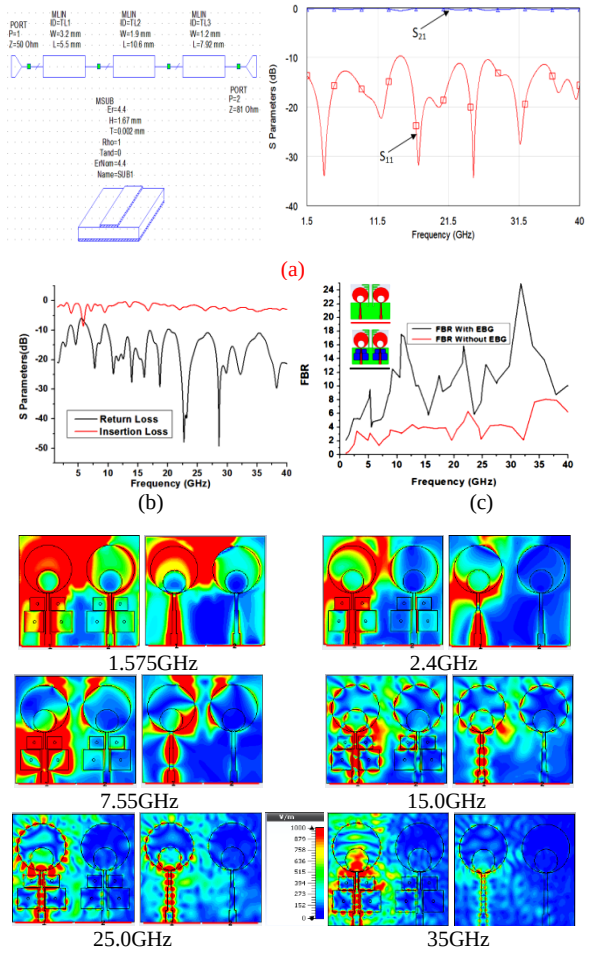


Fig.6. *Electromagnetic field analysis of stepped impedance feeding line and EBG structures (a) Simulated scattering results of stepped impedance feeding line only (b) Simulated scattering results of stepped impedance feeding line with EBG designs (c) Simulated FBR of SWB-antenna with and without EBG structures (d) Field distribution with and without EBG at 1.575, 2.4, 7.55, 15, 25, 35 GHz frequency when port-1 excited and port-2 terminated with a matched load*

FBR (Front to Back Ratio) is calculated with and without EBG to represent the existence of EBG in proposed antenna design as shown in Fig. 6 (c). As EBG is used to prevent the surface wave loss and direct the radiated power in the intended direction, therefore, the FBR of the antenna with EBG is higher than that of without EBG throughout the SWB bandwidth. Fig. 6 (d) represents the electric field distribution over the antenna, with EBG, and without EBG. It is shown that more fields are confined around EBG which improves the FBR in intended bandwidth.

3. Results And Discussion

Measured and simulated return loss, isolation, constant group delay, and peak gain of the designed antenna, as represented in Fig. 7 (a) – (b), are less than -10dB, -20dB, almost 1.0ns, and 7.6dB respectively, throughout the SWB bandwidth except notched band. The return loss begins to deteriorate at high frequencies especially above 35GHz because at high frequency, inductance effect starts to dominate in the proposed impedance matching mechanism. Also, the current is confined on the surface of the conductor due to skin effect which creates impedance mismatch and provides poor return loss. But the simulated return loss is still less than -10dB. Mobile phone and other movable handheld device antennas require at least half power in every plane for effective communication link so circular polarization of antenna for GPS and mm-wave 5G applications is planned. The proposed antenna is circularly polarized in dual-band which is achieved by etching circular slot (Closed Ring Resonator: CRR) out of the radiating patch.

All steps of the proposed antenna design from linear to circular polarization is given in Fig. 3 with its axial ratio shown in Fig. 7(c)-(d). It is observed from the results of antennas (Ant-1 to Ant-5) that the axial ratio of antennas in Fig.7 (c) and (d) w.r.t. frequency is greater than 6dB throughout the Superwideband. Hence, all antenna designs (Ant-1 to Ant-5) are linearly polarized. Although, the proposed Ant-6 with an etched CRR of the optimized radius produces circular polarization at 1.575GHz and 26.0GHz having an axial ratio less than 3dB. The optimized radius of the circular patch and etched CRR are 11.0mm (R_1) and 5.3mm (R_2) respectively.

The simulated current density of antenna at 1.575, 2.40, 2.45, 3.50, 5.50, 8.69, 15.0, 25.0, 35.0GHz frequencies is represented in Fig. 8. The main cause behind the mutual coupling between the adjacent antenna elements is the flow of surface current from the energized antenna 1 (left) to antenna 2. Extremely low surface current on antenna-2 is observed due to the excitation of antenna-1 as shown in Fig.8. Hence, the level of mutual coupling between antenna-1 and antenna-2 is very low. Means, this proposed decoupling mechanism is useful in various antenna applications without affecting the adjacent antenna. In the present case, E-shaped stubs in a ground plane are used for achieving reduced mutual coupling between antennas. Therefore, better isolation is acquired in the intended frequency band.

The radiation distinctiveness of the MIMO antenna is investigated in the SWB bandwidth. To measure radiation from antenna-1, antenna-2 is terminated with a matched load. 3D radiation patterns are shown in Fig. 9 (a), where the peak of the radiation patterns of an antenna is in one direction at all frequencies. When antenna 2 (right) is excited, the power is radiating towards the right of the yz- plane. Thus, the proposed antenna gets pattern diversity which contributes to better isolation between the two antennas. The designed MIMO antenna is fabricated on an FR4 substrate using photolithography technique, shown in Fig. 9 (b). Antenna under test is placed in an anechoic chamber associated with VNA (Vector Network Analyzer) for measurement of antenna characteristics, represented in Fig. 9 (c).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited.
Content may change prior to final publication in an issue of the journal. To cite the paper please use the doi provided on the Digital Library page.

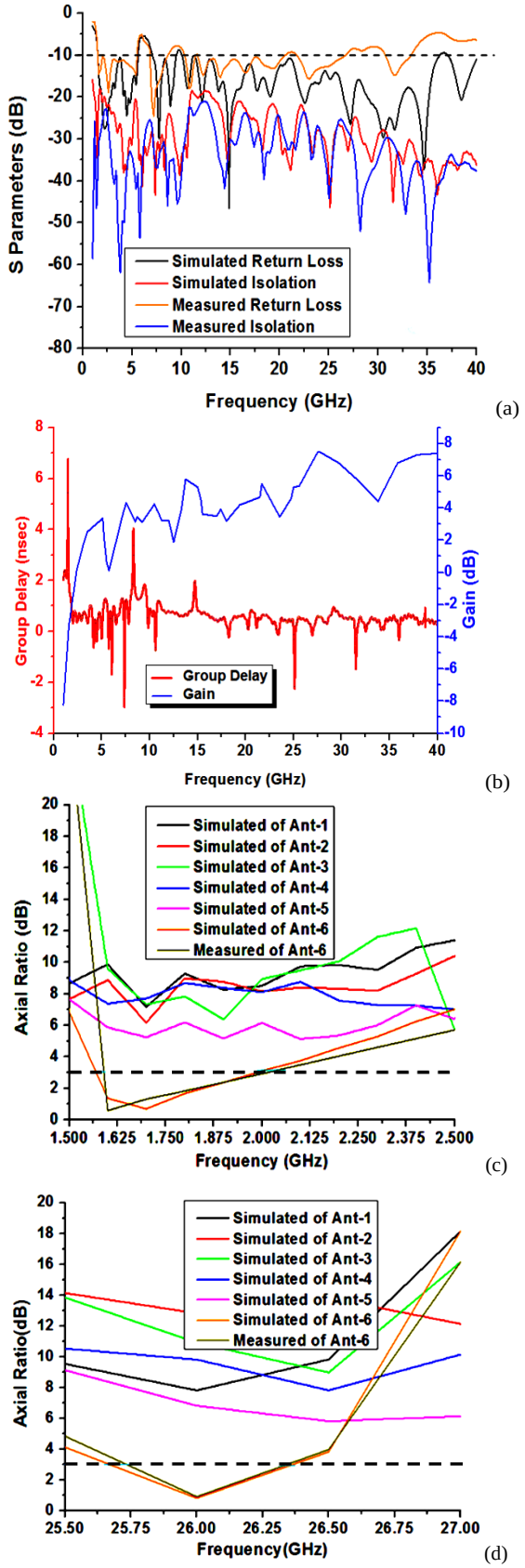


Fig.7. Simulated and measured (a) S-parameters (return & isolation) (b) group delay and gain (c) Parametric analysis of axial ratio in L_1-band (d) Parametric analysis of axial ratio in K-band of SWB-MIMO antenna

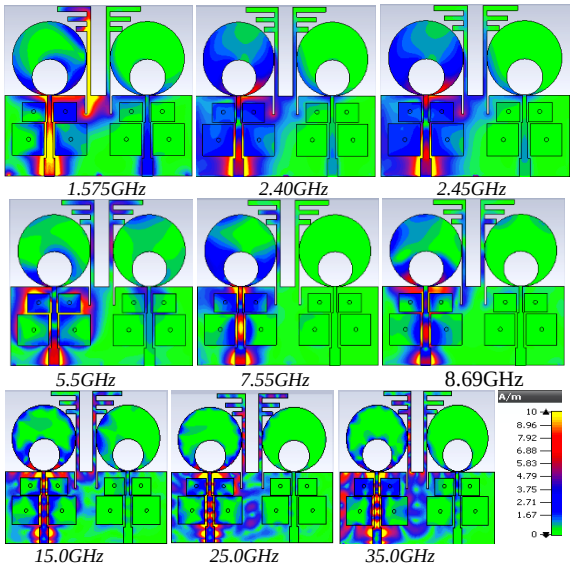


Fig.8. Surface current density distribution at 1.575, 2.40, 2.45, 5.50, 7.55, 8.69, 15.0, 25.0, & 35.0GHz when antenna 1 (left) is energized and antenna 2 (right) terminated with match load

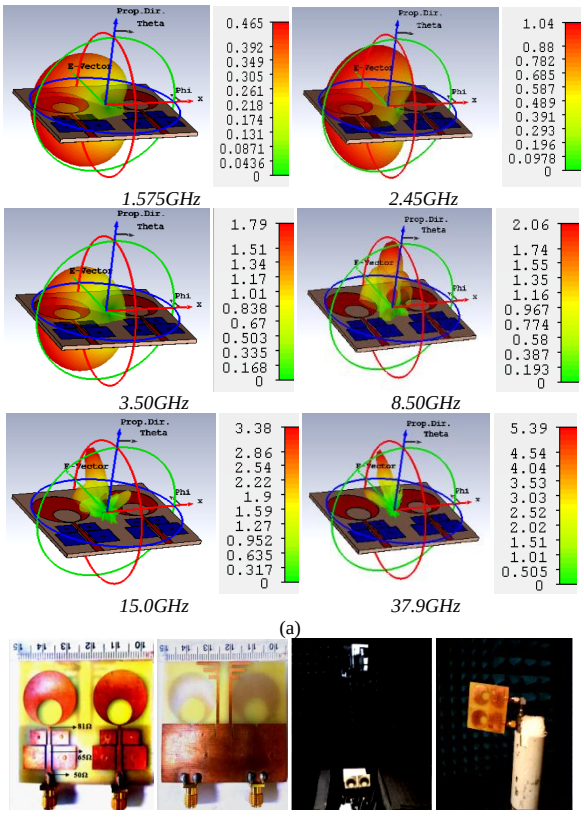


Fig.9. 3-D radiation pattern of designed MIMO antenna at (a) 1.575, 2.45, 3.50, 8.50, 15.0, & 37.9GHz (b) fabricated prototype top and bottom view (c) antenna under test in anechoic chamber for radiation pattern measurement

Fig.10 presents the measured & simulated yz-plane (left) and xz-plane (right) 2D radiation patterns of the designed antenna at 1.575, 2.40 and 28.0GHz frequencies respectively with an acceptable agreement. It is noted that the patterns at operating frequency band are omnidirectional at $\Phi = 0^o$ (*xz*)-plane and dipole like at $\Phi = 90^o$ (*yz*)-plane except notched band (5.9-6.9GHz). At higher frequency, the splitting of the radiation lobes is also noted

This article has been accepted for publication in a future issue of this journal, but has not been fully edited.
Content may change prior to final publication in an issue of the journal. To cite the paper please use the doi provided on the Digital Library page.

due to higher-order mode propagation. The EBG structure is helpful in preventing cross-polarization in the proposed antenna.

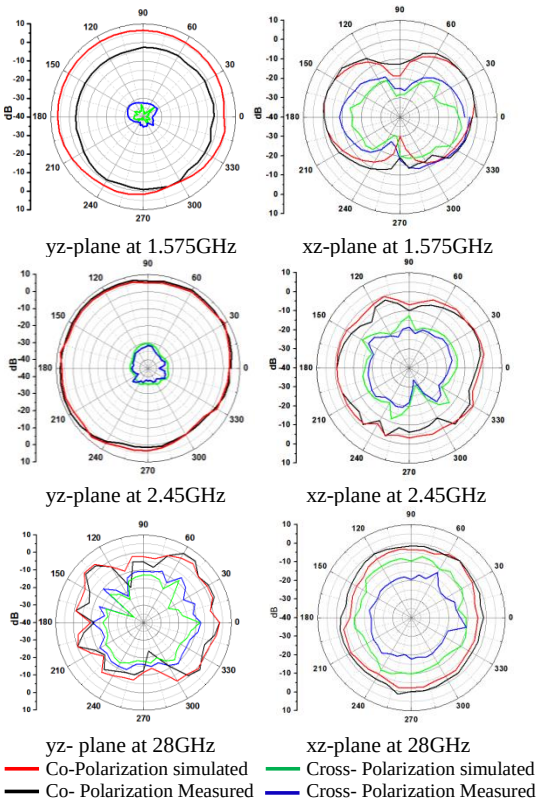


Fig.10. *Measured & simulated 2D radiation patterns at 1.575, 2.45, and 28.0GHz for yz- plane (left) and xz- plane (right), antenna 1 is energized and antenna 2 is terminated by a matched load*

Specific Absorption Rate: SAR is a vital parameter in the antenna at which electromagnetic energy is consumed by the human body when it is exposed, as shown in Fig 11. SAR is explained as "the power absorbed per mass of tissue". SAR is usually averaged either over a small sample volume (typically 1g or 10g of tissue) or over the whole body. The SAR value of any RF and microwave electronic devices is obtained by penetration of electric field intensity in the human head tissue (Table 4) near field environment [37-38]. Furthermore, SAR estimation on the head at 1.575, 2.40 and 5.50GHz frequencies are listed in Table 5. For the estimation of SAR, the values used for radius of skin, bone, and brain are 79.324, 75, and 65mm respectively.

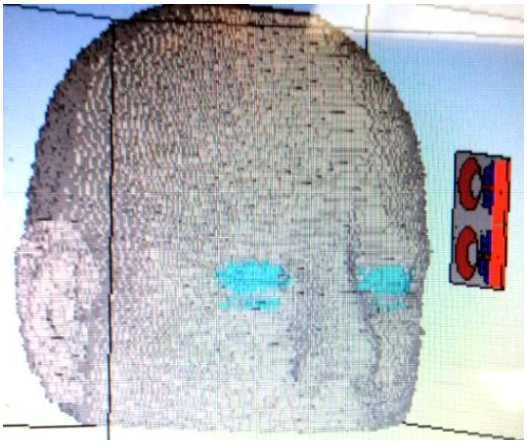


Fig.11. *Calculation of specific absorption rate near the human head in presence of the designed MIMO antenna*

Table 4 Parameters of the human head used for calculation of specific absorption rate by the proposed antenna at constant permeability

Type	Density (ρ) kg/m^3	Thermal Conductivity (σ) S/m	Heat capacity kJ/K/kg	Blood flow W/K/m^3	Metabolism rate W/m^3
Bio-tissue skin	1100	0.293	3.5	9100	1620
Bio-tissue Bone	1850	0.41	1.3	3400	610
Bio-tissue Brain	1030	1.13	3.675	40000	7100

Table 5 SAR values near the human head at the different resonant frequencies when the distance between the human head and MIMO antenna is 15mm

Specific Absorption Rate (W/kg)					
1.575GHz		2.4GHz		5.5GHz	
1(g)	10(g)	1(g)	10(g)	1(g)	10(g)
0.041	0.021	0.048	0.026	0.053	0.032

The calculated value of SAR is less than 1.6W/kg for the proposed antenna as shown in Table 5, for all parameters listed in Table 4. This SAR performance is investigated without any plastic jacket or cover. The SAR of electromagnetic energy is calculated by equation (6)

$$\text{Specific Absorption Rate} = \frac{1}{V}\int \frac{\sigma(r)|E(r)|^2}{\rho(r)} dr \qquad (6)$$

Where σ(r) is thermal conductivity of model in S/m, $E_{rms}(r)$ is rms electric field in V/m, ρ(r) is sample density in kg/m^3, V is the volume of the sample in m^3. SAR is the function of the input power and 100mW power is used for calculation.

4. MIMO Diversity Performance

MIMO diversity performance of the designed antenna is evaluated in terms of Envelope Correlation Coefficient (ECC), Total Active Reflection Co-efficient (TARC), Diversity Gain (DG), Channel capacity loss (CCL). The mutual coupling between two radiators and the amount of correlation between them is calculated by ECC.

It is calculated through S-parameters by equation (7) (a) and this equation is only valid for lossless uniformly distributed power through antennas

$$ECC = \frac{\left|S_{11}^{*}S_{12}+S_{21}^{*}S_{22}\right|^2}{\left(1-|S_{11}|^2+|S_{21}|^2\right)\times\left(1-|S_{22}|^2+|S_{12}|^2\right)} \qquad 7(a)$$

But it provides a hypothetical result which varies from empirical result so, the computation the ECC of the MIMO antenna by equation (7) (b) in terms of radiated fields

$$ECC = \frac{\left|\iint_{4\pi}\left[E_i(\theta,\phi)\times E_j(\theta,\phi)\right]d\Omega\right|^2}{\iint_{4\pi}\left|E_i(\theta,\phi)\right|^2 d\Omega \times \iint_{4\pi}\left|E_j(\theta,\phi)\right|^2 d\Omega} \qquad 7(b)$$

where E_i & E_j are the radiated electric fields vector of the ith and jth elements of the MIMO antenna system.

The ideal value of ECC is zero for uncorrelated MIMO antenna but their acceptable value for practical MIMO antenna is ≤0.5. In Fig.12 (a), the value of ECC of designed MIMO antenna by using S-parameters and radiated fields in super-wideband is observed to be less than 0.005 and 0.03 respectively, excluding the notched-band. Hence, the above results show a pretty low

This article has been accepted for publication in a future issue of this journal, but has not been fully edited.
Content may change prior to final publication in an issue of the journal. To cite the paper please use the doi provided on the Digital Library page.

correlation between the two antennas, which indicates excellent diversity performance.

The diversity gain of the MIMO antenna is calculated approximately by the formula given by $DG = 10 \times \sqrt{(1 - ECC^2)}$ and for satisfactory operation of MIMO antenna, diversity gain should be close to 10dB. The value of diversity gain of the designed super-wideband antenna is 9.99dB apart from the notched band. The ECC and DG with and without EBG structures are varying from 0.0125 to 0.005 and 9.7dB to 9.99dB respectively as depicted in Fig.12 (b). In order to optimize, the channel capacity and multiplexing efficiency of an antenna is a maneuver parameter which is not only account the total antenna efficiency but also the correlation and efficiency imbalance. In addition, multiplexing efficiency (η_{MUX}) is the ratio of the power requirement of an antenna to the reference antenna. The multiplexing efficiency (η_{MUX}) of the designed MIMO antenna is calculated by $\sqrt{(1 - |\rho_c|^2)\eta_1\eta_2}$, where ρ_c is a complex correlation between two side by side antenna i.e. ECC = ρ_c^2 and η_1, η_2 is the total efficiency of the designed MIMO antenna elements. Hence, the total efficiency and multiplexing efficiency are almost identical and its value goes down at the notched band only.

Channel Capacity: Channel capacity parameter of the MIMO, used to evaluate the system performance, is a function of bandwidth and the SNR (signal to noise) ratio, similar to the SISO system. Power is equally feed-in transmitting antenna elements without knowing the channel conditions. The designed 2 element array SWB antenna channel capacity is verified by Ergodic MIMO channel capacity and it is calculated by equation 8. From Fig.13, the minimum and maximum channel capacity of the 2 element array antenna system is 7.5bps/Hz and 11.34bps/Hz respectively at 20dB SNR in a uniform environment.

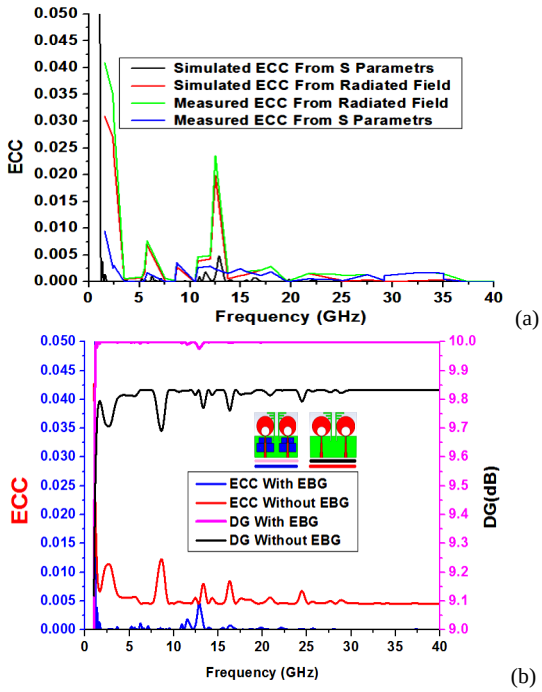


Fig.12. *Diversity performance of proposed SUWB antenna (a) Measured and simulated results of envelope correlation coefficient and diversity gain (b) Simulated results of ECC and DG with and without EBG*

The correlation matrix of MIMO system is R_{xx} (α) = (α) [I] + (1-α) [H][H*], where, 0< α < k / (k-1) and k is the number of antenna elements, For minimum channel Capacity, the value of channel correlation coefficient (α) is zero, therefore, all channels are correlated to each other. Furthermore, the value of α is one for maximum channel capacity, hence all channel has no correlation and fading matrix [H][H*] is converted into an identity matrix. The maximum value of the channel capacity of the 2×2 MIMO antenna system is calculated by the equation 8 (a)

$$C_{2\times2\ MIMO\ Max.} = k\left\{\log_2\left[\det\left([I]+\frac{SNR}{k}[H][H^*]\right)\right]\right\} \qquad (8)$$

= 4 × $\log_2$ (1+50) = **22.68 bits/sec/Hz**

Where, SNR is 20dB for Rayleigh fading environment i.e. SNR = 100. As the permissible limit of channel capacity of the 2×2 MIMO antenna system is 65% of its maximum value [39] (22.68 bits/sec/Hz) = 22.68 ×0.65 = **14.74 bits/sec**/Hz by using Monte Carlo simulation for correlated Rayleigh channel.

The upper level of channel capacity of proposed 2 elements MIMO is calculated by equation (8)

$C_{2\text{-element MIMO}}$ = 2×$\log_2$ (1+50) =11.344bits/sec/Hz

Permissible level of channel capacity is 65% of its maximum value (11.344 bits/sec/Hz) = 11.344 × 0.65 = 7.38 bits/sec/Hz [39]. [I] is the identity matrix, SNR is the ratio of signals of device terminal and distributed channels (independent & identical) with Rayleigh fading environment i.e. SNR = 20dB. k is the number of transmitting antenna elements. [H] is the fading matrix and [H*] is Hermitian transpose of a fading matrix.

Channel capacity of MIMO system is linearly increasing with increasing the number of elements; the element matrix properly characterizes the performance of a MIMO system as it provides the information about the phase and gain characteristic of the transmitting and receiving elements. However, in the presence of the correlation factor in MIMO antennas, the channel capacity losses are also increased. Therefore, the diversity performance of the proposed antenna is affected due to channel capacity losses. It is calculated by the equation (9)

$$CCL\ (dB) = -\log_2[\det(\varphi^c)] \qquad (9)$$

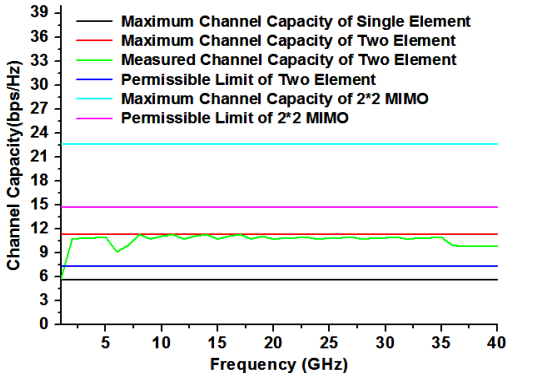


Fig.13. *Comparison of channel capacity of 2 element antennas array with a maximum and permissible limit of a 2×2 MIMO antenna system*

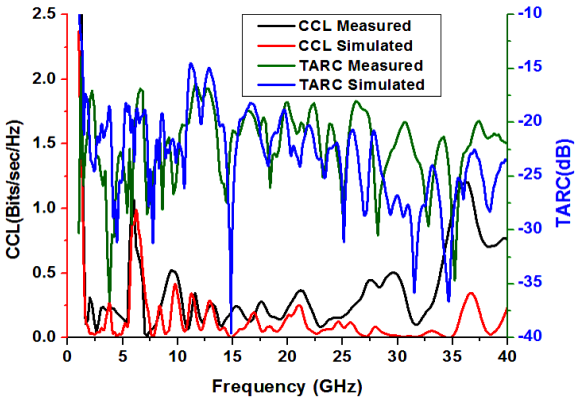


Fig.14. *Simulated and measured results of channel capacity loss (CCL) and total active reflection coefficient (TARC) of proposed MIMO antenna*

This article has been accepted for publication in a future issue of this journal, but has not been fully edited.
Content may change prior to final publication in an issue of the journal. To cite the paper please use the doi provided on the Digital Library page.

Table 6 Simulated MEG results of proposed MIMO antenna at various frequencies

Frequency (GHz)	MEG(dB) Isotropic medium		MEG(dB) Gaussian medium	
	XPR = 1dB	XPR = 6dB	XPR = 1dB	XPR = 6dB
1.575	-3.4	-4.7	-4.2	-6.6
15	-3.1	-3.4	-3.4	-4.5
30	-3.0	-2.5	-4.3	-2.9
39	-3.0	-2.3	-3.8	-2.8

where φ^c is 2×2 S-parameter correlation matrix $\varphi_{ii} = 1 - |S_{ii}|^2 - |S_{ij}|^2$, $\varphi_{ij} = -(S_{ii} \times S_{ij} + S_{ji} \times S_{jj})$. Fig.14 shows the CCLs and it is less than 0.35bps/Hz over the entire bandwidth except for the notched band, which offer a better diversity result of the proposed MIMO antenna.

In the two-port antenna system, contiguous antenna elements impinge on each other when they are working concurrently. This influences the overall desirable gain, efficiency, and bandwidth. Hence, actual MIMO antenna system performance cannot be predicted by s-parameters only, so TARC is also calculated to account for this effect. "TARC is the square root of the ratio of total reflected power to the total incident power" and hence, it gives the information about the apparent return loss of the MIMO antenna system. For side by side two-port radiated antenna system, it can be evaluated by equation (10)

$$\text{TARC} = \sqrt{\frac{(S_{ii}+S_{ij})^2+(S_{jj}+S_{ji})^2}{2}} \quad (10)$$

Ideally, TARC should be less than 0dB for MIMO antenna system. The simulated and measured TARC of the designed MIMO antenna is less than -15dB in the complete super-wideband bandwidth except for the notched band, again it shows the excellent diversity performance of a designed antenna.

Moreover, the mean effective gain (MEG) is the ratio of the mean received power by the i[th] antenna to the mean incident power of the j[th] antenna with the same route, $MEG_i = \varepsilon_{Total}/2$, where ε_{Total} is implying total effective efficiency of the i[th] antenna, *i.e.* $\varepsilon^i_{Total} = \varepsilon^i_{mis} \times \varepsilon^i_{rad}$, and where $\varepsilon^i_{mis} = 1 - \sum_{j=1}^{N}|S_{ij}|^2$, $\varepsilon^i_{rad} = \sum_{j=1}^{N}|S_{ij}|^2$ where, ε^i_{mis}, and ε^i_{rad} are mismatch efficiency and radiation efficiency of i[th] antenna. MEG ratio for the proposed MIMO antenna is extracted by equation (11)

$$MEG_i = \frac{P_{rec}}{P_{inc}} = \iint\left[\frac{XPR \times G_{\theta i}(\Omega)+G_{\phi i}(\Omega)\times P_{\phi}(\Omega)}{1+XPR}\right]d\Omega \quad (11)$$

Where XPR is the cross-polarization ratio and $G_{\theta i}(\Omega)$, $G_{\Phi i}(\Omega)$ & $P_{\Phi}(\Omega)$ is the gain and power density functions of the incident wave, Ω is beam area. The acceptable ratio of mean effective gain of i[th] and j[th] antenna elements should be less than or equal to ±3dB.

The MEG is also represented in isotropic and Gaussian mediums, to investigate the diversity performance at XPR = 1.0dB (outdoor) and 6.0dB (indoor) as shown in Fig. 15 and Table 6. The MEG for the isotropic medium at 1.0dB XPR is varying in the span of -3.4dB to -3.0dB and at 6.0dB XPR is lying between of -4.7 to -2.3dB. Similarly, the MEG for Gaussian medium at 1.0dB XPR and 6.0dB XPR is lying among -3.4dB to -4.2dB, and -6.6 to -2.8dB respectively for entire super-wideband bandwidth. From Fig.15, the MEG ratio of the proposed antenna is found in acceptable limit throughout the super-wideband bandwidth, Hence, it is as well a high-quality channel performance of antenna for wireless communication.

Fig.16 shows the power loss, power accepted, power outgoing to all ports, power radiated and power stimulated of proposed MIMO antenna for SAR calculations. 0.5W stimulated power is used to calculate the diversity performance parameters. Power loss in metal is very less in mW order so most of the power is radiated. Power accepted is almost the same as input power or stimulated power except for notched band because, in that band, maximum power returns back to the same ports. The radiated power is high in super-wideband range except for notch band. The radiated power is always less than stimulated and accepted power because of dielectric, conductor, and surface wave loss in a microstrip line. Power going to all ports is highly interfere at notched-band (5.9-6.9GHz) due to correlation.

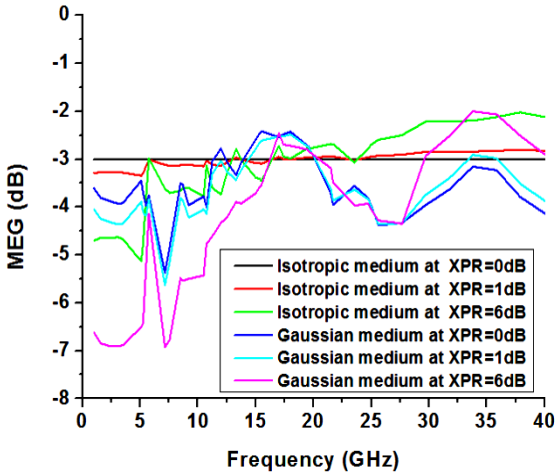


Fig.15. *Simulated results of mean effective gain in Isotropic and Gaussian medium at different XPR values of the proposed antenna*

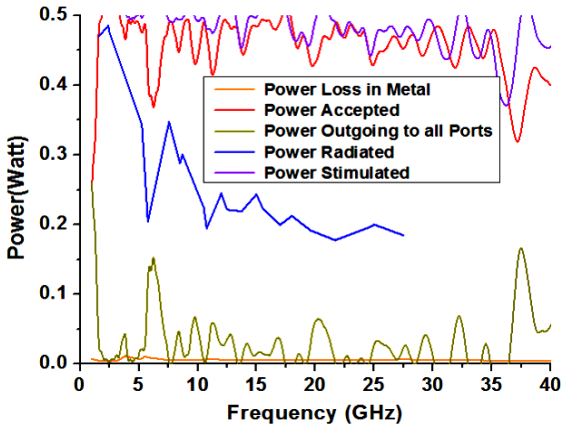


Fig.16. *Power level of 2 elements MIMO antenna in various states*

Table 7 Comparison of designed SWB-MIMO antenna with existing references in various aspects

Size (mm^2)	Impedance Bandwidth (GHz)	Isolation (dB)	ECC	Maximum Gain (dBi)
45×25	3.1-11.5	-19	0.20	5.5 [1]
27.4×27.4	12-22	-18	0.035	10.94 [2]
60×70	0.89-0.97,1.88—1.98, 2.2-2.4, 2.5-2.9	-15	0.20	3.6 [3]
69.5×44.3	1.575,2.4,5.4	-10	0.20	5.5 [5]
33×26	3.1-10.6	-15	0.03	6.0 [10]
26×15	3.1-35	-24	0.01	7.8 [16]
25×25	3.2-3.8	-12	0.01	4.2 [17]
50×30	2.5-14.5	-20	0.04	4.2 [18]
100×50	2.45-5.8	-20	0.04	5.18 [19]
60×50	1.27-1.43,1.8-2.133	-15	0.0785	5.0 [20]
68×136	2.55-2.65	-12	0.20	3.0 [21]
150×110	0.699-0.798,1.92-2.17	-15	0.085	3.0 [22]
54.98×76	2.1-4.0	-12	0.026	2.83 [23]
8×8	2.2-13.3	-15	0.04	6.0 [24]
8×8	2.2-3.2	-30	0.03	3.1 [33]
55.6×50.5	**1.5-40**	**-20**	**0.005**	**7.5 [P]**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited.
Content may change prior to final publication in an issue of the journal. To cite the paper please use the doi provided on the Digital Library page.

Table 8 Comparison of SAR and CCL of designed SWB-MIMO antenna with existing references

Frequency (GHz)	Specific Absorption Rate (SAR)		Channel Capacity Loss (CCL)	
	1g (W/Kg)	10g (W/Kg)	Frequency (GHz)	CCL (bps/Hz)
2.45,5.8	0.246,0.122	0.124, 0.045 [19]	8-18	0.35 [7]
2.1,2.4,4.0	0.8,0.9,1.2	0.6,0.8, 0.8 [23]	0.95–6.13	0.50 [40]
2.4	0.0368	0.0138 [32]	3.08-10.98	0.35 [41]
1.575,2.4, 5.5	**0.041,0.048, 0.053**	**0.021,0.026, 0.032 [P]**	**1.5-40**	**0.35 [P]**

It is noted from the Table 7 & 8 that the designed antenna achieves better results in comparison with the referred MIMO antennas at various specifications, which is highly appreciable as the designed antenna is small in size with super-wideband impedance bandwidth. Group delay (ns) is also calculated to verify the time-domain characteristic of the MIMO antenna. The two identical antennas are aligned in front of each other as a receiver and transmitter kept 100cm apart to obtain group delay deviation. It is found to be almost constant in the super-wideband bandwidth excluding notched band.

5. CONCLUSION

In this article, a dual-band circularly polarized two-element MIMO antenna with band-notch characteristic has been proposed for super-wideband applications having -20dB isolation. All simulated results of antenna have been verified with measured results to the entire impedance bandwidth including the notched-band (5.9-7.1GHz) which mitigates the interference from E-S satellite communication. The antenna achieved 7.6dBi peak gain including stable omnidirectional radiation patterns with excellent ECC, less than 0.005. The SAR performance of the antenna is also demonstrated which under the permissible limits and the average channel capacity is 11.34bps/Hz. The performance of the proposed antenna proved that it is a good candidate in various portable wireless applications.

6. REFERENCES

1. Roshna, T.K., Deepak, U., and Mohanan, P., 'Compact Uwb MIMO Antenna for Tridirectional Pattern Diversity Characteristics', IET Microwaves, Antennas & Propagation, 2017, 11, (14), pp. 2059-2065.
2. Satam, V. and Nema, S., 'Compact, High Gain, Linearly Polarized Diversity Antenna for Ku-Band Applications', Microwave and Optical Technology Letters, 2018, 60, (11), pp. 2843-2849.
3. Moustafa, L. and Jecko, B., 'Design of a Wideband Highly Directive EBG Antenna Using Double-Layer Frequency Selective Surfaces and Multifeed Technique for Application in the Ku-Band', IEEE Antennas and Wireless Propagation Letters, 2010, 9, pp. 342-346.
4. Jin, Y., Ko, M., O, Y., and Choi, J., 'A Planar UWB MIMO Antenna with Gain Enhancement and Isolation Improvement for the 5G Mobile Platform', Microwave and Optical Technology Letters, 2019, 61, (4), pp. 990-998.
5. Kwon, O. Y., Song, R., and Kim, B.S., 'A Fully Integrated Shark-Fin Antenna for MIMO-LTE, GPS, WLAN, and Wave Applications', IEEE Antennas and Wireless Propagation Letters, 2018, 17, (4), pp. 600-603.
6. Jehangir, S.S., Sharawi, M.S., and Shamim, A., 'Highly Miniaturised Semi-Loop Meandered Dual-Band MIMO Antenna System', IET Microwaves, Antennas & Propagation, 2018, 12, (6), pp. 864-871.
7. Saxena, G., Jain, P. and Awasthi, Y.K., 'High Isolation EBG Based MIMO Antenna for X-Band Applications'. IEEE 6th International Conference on Signal Processing and Integrated Networks (SPIN), India, March 2019, pp. 97-100.
8. Exposito-Dominguez, G., Fernandez-Gonzalez, J.M., Padilla, P., and Sierra-Castaner, M., 'Mutual Coupling Reduction Using EBG in Steering Antennas', IEEE Antennas and Wireless Propagation Letters, 2012, 11, pp. 1265-1268.
9. Payandehjoo, K. and Abhari, R., 'Employing EBG Structures in Multiantenna Systems for Improving Isolation and Diversity Gain', IEEE Antennas and Wireless Propagation Letters, 2009, 8, pp. 1162-1165.
10. Liu, G., Liu, Y., and Gong, S., 'Compact Uniplanar UWB MIMO Antenna with Band-Notched Characteristic', Microwave and Optical Technology Letters, 2017, 59, (9), pp. 2207-2212.
11. Khan, A.A., Jamaluddin, M.H., Aqeel, S., Nasir, J. and Owais, O.,'Dual-band MIMO dielectric resonator antenna for WiMAX/WLAN applications. IET Microwaves, Antennas & Propagation, 2017,11, (1), pp.113-120.
12. Kan, G., Lin, W., Liu, C., and Zou, D., 'An Array Antenna Based on Coplanar Parasitic Patch Structure', Microwave and Optical Technology Letters, 2018, 60, (4), pp. 1016-1023.
13. Madhav, B.T.P., Usha Devi, Y., and Anilkumar, T., 'Defected Ground Structured Compact MIMO Antenna with Low Mutual Coupling for Automotive Communications', Microwave and Optical Technology Letters, 2019, 61, (3), pp. 794-800.
14. Ramachandran, A., Mathew, S., Rajan, V., and Kesavath, V., 'A Compact Triband Quad-Element MIMO Antenna Using Srr Ring for High Isolation', IEEE Antennas and Wireless Propagation Letters, 2017, 16, pp. 1409-1412.
15. Asadpor L, Rezvani M. 'Multiband microstrip MIMO antenna with CSRR loaded for GSM and LTE applications, Microwave and Optical Technology Letters,2018; pp.3076–3080.
16. Gautam, A.K., Yadav, S., and Rambabu, K., 'Design of Ultra-Compact Uwb Antenna with Band-Notched Characteristics for MIMO Applications', IET Microwaves, Antennas & Propagation, 2018, 12, (12), pp. 1895-1900.
17. Saxena, S., Kanaujia, B.K., Dwari, S., Kumar, S., and Tiwari, R., 'MIMO Antenna with Built-in Circular Shaped Isolator for Sub-6 GHz 5G Applications', Electronics Letters, 2018, 54, (8), pp. 478-480.
18. Iqbal, A., Saraereh, O.A., Ahmad, A.W., and Bashir, S., 'Mutual Coupling Reduction Using F-Shaped Stubs in Uwb-Mimo Antenna', IEEE Access, 2018, 6, pp. 2755-2759.
19. Singh, H.S., Pandey, G.K., Bharti, P.K., and Meshram, M.K., 'A Compact Dual-Band Diversity Antenna for WLAN Applications with High Isolation', Microwave and Optical Technology Letters, 2015, 57, (4), pp. 906-912.
20. Yang, X., Liu, Y., Xu, Y.-X., and Gong, S.X., 'Isolation Enhancement in Patch Antenna Array with Fractal Uc-Ebg Structure and Cross Slot', IEEE Antennas and Wireless Propagation Letters, 2017, 16, pp. 2175-2178.
21. Li, M.-Y., Xu, Z.-Q., Ban, Y.-L., Sim, C.-Y.-D., and Yu, Z.-F., 'Eight-Port Orthogonally Dual-Polarised MIMO Antennas Using Loop Structures for 5G Smartphone', IET Microwaves, Antennas & Propagation, 2017, 11, (12), pp. 1810-1816.
22. Morsy, M.M. and Morsy, A.M., 'Dual-Band Meander-Line MIMO Antenna with High Diversity for LTE/UMTS Router', IET Microwaves, Antennas & Propagation, 2018, 12, (3), pp. 395-399.
23. Chouhan, S., Panda, D.K., Kushwah, V.S., and Mishra, P.K., 'Octagonal-Shaped Wideband MIMO Antenna for Human Interface Device and S-Band Application', International Journal of Microwave and Wireless Technologies, 2018, pp. 1-10.
24. Toktas, A., 'G-Shaped Band-Notched Ultra-Wideband Mimo Antenna System for Mobile Terminals', IET Microwaves, Antennas & Propagation, 2017, 11, (5), pp. 718-725.
25. Katie, M.O., Jamlos, M.F., Mohsen Alqadami, A.S., and Jamlos, M.A., 'Isolation Enhancement of Compact Dual-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited.
Content may change prior to final publication in an issue of the journal. To cite the paper please use the doi provided on the Digital Library page.

Wideband MIMO Antenna Using Flag-Shaped Stub', Microwave and Optical Technology Letters, 2017, 59, (5), pp. 1028-1032.
26. Jha, K.R., Bukhari, B., Singh, C., Mishra, G., and Sharma, S.K., 'Compact Planar Multistandard MIMO Antenna for IoT Applications', IEEE Transactions on Antennas and Propagation, 2018, 66, (7), pp. 3327-3336.
27. Hasan, M.M., Faruque, M.R.I., and Islam, M.T., 'Dual Band Metamaterial Antenna for LTE/Bluetooth/WiMAX System', Sci Rep, 2018, 8, (1), pp. 1240.
28. Pan, Y., Cui, Y., Qi, C.L., and Li, R., 'Evaluation of Dual-Polarised Triple-Band Multi-Beam MIMO Antennas for WLAN/WiMAX Applications', IET Microwaves, Antennas & Propagation, 2017, 11, (10), pp. 1469-1475.
29. Wang, S.-M., Hwang, L.-T., Chang, F.-S., Lee, C.-J., and Hsu, C.-Y, 'MIMO Antenna Design with Built-in Decoupling Mechanism for WLAN Dual-Band Applications', Electronics Letters, 2015, 51, (13), pp. 966-968.
30. Wong, K.-L., Chang, H.-J., and Li, W.-Y., 'Integrated Triple-Wideband Triple-Inverted-F Antenna Covering 617-960/1710-2690/3300-4200 MHz for 4G/5G Communications in the Smartphone', Microwave and Optical Technology Letters, 2018, 60, (9), pp. 2091-2096.
31. Han, Z.-J., Song, W., and Sheng, X.-Q., 'Gain Enhancement and Rcs Reduction for Patch Antenna by Using Polarization-Dependent Ebg Surface', IEEE Antennas and Wireless Propagation Letters, 2017, 16, pp. 1631-1634.
32. Ashyap, A.Y.I., Zainal Abidin, Z., Dahlan, S.H., Majid, H.A., Shah, S.M., Kamarudin, M.R., and Alomainy, A., 'Compact and Low-Profile Textile EBG-Based Antenna for Wearable Medical Applications', IEEE Antennas and Wireless Propagation Letters, 2017, 16, pp. 2550-2553.
33. Abed, A.T., 'Highly Compact Size Serpentine-Shaped Multiple-Input–Multiple-Output Fractal Antenna with Cp Diversity', IET Microwaves, Antennas & Propagation, 2018, 12, (4), pp. 636-640.
34. S.B. Vignesh, Nasimuddin, A. Alphones, "Circularly polarized strips integrated microstrip antenna for energy harvesting applications", Microw. Opt. Technol. Lett., 2016, 58, (5), pp. 1044-1049
35. Z.N.C. Nasimuddin, X. Qing, 'Asymmetric-circular shaped slotted microstrip antennas for circular polarization and RFID application', IEEE Trans. Antennas Propagation, 2010, 58, (12), pp. 3821-3828
36. Awasthi Y K, Singh H, Sharma M, Kumari S, Verma A. K. Computer-aided design based circuit model of microstrip line for terahertz interconnects technology. The Journal of Engineering, 2017, (9), pp. 512-526.
37. Abd-Alhameed, Read A., M. Mangoud, Peter S. Excell, and K. Khalil. 'Investigations of polarization purity and specific absorption rate for two dual-band antennas for satellite-mobile handsets', IEEE Transactions on Antennas and Propagation "2005,53, (6) ,pp. 2108-2110.
38. Iyama, Takahiro, Teruo Onishi, Yoshiaki Tarusawa, Shinji Uebayashi, and Toshio Nojima. "Novel specific absorption rate (SAR) measurement method using a flat solid phantom." IEEE Transactions on Electromagnetic Compatibility,2008 ,50, (1),pp. 43-51.
39. Loyka, S.L., 'Channel capacity of MIMO architecture using the exponential correlation matrix', IEEE Communications letters, 2001, 5, (9), pp.369-371.
40. Rajkumar, S., Sivaraman, N.V., Murali, S. and Selvan, K.T, 'Heptaband swastik arm antenna for MIMO applications', IET Microwaves, Antennas & Propagation, 2017,11, (9), pp.1255-1261.
41. Biswal, S.P. and Das, S, 'A low-profile dual port UWB-MIMO/diversity antenna with band rejectionability', International Journal of RF and Microwave Computer-Aided Engineering, 2018,28, (1), pp.21159.
42. Ang, K. S., Lee, C. H. and Leong, Y. C. 'A broad-band quarter-wavelength impedance transformer with three reflection zeros within passband', IEEE Transaction on Microwave Theory Techniques, 2004, 52, (12), pp. 2640-2644.
43. Row, J.S. 'The design of a squarer-ring slot antenna for circular polarization', IEEE Transactions on Antennas and Propagation, 2005, 53, (6), pp. 1967-1972.
44. Jensen, T., Zhurbenko, V., Krozer, V. and Meincke, P., 'Coupled transmission lines as impedance transformer", IEEE Transaction on Microwave Theory Techniques, 2007, 55, (12), pp. 2957-2965.

View publication stats

Journal of ELECTRONIC MATERIALS
https://doi.org/10.1007/s11664-019-07700-x
© 2019 The Minerals, Metals & Materials Society

Impedance Spectroscopy and Conduction Behavior in $CoFe_2O_4$-$BaTiO_3$ Composites

S. SHANKAR,[1,2] O.P. THAKUR,[3,4] and M. JAYASIMHADRI[2,5]

1.—Experimental Research Laboratory, Department of Physics, ARSD College, University of Delhi, New Delhi 110021, India. 2.—Luminiscent Materials Research Lab, Department of Applied Physics, Delhi Technological University, Delhi 110042, India. 3.—Materials Analysis and Research Laboratory, Department of Physics, NSIT, New Delhi 110078, India. 4.—e-mail: opthakur@yahoo.com. 5.—e-mail: jayaphysics@yahoo.com

Magnetoelectric composites combining ferromagnetic cobalt ferrite (CFO) and ferroelectric barium titanate (BT) are prepared using conventional solid-state reaction. Dielectric measurements reveal that CFO and CFO-BT composites exhibit Maxwell–Wagner polarization and temperature-dependent relaxation originating due to the presence of ions/defects vacancies. Prominently, CFO-30%BT composite showed an unexpected low dielectric loss ≈ 0.5 above 1 kHz and has potential for device application. Impedance spectroscopy confirms the thermally activated relaxation mechanism and high impedance behavior in the samples. Modulus analysis affirms the statistically distributed non-Debye-type behavior existing in CFO-BT composites. The electric conductivity takes place via correlated barrier hopping in CFO and CFO-10BT composite and overlapping large polaron tunneling in CFO-30BT composite. The addition of BT in CFO initiates the structural modification and results in conductivity cross-over with improved conductivity. These results are useful for exploring devices based on CFO-BT composites.

Key words: Dielectric, impedance, conductivity, modulus

INTRODUCTION

Magnetoelectric (ME) composites are promising candidates for transducers, sensing and phase shifting devices owing to their product property.[1–4] These composites are magnetically polarized on application of an electric field and *vice versa* as they contain both electric and magnetic dipoles.[2] Moreover, ME composites offer an additional degree of coupling arising from mechanical coupling between the electric and magnetic phases.[4] This particular feature has attracted much research attention around the globe. The extent of coupling can be modified by tuning the connectivity and orientations between these phases and thereby rendering ME materials useful for scientific and technological devices.[5]

(Received April 23, 2019; accepted September 28, 2019)

Spinel cobalt ferrite ($CoFe_2O_4$) is a vastly studied material because of its remarkable magnetostriction property, which is promising for application in the field of electronics, electrochemical science and magnetic devices.[6–9] Cobalt ferrite exhibits a physically and chemically stable structure with notable electrical properties making it suitable for electronic and magnetic devices. The magnetism of cobalt ferrite is grain size and temperature dependent.[10] Recently, substantial research has been made to enhance the structure and electrical properties of CFO-based systems. Significant efforts have been dedicated to partially substituting Fe^{3+} ions,[11,12] composites with ferroelectric/piezoelectric materials,[13–15] nanoparticles[16] and core/shell structures.[17] Among these heterostructural systems, CFO-based composites combining the ferromagnetic CFO with ferroelectric phase stand out distinctively because of their outstanding ME coupling. Interestingly, CFO and barium titanate ($BaTiO_3$) are observed to give promising results for

Published online: 21 October 2019

magnetoelectricity in both bulk and nanocomposite systems.[18–20] The incorporation of pervoksite $BaTiO_3$ in spinel CFO is observed to bring structural modification and lattice disorder,[21] which may effect the electrical behavior and impedance properties in CFO-BT composites. Complex impedance spectroscopy (CIS) is a very elegant technique used to understand the electrical behavior by separating the grain impedance from impedances of other sources like grain boundary and electrode interface.[22] This measurement provides information regarding distribution of relaxation processes over a frequency range.

In our previous work, we have reported the structural, ferroelectric, ferromagnetic and enhanced magnetoelectric behavior in CFO-BT composites.[21] The presence of cubic as well as tetragonal structures in CFO-BT composites that display a magnetoelectric coupling coefficient of 22 mV/cm/Oe has been already established.[21] In this work we report the comprehensive investigation on the electrical behavior and conductivity phenomena using impedance spectroscopy of the CFO-BT composites as a function of BT content. In the subsequent sections the experimental methodology, microstructural analysis and electrical properties of CFO-BT composites are presented. This work will aid the understanding of electrical behavior and will be useful in exploring the application of CFO-BT composites in high-frequency applications with suitable tuning for device formation.

EXPERIMENTAL

$CoFe_2O_4$-$BaTiO_3$ (CFO-BT) composites with composition CFO, $CoFe_2O_4$-10%$BaTiO_3$ (CFO-10BT) and $CoFe_2O_4$-30%$BaTiO_3$ (CFO-30%BT) have been prepared from bulk CFO and BT using conventional solid-state reaction and have been labeled as CFO, CFO-10BT and CFO-30BT.[21] Surface morphological studies are carried out using scanning electron microscopy (SEM). The electrical measurements were made using a Nova Control Impedance Analyzer with frequency ranging from 100 Hz to 1 MHz and over temperatures 100–400°C.

RESULTS AND DISCUSSION

Microstructure and Dielectric Studies

Figure 1a, b, and c shows the SEM images of CFO, CFO-10BT and CFO-30BT composites, respectively. The SEM images display a uniform, homogenously dispersed structure with more or less narrow grain size distribution. In addition, some amount of coagulation can be observed in the SEM pictures

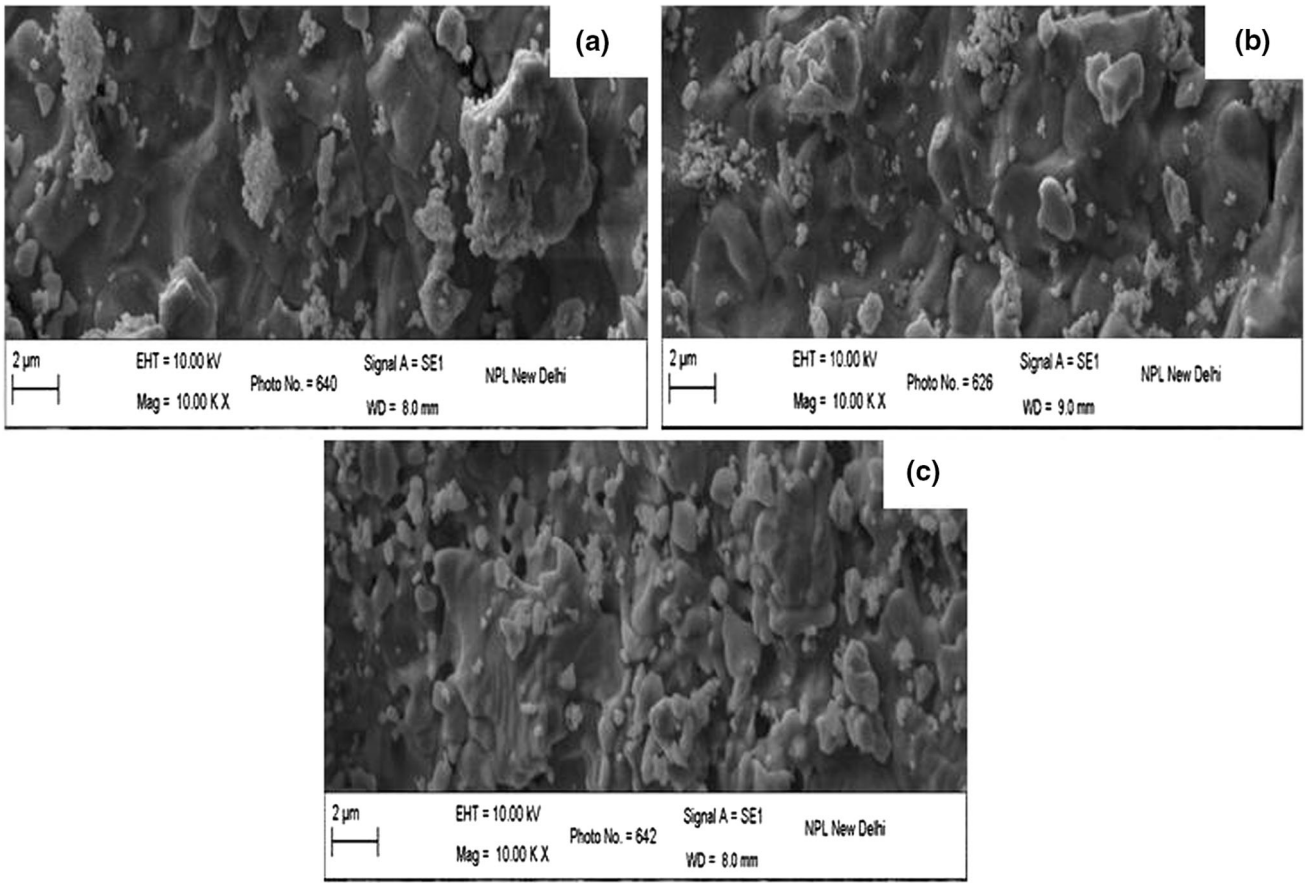


Fig. 1. SEM micrographs of (a) CFO, (b) CFO-10BT, and (c) CFO-30BT composites.

due to strong magnetism in CFO. It can also be observed from the SEM images that all the composites are tightly bound with low porosity and have an average grain size $\approx (5 \pm 1)\ \mu$m.

The variation of dielectric constant (ε') with frequency for CFO, CFO-10BT and CFO-30BT composites at different fixed temperatures is shown in Fig. 2a, b, and c, respectively. The composites of CFO-BT display a decrease in dielectric permittivity with frequency arising from the presence of a conventional dielectric relaxation process and following Maxwell–Wagner polarization. The decrease is mainly because the electric dipoles cannot adjust to the change in frequency of applied electric field.[23] It can be observed from Fig. 2 that the permittivity increases with temperature and BT content in CFO. The dispersion in permittivity in the low and high frequency regions confirms the presence of space charge polarization.[23] The composites of CFO-BT display higher permittivity than pure CFO. The dielectric constant increases from 450 to 700 at 100 Hz for CFO and CFO-30BT composite, respectively. Figure 3a, b, and c displays change of loss (tan δ) with frequency for CFO, CFO-10BT and CFO-30BT composites, respectively. The high dielectric loss in CFO is attributed to hopping of divalent charge carrier Fe^{2+}/Fe^{3+} ions. This prominent feature decides the dielectric parameters in CFO-BT composites despite the strong dielectric nature of BT.[24] It is surprising to find that the dielectric loss decreases prominently from 10 to less than unity for CFO and CFO-30BT composites, respectively. The minimum loss obtained is 0.5 for CFO-30 BT composite at 1000 Hz which remains more or less constant over the experimentally studied frequency range.

Impedance and Modulus Spectroscopy

Complex impedance spectroscopy is a useful technique to understand electrical behavior and electrical relaxation processes by dividing the grain–grain boundary contributions and other sources. The Nyquist Cole–Cole plots of CFO, CFO-10BT and CFO-30BT composites, respectively, are displayed in Fig. 4. All Nyquist plots start from the origin at 100°C suggesting the dominant grain

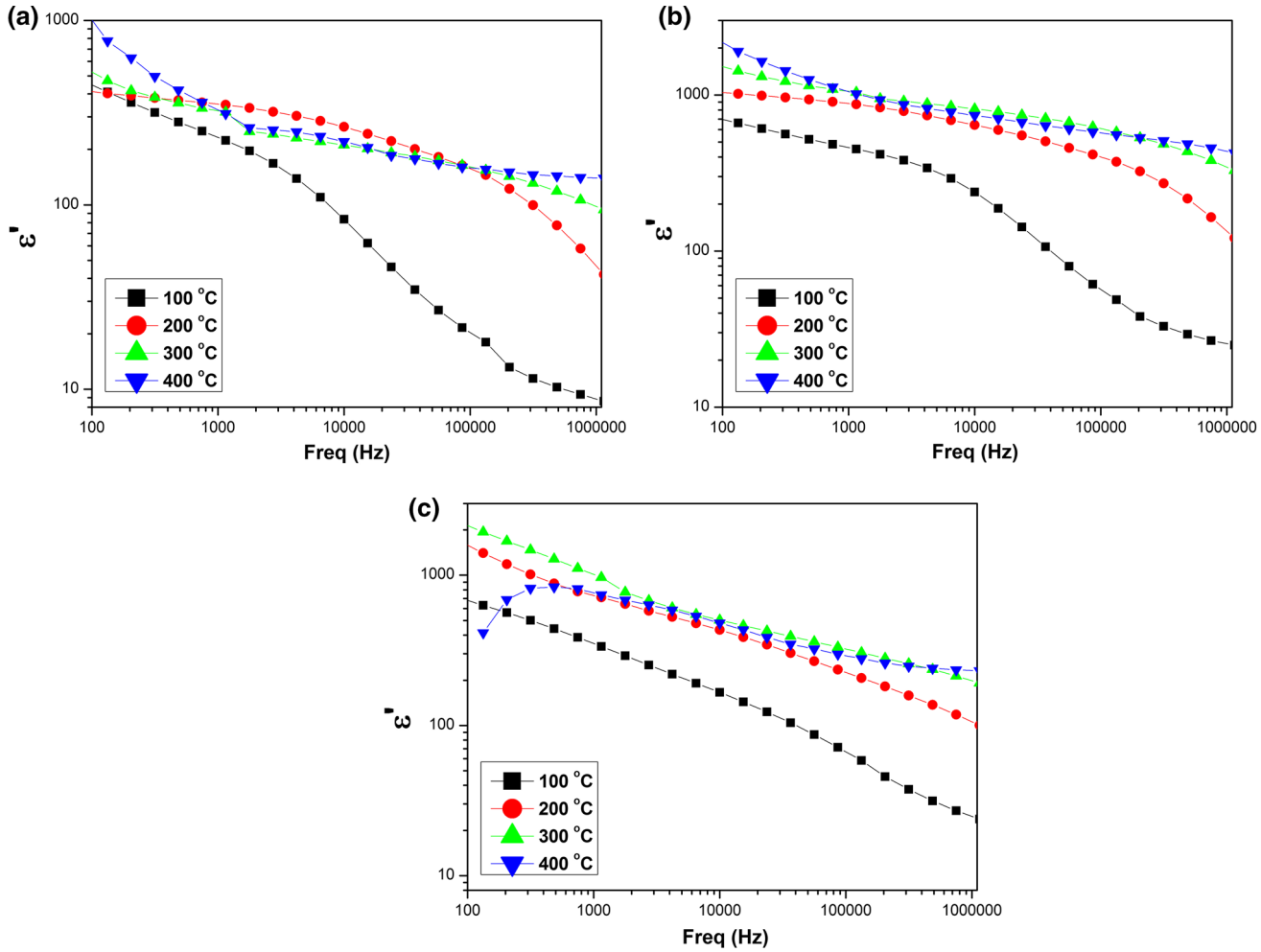


Fig. 2. Frequency dependence of real part of dielectric constant (ε') for (a) CFO, (b) CFO-10BT, and (c) CFO-30BT composites.

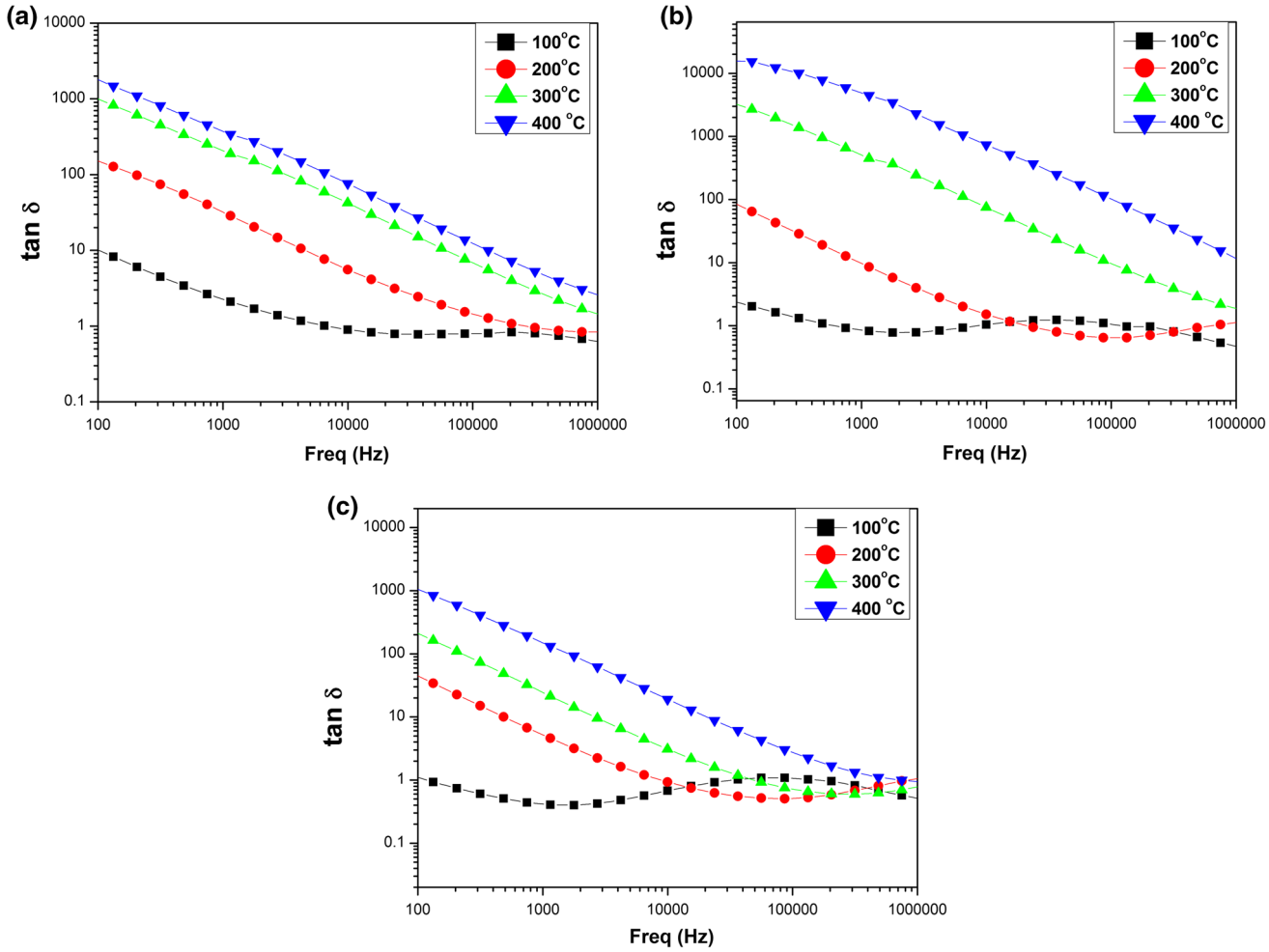


Fig. 3. Frequency dependence of dielectric loss (tan δ) for (a) CFO, (b) CFO-10BT, and (c) CFO-30BT composites.

contributions. The increase in the temperature from 100°C to 200°C sets significant contributions of grains and grain boundaries in all the compositions. As the temperature reaches 400°C, the samples CFO and CFO-BT composites tend to exhibit dominant contributions from the grain boundaries along with a small extension in the low frequency region. The radius of the semicircle increases on addition of BT in CFO up to 10% due to the dielectric nature of BT. However, the CFO-30BT composite displays a relative decrease in the Nyquist semicircle radius as compared to the CFO-10BT composite which could be possibly due to extensive hopping of Fe ions. Furthermore the radius of the semicircles in Nyquist plots decrease invariably with temperature for CFO and CFO-BT composites owing to the presence of statistically distributed relaxation processes.[25,26] The centers of semicircles in Nyquist plots for CFO and CFO-BT composites lie below the Z' axis and suggest a typical non-Debye behavior in CFO-based materials.[22] Moreover, based on the assumption that the semicircle of the Nyquist plot is composed of combinations of RC circuits connected in parallel and using the frequency equation $f = 1/2\pi RC$, the relaxation time (τ) can be estimated as the reciprocal of frequency in the maxima of Nyquist plots.[27]

The non-Debye behavior and grains/grain boundary contributions present in CFO and CFO-BT composites can be assessed by modeling the Nyquist plots as R–C equivalent circuits respectively employing altered brick layer model.[28–30] The equivalent electric circuit may consist of an arrangement of more than one circuit, comprising a resistor and/or CPE and/or capacitor connected in parallel, representing the contributions arising

from the grains, grain boundaries and interfaces. The Nyquist plots of CFO and CFO-BT composites are fitted using Zview software, and two equivalent circuits containing resistance and a constant phase element are found to be suitable in following the plots. Figure 5a, b, c, d, e, f, g, h, and i shows correspondence between the Nyquist plots and fitted data indicated by the black circle and solid red line, respectively, for CFO (Fig. 5a, b, and c), CFO-10BT (Fig. 5d, e, and f) and CFO-30BT (Fig. 5g, h, and i) composites, respectively. The parameters obtained from fitting of Nyquist plots are shown in Table I. The experimental results and fitted graphs as seen in Fig. 5 are in consonance with each other. The occurrence of CPE combinational electric circuit governed by equation $Z_{CPE} = 1/(j\omega)^n$CPE confirms the non-Debye behavior exhibited by CFO and CFO-BT composites.

The change in impedance Z' with frequency for CFO, CFO-10BT and CFO-30BT composites, respectively, at fixed temperatures is shown in Fig. 6a, b, and c. The plots in Fig. 6 show that Z' is large at low frequencies region and decreases continually with frequency for the studied temperature range. It can be observed that Z' coincides at higher temperatures and higher frequencies which is possibly due to space charge release originating because of reduction in hopping potential barrier in CFO and CFO-BT composites. With the increase in BT content and temperature Z' decreases confirming the thermal dependence in dispersion of conductivity existing in CFO-BT composites. This is dominantly attributed to the decrease in resistance of

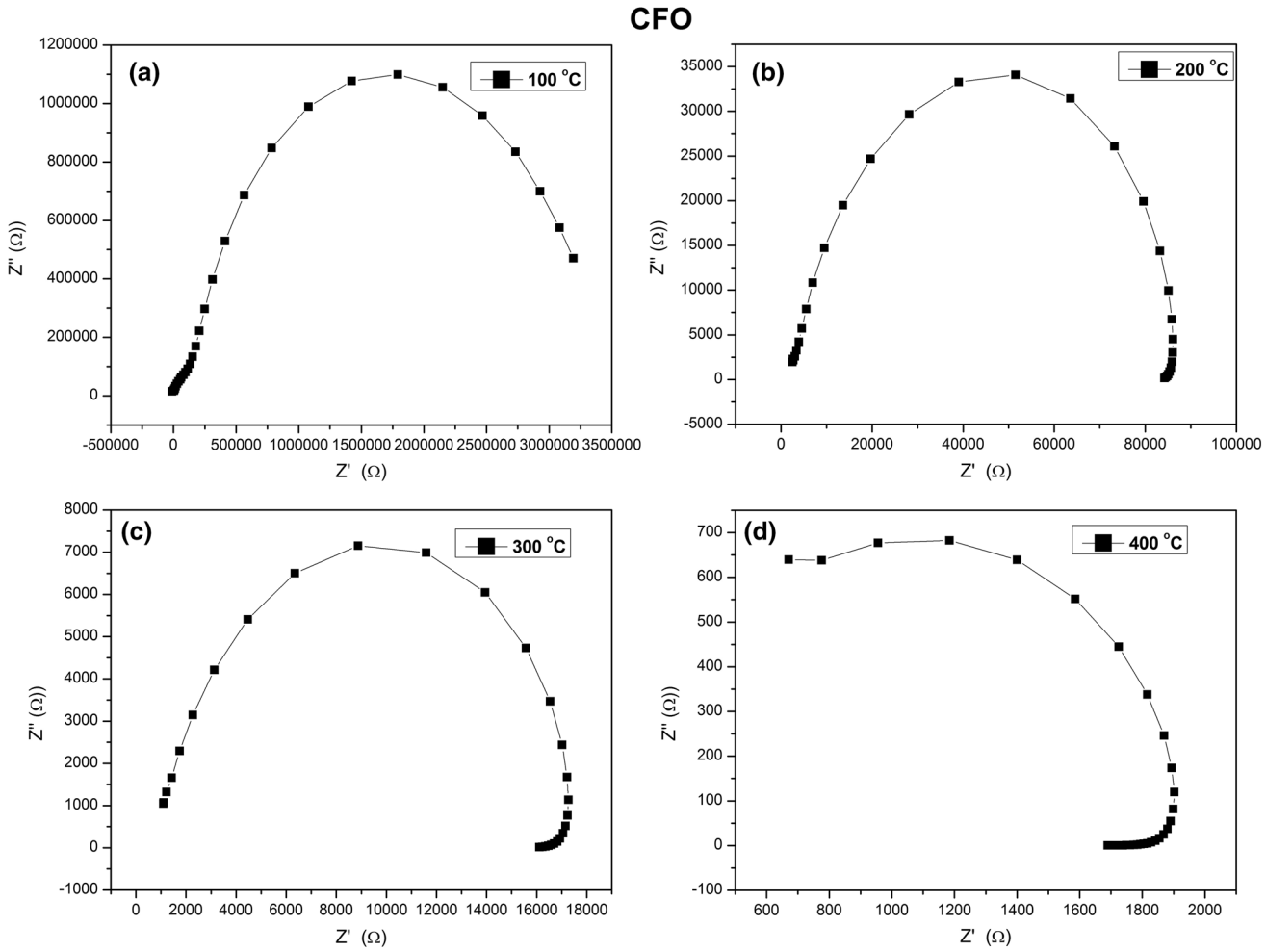


Fig. 4. Nyquist plot of CFO-BT composites measured at different temperatures.

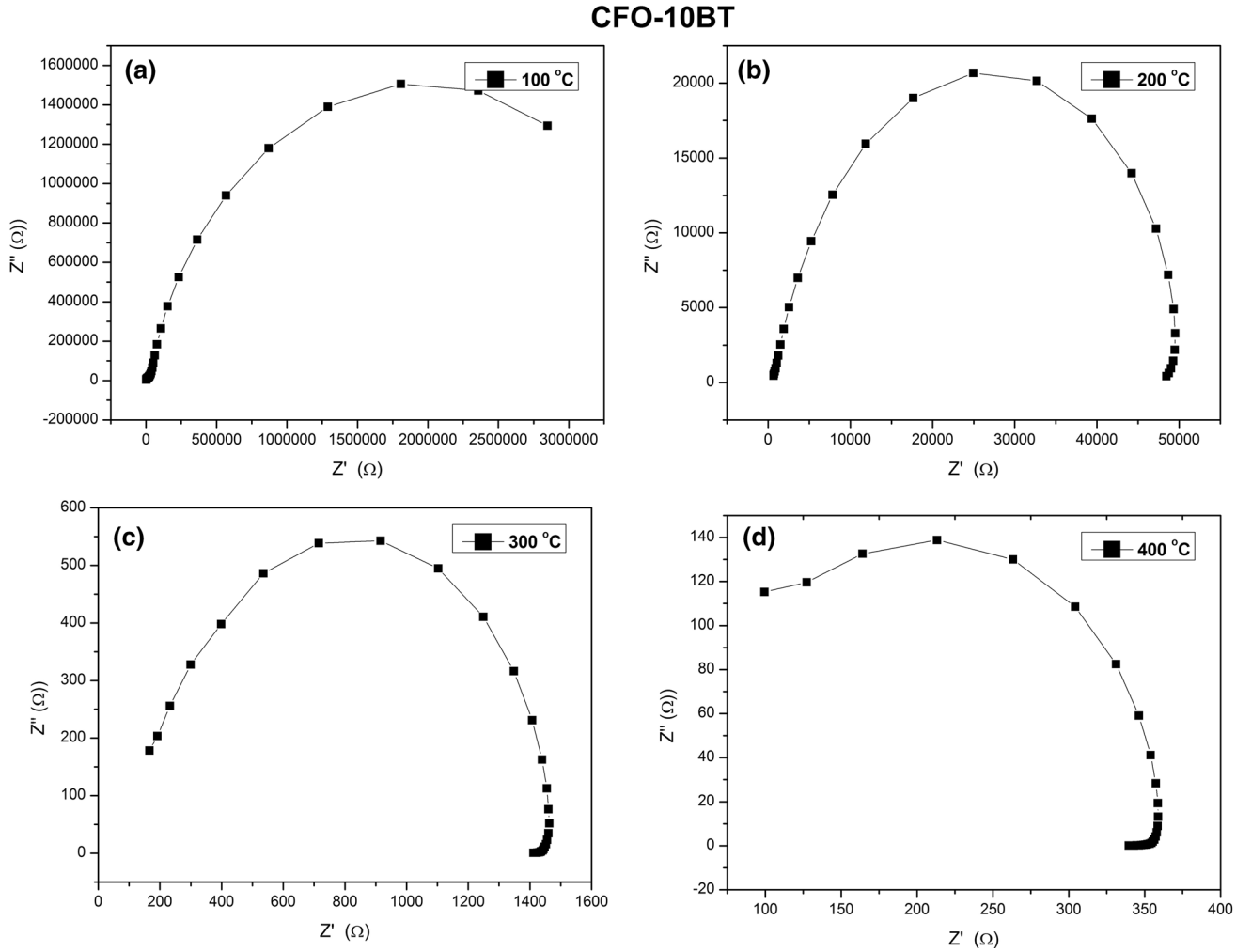


Fig. 4. continued.

grain/grain boundaries.[21,31] Figure 6 also shows the variation in complex part of impedance Z'' with frequency for CFO, CFO-10BT and CFO-30BT composites, respectively. Variation of complex Z'' is essential in examining presence of impedance phase in composites. The graphs of Fig. 6 attain peaks at all measured temperatures for all the compositions. The maxima of Z'' (Z''_{max}) drifts to the high-frequency regime, indicating the existence of thermally activated relaxation processes in CFO and CFO-BT composites arising from thermal agitation of charges present in the grains and at the grain boundaries.[31] The Z'' peaks exhibit thermal decrement with broadening which confirms the thermally activated relaxation mechanism related to the existence of static charges and defects/vacancies at low and high temperatures, respectively.

The electrical response can be analyzed using the modulus spectroscopy analysis. Figure 7a, b, and c depicts the frequency dependency of real part of modulus (M') for CFO, CFO-10BT and CFO-30BT composites, respectively. Figure 7 shows that M' is negligible at low frequencies and becomes saturated at high frequencies owing from the movement of short range charge carriers. As the amount of BT is increased in CFO, a decrease in the saturated value is observed due to the localization of charge carriers. The variation of complex modulus (M'') with

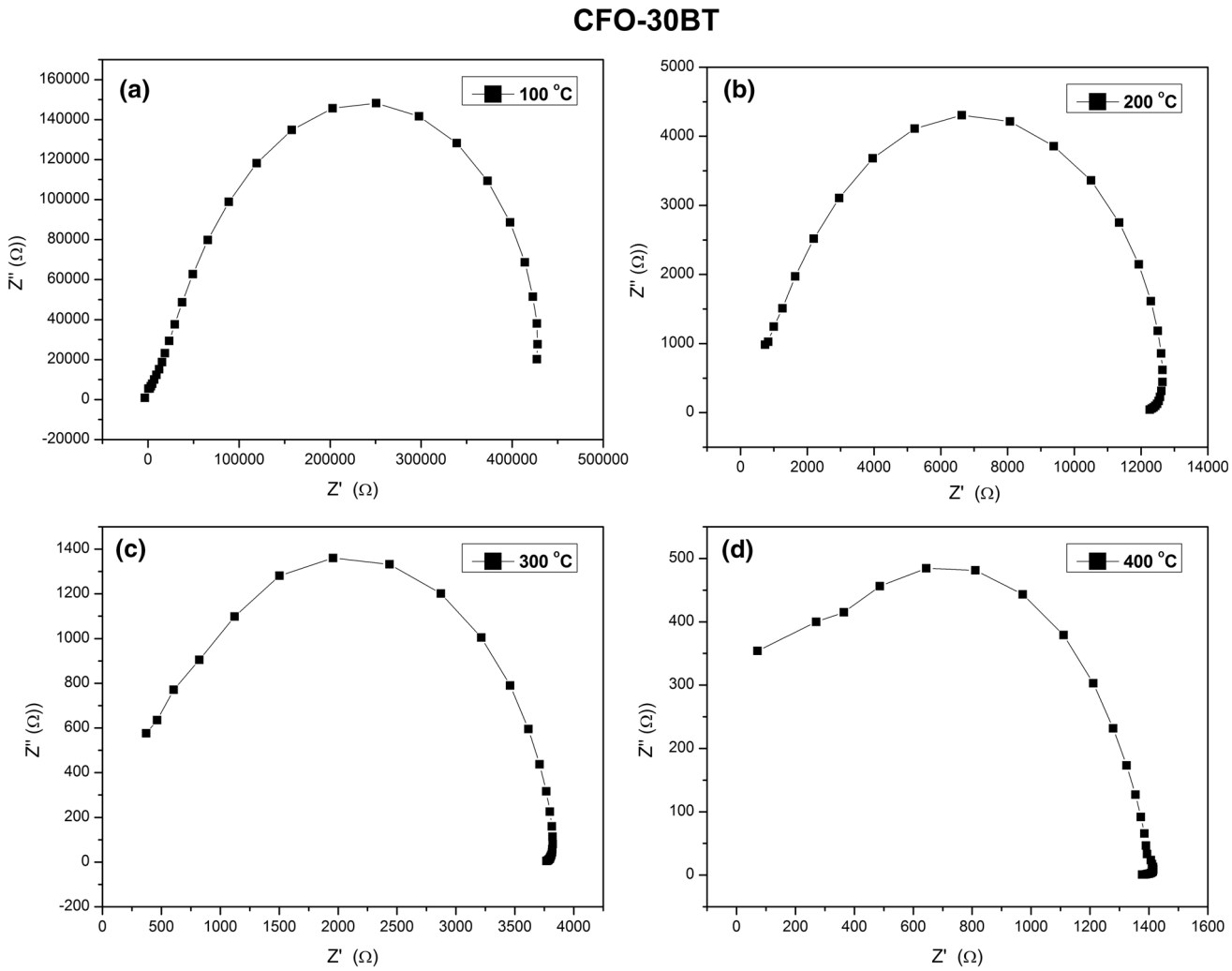


Fig. 4. continued.

frequency is also shown in Fig. 7. It can be seen that all samples display a maxima at 100°C which broadens and shifts to high frequency with the increase in temperature and confirms the presence of temperature-dependent conductivity behavior in CFO and CFO-BT composites. All the peaks are asymmetric in nature owing to the presence of a multi-relaxation process and usual non-Debye nature of CFO based samples.

Electrical Properties

Variation of the real part of conductivity (σ') for CFO, CFO-10BT and CFO-30BT composites, respectively with temperature as well as frequency is shown in Fig. 8a, b, and c. At low frequencies up to 1 kHz, the conductivity spectrum shows negligible variation and is frequency independent. This contribution originates from the dc conductivity and is attributed to the movement of charge around grain boundaries. The data show dispersion at higher frequencies for CFO and CFO- BT composites, suggesting the onset of ac conductivity. This dispersion arises from increased mobility in charges. For temperatures except 100°C, the spectra tend to merge at the high-frequency regime which is indicative of structural disordering. At higher temperatures, the frequency dispersion decreases radically due to dominant dc conductivity part. It can also be observed from the data that conductivity increases with BT content as a result of an increase in the conduction path due to structural distortion in CFO-BT composites.[20] Moreover, all the composites display a thermally activated conduction mechanism.

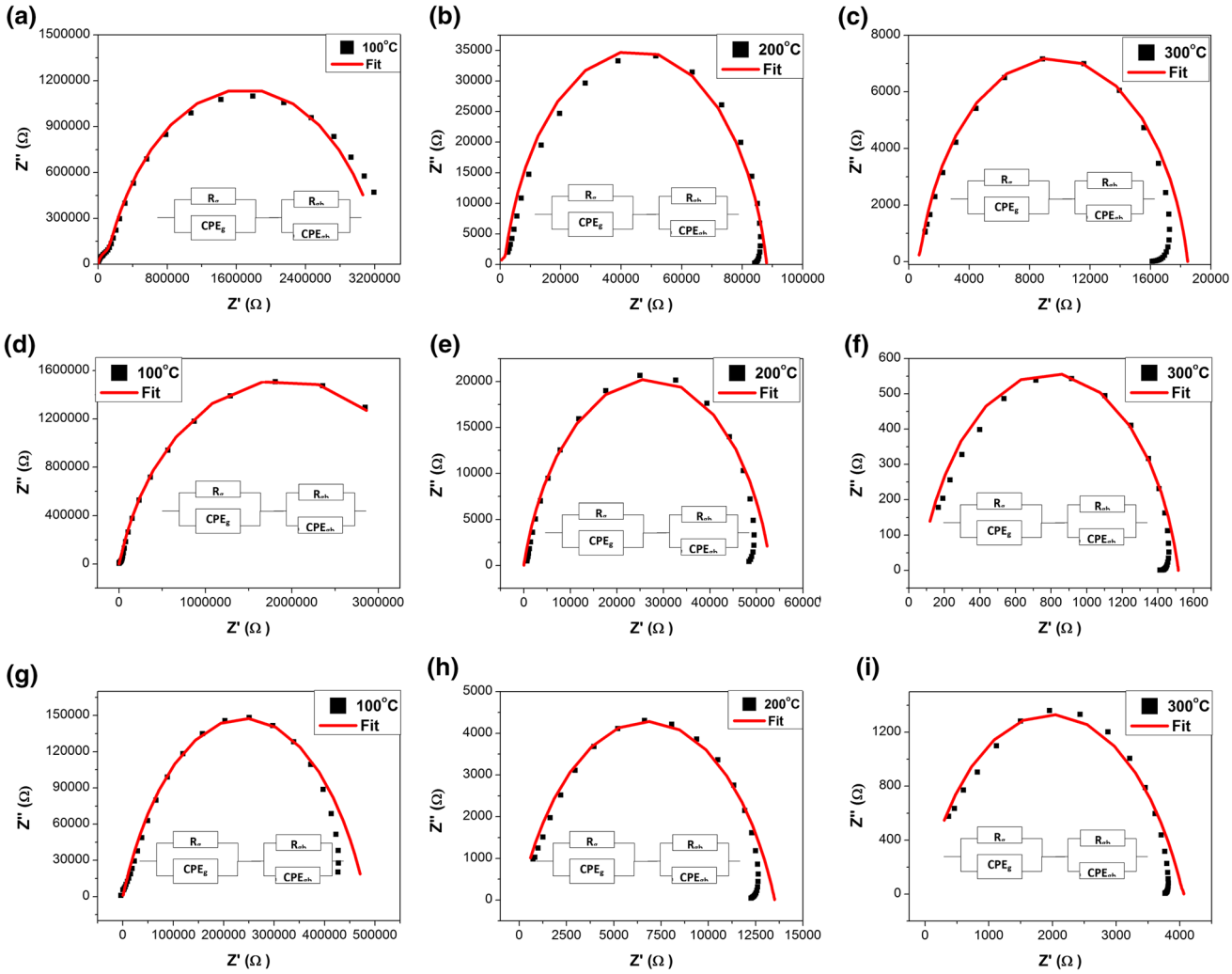


Fig. 5. Combined Nyquist plots of (a–c) CFO, (d–f) CFO-10BT, and (g–i) CFO-30BT composites respectively at different temperatures and their equivalent circuit measured.

Table I. Fitted parameters for Cole–Cole plots of CFO-BT composites

Temp (°C)	R_g (Ω)	CPE_g	n_g	R_{gb} (Ω)	CPE_{gb}	n_{gb}
x = 0						
100	3,200,000	0.0000000125	0.79	100,000	0.000000002	0.86
200	86,600	0.0000000069	0.865	1500	0.0000000009	0.9
300	17,660	0.0000000075	0.87	800	0.0000006	0.4
x = 10						
100	3,000,000	0.000000002	0.9	800,000	0.00000002	0.7
200	9500	0.00000002	0.7	43,500	0.00000002	0.80
300	80	0.000000008	0.7	1435	0.000000002	0.84
x = 30						
100	480,000	0.000000003	0.7	1300	0.0000002	0.9
200	13,500	0.000000002	0.72	13	0.000000002	0.5
300	4050	0.000000004	0.7	13	0.00002	0.99

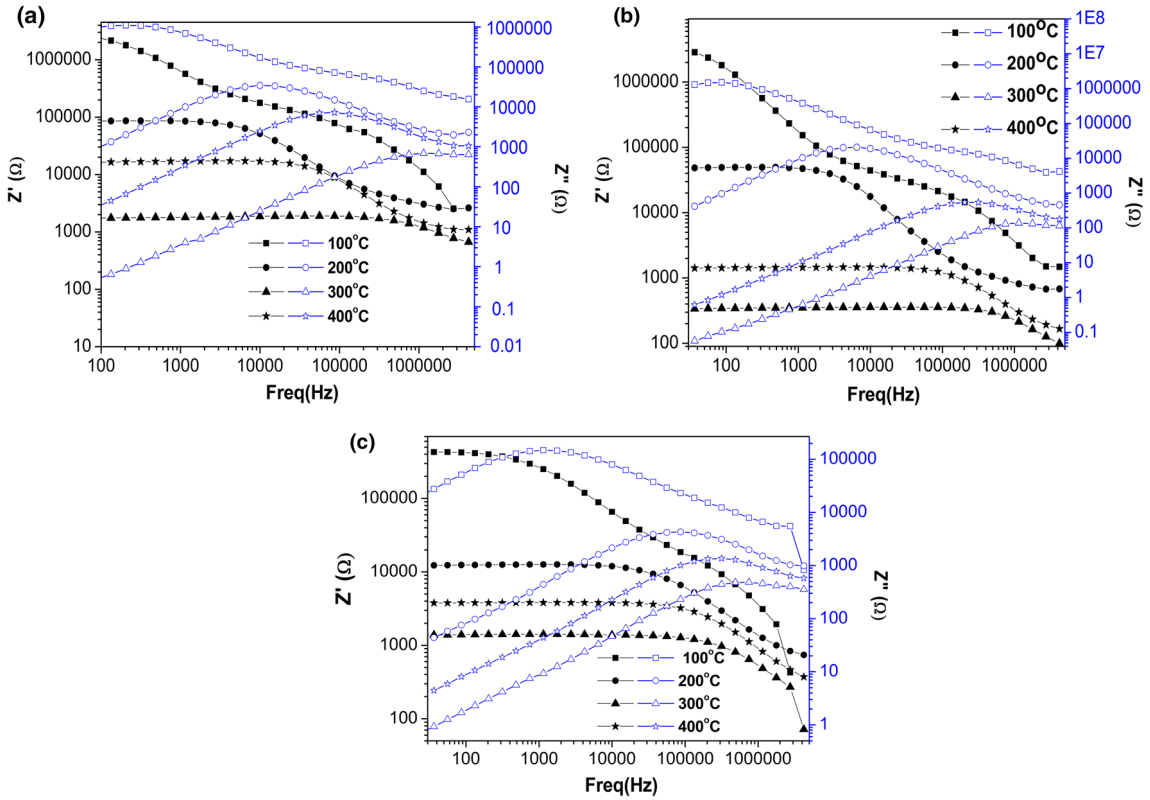


Fig. 6. Frequency dependence of real part of impedance (Z') for (a) CFO, (b) CFO-10BT, and (c) CFO-30BT composites. Inset shows frequency dependence of complex part of impedance (Z'') for (a) CFO, (b) CFO-10BT, and (c) CFO-30BT composites.

It can be assumed that whenever the conductivity increases and saturates at high frequencies, then a single particle is hopping around in a double well with infinite barriers.[32,33]

The conductivity behavior in CFO, CFO-10BT and CFO-30BT can be understood by separating the dc and ac conductivity from the real part of conductivity (σ') as:

$$\sigma' = \sigma_{dc} + \sigma_{ac} \tag{1}$$

The thermal variation of σ_{dc}—dc conductivity can be exploited to determine the activation energy (E_a) in CFO and CFO-BT composites and is governed by the Arrhenius equation as

$$\sigma_{dc} = \sigma_o e^{-E_a/k_B T} \tag{2}$$

where the terms σ_o, k_B, T represent usual notations.[12] The thermal dependence of σ_{dc} and its linear fit is represented in Fig. 9. On the studied temperatures separate slope values and activation energies have been obtained for CFO, CFO-10BT and CFO-30BT composites, respectively, and are tabulated in Table II. The activation energies (E_a) obtained are found to be less than one in case of CFO and CFO-BT composites and affirms that thermal activated conductivity is controlled by singly ionized oxygen.[34–36] These oxygen ions/defect vacancies are produced during the synthesis at high temperatures which reduces the Ti^{4+} ions to Ti^{3+} ions following the Kroger–Vink equation of defects:

$$O_o^x \rightarrow \frac{1}{2}O_2 + V_o^{**} + e^- \tag{3}$$

where V_o^{**} is the ionized metal. These defects produced affect the inter-granular impedance.

The ac conductivity term usually follows Jonscher's power law of frequency given by

$$\sigma_{ac} = A\omega^n \tag{4}$$

where A and n vary with temperature. However, for the case of CFO and CFO-BT composites, Jonscher's law alone did not yield a completely positive fit for ac conductivity. To supersede this condition,

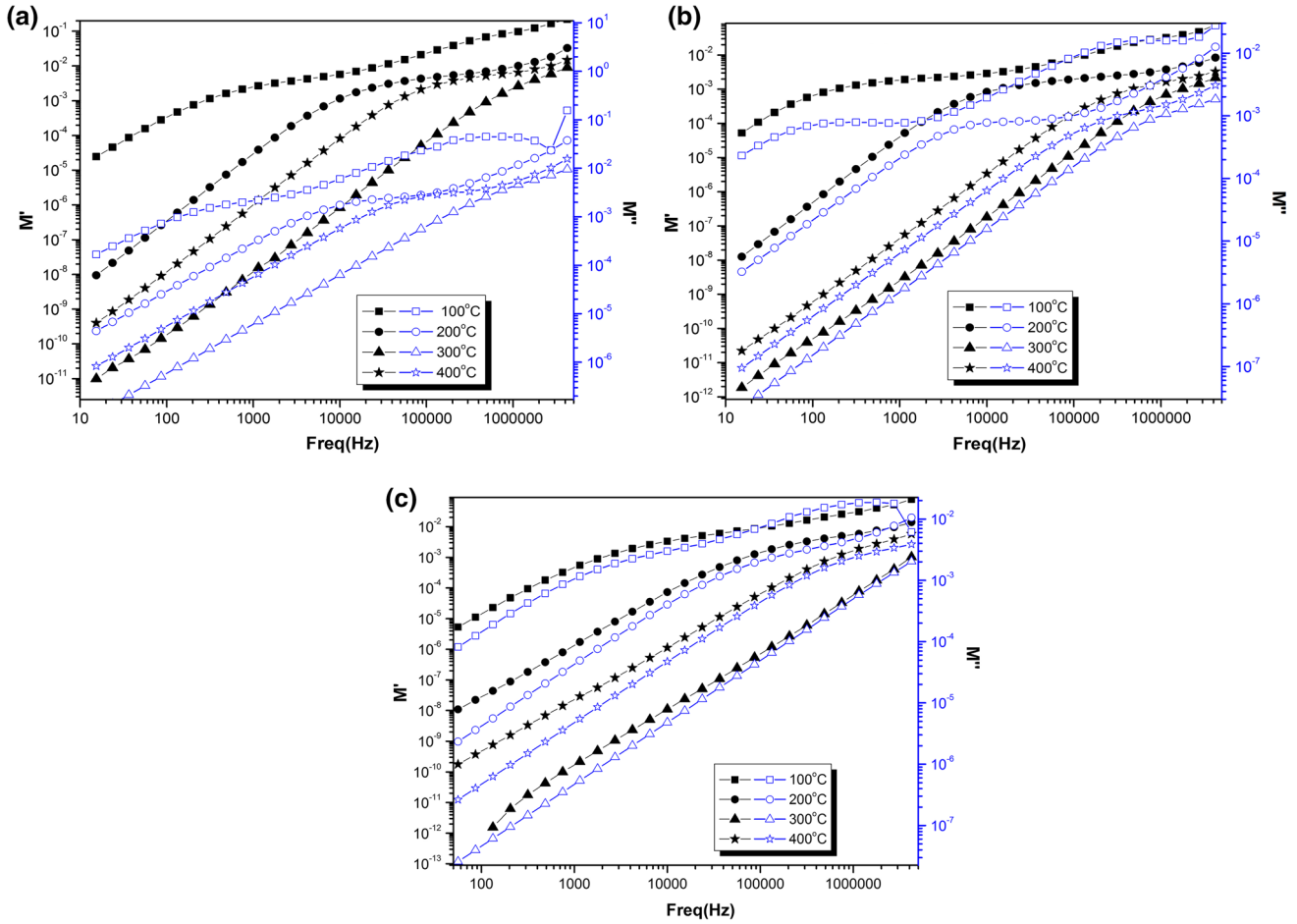


Fig. 7. Variation of M' with respect to frequency of (a) CFO, (b) CFO-10BT, and (c) CFO-30BT composites. Inset shows variation of M'' with respect to frequency of (a) CFO, (b) CFO-10BT, and (c) CFO-30BT composites.

modified Jonscher's law, consisting of multiple relaxations was used:

$$\sigma_{ac} = A\omega_1^{n1} + B\omega_2^{n2} \tag{5}$$

where the constants A, B, $n1$ and $n2$ are temperature dependent. These terms ω_1 (= ω/ω_1) and ω_2 (= ω/ω_2) denote the low and high hopping frequencies. The exponents $n1$ and $n2$ represent restricted ion hopping and two level systems.[37] Modified Jonscher's law yielded a satisfactory fit. The obtained values of $n1$ and $n2$ are plotted in Fig. 10. The conductivity phenomena present in CFO-BT composites can be understood by observing the frequency and temperature dependence of exponents n1. The temperature-independent nature up to ω_1 is due to the diffusive nature of charge carriers in uniform potentials which gets sub-diffusive above and below ω_1 and ω_2. It can be realized that the vacancies accumulated at interfaces become conducting by recovering the oxygen vacancies from interfacial traps. Jonscher's law for conductivity describes that variation in exponent n is used for understanding the type of conduction behavior present in the samples. Usually, there are three types of conductivity phenomena which exist viz., quantum mechanical tunneling (QMT), correlated barrier hopping (CBH) and overlapping large-polaron tunneling (OLPT) depending on exponent n variation. In the case when the value of exponent n is nearly 0.8, remains constant with temperature, and changes with frequency alone, then QMT behavior occurring via barrier localized sites is present. The thermal dependence of exponent n can be explained on the basis of QMT by assuming that the charge carriers form and move by means of

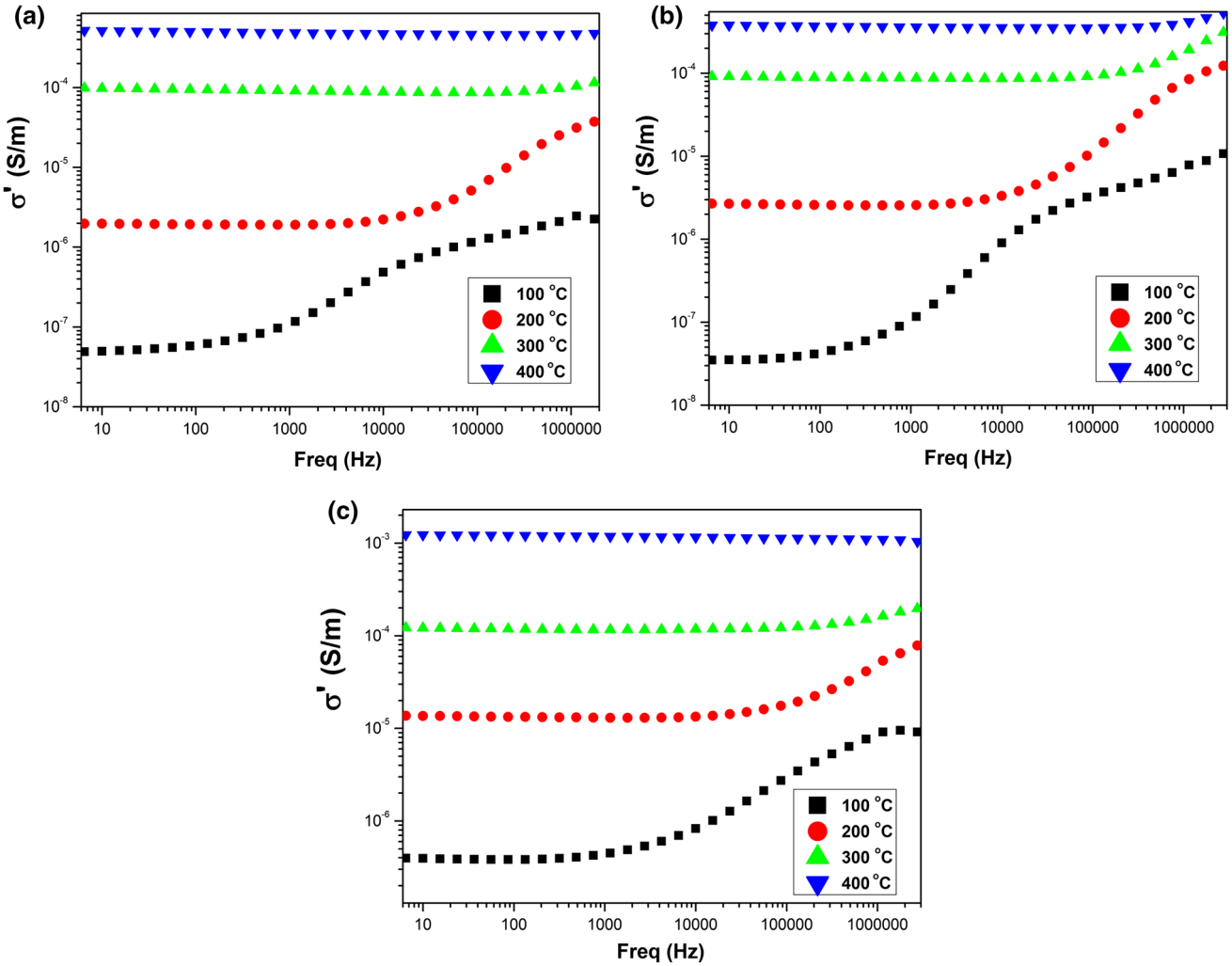


Fig. 8. Plot of conductivity (σ) as function of frequency measured at different temperatures for (a) CFO, (b) CFO-10BT, and (c) CFO-30BT composites.

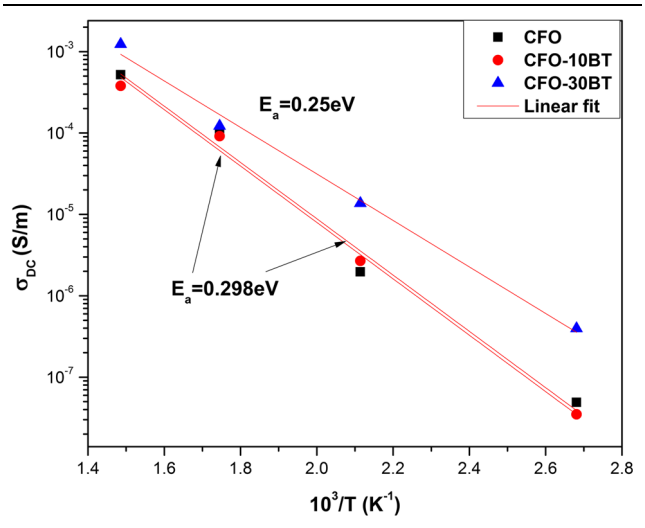


Fig. 9. Plot of dc conductivity as a function of temperature ($10^3/T$) for CFO, CFO-10BT, and CFO-30BT composites, respectively.

Table II. Values of activation energy (E_a) in (eV) obtained from fitting of dc conductivity with $1000/T$ (K^{-1})

Activation energy (eV)	CFO	CFO-10BT	CFO-30BT
E_a	0.25	0.298	0.298

non-overlapping small polaron tunneling (NSPT). As per this theory, n increases with temperature. However, a decreasing variation in exponent n with temperature, implied the CBH conduction existing in the sample. Lastly, whenever the exponent n exhibits frequency as well as thermal variation by obtaining a minima on increasing the temperature and increases thereafter implies an OLPT dominated conduction behavior.[21,38]

It is seen from Fig. 10 that exponent n displays thermally dominated dispersion and confirms the

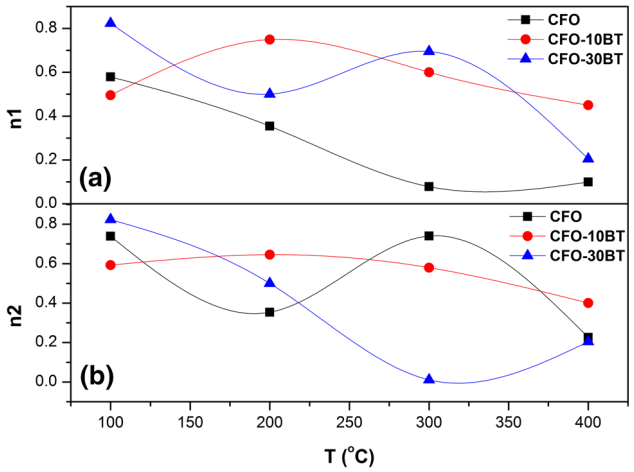


Fig. 10. Variation of exponent *n*1 and *n*2 as a function of temperature for CFO, CFO-10BT, CFO-30BT composites.

existence of two or more different conductivity phenomena in CFO-BT composites. The conduction phenomena for CFO and CFO-10BT composite display the usual CBH of charges between barriers.[39,40] Interestingly, in the case of the CFO-30BT composite, a crossover from CBH to OLPT in the conductivity behavior is observed arising from structural distortion.[41] Most importantly, the values of exponent n are below 1 and confirms the dominant translational motion.

CONCLUSIONS

Ferromagnetic-ferroelectric $CoFe_2O_4$-$BaTiO_3$ (CFO-BT) composites were prepared using conventional solid-state synthesis. SEM images display homogenously dispersed structure with average grain size of $\approx (5 \pm 1)\ \mu$m. The variation of dielectric properties with frequency display Maxwell–Wagner polarization and temperature dependent relaxation phenomena in CFO-BT composites. Electrical conduction studies exhibit a shift in conductivity behavior from CBH to OLPT mechanism from CFO to CFO-30BT composite occurring due to structural disordering. The activation energy obtained from the Arrhenius equation is below unity in CFO-BT composites and confirm presence of singly ionized ions determining conductivity phenomena. The complex impedance spectroscopy and modulus analysis confirm the presence of thermally reliant non-Debye type relaxation mechanism in the CFO-BT system. The current outcomes offer important information for sophisticated examinations in CFO-BT-based composites appropriate for high-frequency devices and other industrial applications.

ACKNOWLEDGMENTS

This work is supported by the Star Innovation Project (ARSD-SIP-01) funded by the University of Delhi, New Delhi, and USIC, University of Delhi, New Delhi, India, for characterization facilities.

REFERENCES

1. D. Khomskii, *Physics* 2, 20 (2009).
2. J.F. Scott, *Science* 315, 954 (2007).
3. S. Kumari, D.K. Pradhan, N. Ortega, K. Pradhan, C. DeVreugd, G. Srinivasan, A. Kumar, T.R. Paudel, E.Y. Tsymbal, A.M. Bumstead, J.F. Scott, and R.S. Katiyar, *Phys. Rev. B* 95, 214109 (2017).
4. W. Eerenstein, N.D. Mathur, and J.F. Scott, *Nature* 442, 759 (2006).
5. M. Kumar, S. Shankar, G.D. Dwivedi, A. Anshul, O.P. Thakur, and A.K. Ghosh, *Appl. Phys. Lett.* 106, 072903 (2015).
6. L. Ajroudi, N. Mliki, L. Bessais, V. Madigou, S. Villain, and C. Leroux, *Mater. Res. Bull.* 59, 49 (2014).
7. K. Inomata, N. Ikeda, N. Tezuka, R. Goto, S. Sugimoto, M. Wojcik, and E. Jedryka, *Sci. Technol. Adv. Mater.* 9, 014101 (2008).
8. Y. Ghen, J.E. Snyder, C.R. Schwichtenberg, and K.W. Dennis, *IEEE Trans. Magn.* 35, 3652 (1999).
9. X. Lu, G. Liang, Q. Sun, and C. Yang, *Mater. Lett.* 65, 674 (2011).
10. M.M. El-Okr, M.A. Salem, M.S. Salim, R.M. El-Okr, M. Ashoush, and H.M. Talaat, *J. Magn. Magn. Mater.* 323, 920 (2011).
11. S.H. Song, C.C.H. Lo, S.J. Lee, S.T. Aldini, J.E. Snyder, and D.C. Jiles, *J. Appl. Phys.* 101, 09C517 (2007).
12. K. Kamala Bharathi and C.V. Ramana, *J. Mater. Res.* 26, 584 (2011).
13. U. Acevedo, T. Gaudisson, R. López-Noda, S. Ammar, S. Nowak, and R. Valenzuela, *J. Spintron. Magn. Nanomater.* 1, 1 (2012).
14. L. Hao, D. Zhou, Q. Fu, and Y. Hu, *J. Mater. Sci.* 48, 178 (2013).
15. S. Kuila, J. Ray, A.K. Biswal, and P.N. Vishwakarma, *Solid State Phys. AIP Conf. Proc.* 1665, 130054 (2015).
16. M. Kumar, S. Shankar, O.P. Thakur, and A.K. Ghosh, *Mater. Lett.* 143, 241 (2015).
17. P.N. Oliveira, D. Alanis, R.D. Bini, D.M. Silva, G.S. Dias, I.A. Santos, L.F. Cótica, R. Guo, and A.S. Bhalla, *Integr. Ferroelectr.* 174, 88 (2016).
18. C.S. Antoniak, D. Schmitz, P.B. Frank, M.F. de Groot, S. Stienen, A. Warland, B. Krumme, R. Feyerherm, E. Dudzik, W. Kleemann, and H. Wende, *Nat. Commun.* 4, 2051 (2013).
19. A. Manjeera, M. Vittal, V.R. Reddy, G. Prasad, and G.S. Kumar, *Ferroelectrics* 519, 15 (2017).
20. S. Shankar, M. Kumar, A.K. Ghosh, O.P. Thakur, and M. Jayasimhadri, *J. Alloys Compd.* 779, 918 (2019).
21. S. Shankar, O.P. Thakur, and M. Jayasimhadri, *Mater. Chem. Phys.* 234, 110 (2019).
22. M. Rahman and C.V. Ramana, *J. Appl. Phys.* 116, 164108 (2014).
23. S. Shankar, M. Kumar, S. Kumar, O.P. Thakur, and A.K. Ghosh, *J. Alloys Compd.* 694, 715 (2017).
24. S. Chandarak, A. Ngamjarurojana, S. Srilomsak, P. Laoratanakul, S. Rujirawat, and R. Yimnirun, *Ferroelectrics* 410, 75 (2011).
25. T. Kimura and Y. Yoshida, *J. Am. Ceram. Soc.* 89, 869 (2006).
26. M. Suzuki, M. Miyayama, Y. Noguchi, and T. Uchikoshi, *J. Appl. Phys.* 104, 014102 (2008).
27. D.C. Sinclair and A.R. West, *J. Appl. Phys.* 66, 3850 (1989).
28. Y.D. Kolekar, L.J. Sanchez, and C.V. Ramana, *J. Appl. Phys.* 115, 144106 (2014).
29. S. Mohanty, R.N.P. Choudhary, R. Padhee, and B.N. Parida, *Ceram. Int.* 40, 9017 (2014).
30. C. Nayek, P. Murugavel, S. D. Kumar, and V. Subramanian, *Appl. Phys. A* 120, 615 (2015).
31. P. Ganguly, A.K. Jha, and K.L. Deori, *Solid State Commun.* 146, 472 (2008).

32. E. Barsoukov and J.R. Macdonald, *Impedance Spectroscopy, Theory, Experiment, and Applications* (New York: Wiley, 2005).
33. W. Li and R.W. Schwartz, *Appl. Phys. Lett.* 89, 242906 (2006).
34. Y.H. Gao, J. Yang, H. Shen, J.L. Sun, X.J. Meng, and J.H. Chu, *Appl. Phys. Lett.* 104, 122902 (2014).
35. H.S. Shulman, D. Damjanovic, and N. Setter, *J. Am. Ceram. Soc.* 83, 528 (2000).
36. T. Mondal, S. Das, T. Badapanda, T.P. Sinha, and P.M. Saruna, *Physica B* 508, 124 (2017).
37. A. Avogadro, F. Tabak, M. Corti, and F. Borsa, *Phys. Rev. B* 41, 6137 (1990).
38. M.F. Kotkata, F.A. Abdel-Wahab, and H.M. Maksoud, *J. Phys. D Appl. Phys.* 39, 2059 (2006).
39. S. Sahoo, P.K. Mahapatra, R.N.P. Choudhary, M.L. Nandagoswami, and A. Kumar, *Mater. Res. Express* 3, 065017 (2016).
40. A. Kaur, L. Singh, and K. Asokan, *Ceram. Int.* 44, 3751 (2018).
41. R.K. Panda, R. Muduli, S.K. Kar, and D. Behera, *J. Alloys Compd.* 615, 899 (2014).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

The current issue and full text archive of this journal is available on Emerald Insight at:
www.emeraldinsight.com/0265-671X.htm

RELIABILITY PAPER

Improving system reliability by optimal allocation of resources for discovering software vulnerabilities

Navneet Bhatt
Department of Operational Research, Faculty of Mathematical Sciences, University of Delhi, New Delhi, India
Adarsh Anand
Department of Operational Research, University of Delhi, New Delhi, India, and
Deepti Aggrawal
University School of Management and Entrepreneurship, Delhi Technological University, New Delhi, India

Received 29 July 2019
Revised 13 September 2019
Accepted 21 September 2019

Abstract

Purpose – The purpose of this paper is to provide a mathematical framework to optimally allocate resources required for the discovery of vulnerabilities pertaining to different severity risk levels.
Design/methodology/approach – Different sets of optimization problems have been formulated and using the concept of dynamic programming approach, sequence of recursive functions has been constructed for the optimal allocation of resources used for discovering vulnerabilities of different severity scores. Mozilla Thunderbird web browser data set has been considered for giving the empirical evaluation by working with vulnerabilities of different severities.
Findings – As per the impact associated with a vulnerability, critical and high severity level are required to be patched promptly, and hence, a larger amount of funds have to be allocated for vulnerability discovery. Nevertheless, a low or medium risk vulnerability might also get exploited and thereby their discovery is also crucial for higher severity vulnerabilities. The current framework provides a diversified allocation of funds as per the requirement of a software manager and also aims at improving the discovery of vulnerability significantly.
Practical implications – The finding of this research may enable software managers to adequately assign resources in managing the discovery of vulnerabilities. It may also help in acknowledging the funds required for various bug bounty programs to cater security reporters based on the potential number of vulnerabilities present in software.
Originality/value – Much of the attention has been focused on the vulnerability discovery modeling and the risk associated with the security flaws. But, as far as the authors' knowledge is concern, there is no such study that incorporates optimal allocation of resources with respect to the vulnerabilities of different severity scores. Hence, the building block of this paper contributes to future research.

Keywords Reliability, Resource allocation, Severity, Dynamic programming, Software vulnerability
Paper type Research paper

Nomenclature

n	number of severity levels pertaining to a software system	N_i	number of vulnerabilities present in the software in the ith severity level ($i = 1, \ldots, n$)
$\Omega_i(X_i(t))$	cumulative number of vulnerabilities discovered using resources in $(0, t]$ in ith severity level	r_i	vulnerability discovery rate ($i = 1, \ldots, n$)

This research did not receive any specific research funding.

International Journal of Quality & Reliability Management
© Emerald Publishing Limited
0265-671X
DOI 10.1108/IJQRM-07-2019-0246

$x_i(t)$	resource expenditure at time t for ith severity level		(i.e. corresponding to nth stage in dynamic programming algorithm)
X_i, Z	the amount of vulnerability discovery resources to be allocated to the ith severity level	N_{io}	aspiration level of the ith severity level (i.e. number of vulnerabilities desired to be discovered from ith severity level)
T	total time for the discovery process		
X_i^*	optimal value of X_i, $i = 1, \ldots, n$	p_i	the minimum proportion of total vulnerabilities to be discovered from ith severity level
$f_n(Z)$	optimal number of vulnerabilities discovered upto nth severity level		

1. Introduction

Advancement in the secure software system has motivated the software vendors to allocate funds for the discovery and mitigation of vulnerabilities encountered during its operational period. A huge share of vulnerabilities in the wild is reported by end users and based on their severity rating its remedy is provided. Severity also plays an important role in judging a flaw's damage potential. In order to make the system reliable, it is important for the software managers to utilize correct tools and procedures to allocate the resources in order to anticipate all the vulnerabilities present in software. Exploitation of vulnerabilities is not only affecting the vendor but is also a threat for human beings. As the security is criticality linked with a software usage, discovery of vulnerabilities by the security team and the white hat hackers is generally a major concern. Therefore, software managers must allocate appropriate efforts for the discovery of vulnerabilities. The discovery phenomenon is basically carried with the aim of discovering vulnerabilities and it requires empirical illustration. The mathematical linkage between the number of vulnerabilities discovered in a software and the discovery time is regarded as vulnerability discovery models (VDMs) in the literature (Ozment, 2007).

Many researchers have tried to formulate different vulnerability discovery conditions through VDMs. VDMs have been utilized to monitor the security profile of a software, to understand the security trends, deployment of patches and updates. The models are likewise to quantify the vulnerabilities discovery phenomenon in the operational phase. Efforts required in the discovery of vulnerabilities comprise of various factors which can be number of reporters, number of hours worked, usage efforts, etc. To mathematically study the discovery phenomenon of software vulnerabilities, various mathematical models have been proposed in the literature. Researcher considered various discovery patterns by considering different distribution functions such as exponential, Weibull, logistic, hump-shaped, normal and so on (Rescorla, 2005; Alhazmi and Malaiya, 2005; Kapur *et al.*, 2015; Anand and Bhatt, 2016; Anand *et al.*, 2017; Bhatt *et al.*, 2017, 2018; Kaur *et al.*, 2019). The model proposed by Rescorla (2005) is the first and most representative attempt to describe the vulnerability discovery rate. Alhazmi and Malaiya (2005) and Alhazmi *et al.* (2007) discussed that the discovery phenomenon follows s-shaped logistic curve and showed the importance of effort in the discovery of vulnerabilities. Later, Anand and Bhatt (2016) analyzed the vulnerability discovery phenomenon of software showing a hump-shaped vulnerability discovery rate.

It was Alhazmi and Malaiya (2005) who first studied the discovery process using the equivalent effort based on the installed user base in discovering the vulnerabilities more precisely than time alone. Later, Sommestad *et al.* (2012) evaluated the amount of efforts required by a penetration tester in an enterprise-based web application. The VDMs are considered as an interdisciplinary mathematical modeling framework build upon the principles of software reliability growth models (SRGMs) wherein various SRGMs based on testing effort functions have been deduced for the detection of faults underlying in a software (Kapur *et al.*, 2011). Authors such as Ohtera and Yamada (1990) developed an

optimization problem which minimizes the testing resources spent during the testing of modular software. Yamada *et al.* (1995) discussed an optimization problem considering the minimization of resources such that the count of leftover faults present in each module remains equal to the desired level when the resources are constrained. The importance of allocation of resources in understanding the testing and maintaining a software system is shown. Although allocation of resources has been successfully employed to designate faults in software, the usefulness of these methods to indicate vulnerabilities has not been extensively studied yet. Authors such as Berman and Ashrafi (1993) and Khan *et al.* (2008) showed the utility of the dynamic programming approach in the allocation of resources for the software systems. Of late, Anand *et al.* (2019) studied the resource allocation problem for a multi-upgraded software system using this approach.

Numerous studies (Kapur *et al.*, 2010; Cao *et al.*, 2017; Inoue *et al.*, 2017; Wei *et al.*, 2018) showed the importance of allocation of resources in understanding the testing and maintaining a software system. However, few authors associate effective bug bounty policies for the vulnerability discovery. Finifter *et al.* (2013) performed an experimental study to understand the vulnerability rewards programs based on the severity of vulnerability discovered. Zhao *et al.* (2017) developed an economic model that focuses on strategies for enhancing the effectiveness of bug bounty programs. Another study by Maillart *et al.* (2017) uses the notion of Brady *et al.* (1999) by validating the influence of bug bounty programs in the discovery of bugs. As mentioned in a report, the Google Chrome vulnerability program cost around $2.9m in 2017 and has paid incentives to 1,230 individuals based on different severity levels (Keller, 2018). All the aforesaid works have either focused on time-based vulnerability discovery modeling, formulated effort-based modeling or studied regarding the rewards paid to hackers through bug bounty programs. But, as per our knowledge, no previous work describes about optimal allocation of resources for the software vulnerability discovered based on their severity level. This is the first study that explores this ideology.

As per our understanding, it has been assumed that whenever software is released in the market, firms are likely to spend high expected payout for high and critical severity vulnerabilities. Although aiming just a high severity flaw will not only exclude a low or medium severity flaw but also results in an unequal distribution of resources, i.e., many times lower severity flaws are heavily allocated and for the critical severity bugs might remain with a smaller amount of funds for their discovery. Based on the severity level, vulnerabilities can be exploited as a critical/high severe vulnerability may impact the software significantly. It is an important task for a security manager to allocate sufficient funds for the discovery of vulnerabilities corresponding to different severity levels.

Our primary optimization objective is to discover a maximum number of vulnerabilities with different severity levels to be discovered based on the available set of resources. Second, another problem is considered that inculcated advancement in the first problem in terms of least proportion of vulnerabilities sought to be discovered from each severity rating. In doing so, dynamic programming approach is employed (Hadley, 1964). The application of dynamic programming has been utilized for the first time under the concept of vulnerability discovery modeling phenomenon in software. To validate the proposed approach, an empirical evaluation has been performed to show the applicability of resources in the discovery of vulnerabilities.

The major implications of this research are as follows. First, different optimization problems are formulated that will assist software managers in allocating optimal allocations of resources to discover and mitigate the risk pertaining to vulnerabilities of different severity. Another implication can be made from this research is that the technique allows the resources to be allocated in such a way that advocates in discovering a maximum number of vulnerabilities in each severity class.

The rest of the paper is structured as follows. In Section 2, we discuss the mathematical formulation for an effort-based VDM and discuss in detail the proposed

resource allocation problem in discovering vulnerabilities of different severity levels. Section 3 shows an empirical validation of the proposed approach. Finally, we conclude the paper in Section 4.

2. Building blocks for the proposed framework

2.1 Severity levels

A vulnerability discovered in the wild attains different severity levels based on the set of parameters that measures the vulnerability impact and exploitation to software (CVSS/First). As per the CVSS score computed based on different factors defined in FIRST.org, vulnerability is assigned with any of the four different severity ratings as follows: critical, high, medium and low as given in Table I. A vulnerability whose score lies in the critical range can be used by the attackers to run arbitrary code and require no user authorization and have a high likelihood of exploitation. For these types of vulnerabilities, it is advised to provide a patch as soon as possible. A vulnerability having a high-key impact may gather sensitive data and require partial user interaction for the exploitation. A medium severity level requires the attacker to manipulate individual victim via gaining respective privileges for successful exploitation. A low severity level vulnerability has a very little impact on an individual user and requires local or physical access in order to exploit vulnerability.

2.2 Effort-based vulnerability discovery phenomenon

VDMs are time-dependent parametric functions describing the behavior of vulnerability discovery process. The purpose behind the time dependency is that the vulnerabilities can be easily recorded based on their time of discovery. However, this does not inculcate the efforts advocated by a software firm used for the discovery of vulnerabilities in software. A vulnerability model discussed by Alhazmi *et al.* (2007) considered the vulnerability detection rate is proportional to the leftover vulnerabilities with respect to the effort. To mathematically model the discovery phenomenon, following differential equation can be formulated for *i*th severity level:

$$\frac{d\Omega_i(t)}{dt}/x_i(t) = r_i(N_i - \Omega_i(t)), \quad i = 1, \ldots, n. \tag{1}$$

Solving Equation (1) with the initial condition that, at $t = 0$, $X_i(t) = 0$, $\Omega_i(t) = 0$ we get:

$$\Omega_i(t) = N_i\left(1 - e^{-r_i X_i(t)}\right),\ i = 1, \ldots, n. \tag{2}$$

To describe the behavior of effort function, Alhazmi *et al.* (2007) used exponential distribution. In this study, we have also considered the exponential form of discovery resources such that the effort rate is proportional to the available resources. Hence, the mathematical form considering the efforts for *i*th severity level can be stated as follows:

$$\frac{dX_i(t)}{dt} = e_i(\alpha_i - X_i(t)), \tag{3}$$

Table I. Severity rating system

CVSS score range	Severity in advisory
0.1–3.9	Low
4.0–6.9	Medium
7.0–8.9	High
9.0–10	Critical

where e_i corresponds to the rate by which the available resources will be spent in ith severity level and α_i represents the total resources available for ith severity level. Upon solving Equation (3) under the initial condition $X_i(0) = 0$, we get:

$$X_i(t) = \alpha_i\left(1-e^{-e_i t}\right),\ i = 1, \ldots, n. \tag{4}$$

To study the resource allocation for discovering the vulnerabilities above set of equation is utilized for different set of severity levels.

2.3 Resource allocation problem

Software consists of different types of vulnerability having different severity levels. Majority of software firms, such as Mozilla Firefox and Google Chrome web browser, do maintain a Vulnerability Reward Program with the aim to minimize the exploitation likelihood of vulnerabilities by providing adequate incentives to the security community. As mentioned previously, in discovering of vulnerability of different severity levels such as low, medium, high and critical security researchers are awarded with incentives. Hence, in order to maximize the discovery potential of vulnerabilities, firms often schedule resources during the operational period of software.

Enormous amount of funds is spent during the development of software ranging from its initial development to its final deployment. The majority of resources are often consumed during its development phase; however, after release, maintenance has also played an important role in the recent time. The vulnerabilities present in a software are generally discovered after its eventual release. Taking into consideration the fact that a vulnerability belonging to different severity levels can be exploited in the wild. Let us consider a software consisting of vulnerabilities of different severities, which are being discovered independently such that firms wishes to maximize the vulnerabilities discovery, we can state the optimization problem:

$$\left.\begin{aligned} &\text{Max} \sum_{i=1}^{n} \Omega_i \\ &\text{Subject to} \\ &\sum_{i=1}^{n} X_i \leqslant Z,\ X_i \geqslant 0,\ i = 1, \ldots, n \end{aligned}\right\}, \tag{5}$$

where Ω_i is the expected count of vulnerabilities to be discovered of the ith severity level using X_i amount of resources and the Z is the threshold limit of resources available in the operational phase for the discovery.

We can restate the problem (5) by substituting the mean value function as defined in Equation (2) as follows:

$$\left.\begin{aligned} &\text{Max} \sum_{i=1}^{n} \Omega_i(X_i) = \sum_{i=1}^{N} N_i\left(1-e^{-r_i X_i}\right) \\ &\text{Subject to} \\ &\sum_{i=1}^{n} X_i \leqslant Z,\ X_i \geqslant 0,\ i = 1, \ldots, n \end{aligned}\right\}. \tag{6}$$

In order to obtain optimal allocation of resources for discovering vulnerabilities pertaining to different severity levels, we have used the dynamic programming approach by constructing recursive functions. It is an optimization technique that usually simplifies a

decision problem by constructing a sequence of multistage decisions in a recursive fashion. As per Bellman's principle of optimality, we can write the forward recursive functions as follows (Hadley, 1964).

For $i=1$ we have:

$$f_1(Z) = \max_{X_i=Z_1} \{N_1(1-e^{-r_1X_1})\}. \tag{7}$$

And for $i=n$, we have:

$$f_n(Z) = \max_{0 \leqslant X_n \leqslant Z} \{N_n(1-e^{-r_nX_n}) + f_{n-1}(Z-X_n)\};\ \text{for } 2,\ldots,n. \tag{8}$$

Using Equations (7) and (8), allocation of resources can be made to discover vulnerabilities of different severity levels. To calculate the optimal value of resources required in discovering vulnerabilities from each severity level, we make the use of theorem given in the study of Kapur *et al.* (2003) to calculate the values of X_i^*. The allocation of resources is done in a sequential fashion and there might be some instances in which the vulnerabilities index with higher number may not get any allocation due to a lower discovery index (N_ir_i). Even if the allocation of resources is optimal, there might be some cases in which an adequate amount of resources is not allocated to maintain some minimum discovery count of vulnerabilities in each severity level. This may increase the likelihood of any exploitation by the black-hats if the vulnerabilities of any specific severity type are left undetected by the vendor. Therefore, it is expected that a certain proportion of vulnerabilities are desired to be discovered for each severity level. Hence, (6) needs to be reformulated by considering a constraint that will specify a minimum desired level of vulnerabilities to be discovered from each severity level. The resulting problem can be described as follows:

$$\left.\begin{aligned} &\text{Max} \sum_{i=1}^{n} \Omega_i(X_i) = \sum_{i=1}^{n} N_i(1-e^{-r_iX_i}) \\ &\text{Subject to} \\ &\sum_{i=1}^{n} X_i \leqslant Z,\ X_i \geqslant 0,\ i=1,\ldots,n \\ &\Omega_i = N_i(1-e^{-r_iX_i}) \geqslant p_iN_i = N_{i0},\ i=1,\ldots,n \end{aligned}\right\}, \tag{9}$$

where p_i is the least expected proportion of vulnerabilities and N_{i0} denotes a minimum number of vulnerabilities required to be discovered from each severity level. The above problem (9) can be solved either by a method of reducing the dimensionality by using the Lagrange multiplier approach or by considering the substitution process. The prior approach may lead to an increased complexity in finding an optimal solution at each stage and does not provide a closed form solution. Our aim is to alter (9) into a problem that have a comparable structure as of (6) such that an optimal solution can be obtained. For that very reason that add-on constraint in (9) can be altered as follows:

$$\Omega_i(X_i) \geqslant N_{i0} \Rightarrow N_i(1-e^{-r_iX_i}) \geqslant N_{i0}. \tag{10}$$

The above equation can be further simplified as:

$$X_i \geqslant -\frac{1}{r_i} \log\left[1-\frac{N_{i0}}{N_i}\right] = Z_i(\text{say}),\ i=1,\ldots,n. \tag{11}$$

Equation (9) can be further amended by making use of Equation (11) as:

$$\left.\begin{aligned} &\text{Max} \sum_{i=1}^{n} \Omega_i(X_i) = \sum_{i=1}^{n} N_i\left(1-e^{-r_i X_i}\right) \\ &\text{Subject to} \\ &\sum_{i=1}^{n} X_i \leqslant Z,\ X_i \geqslant 0,\ i = 1, \ldots, n \\ &X_i \geqslant Z,\ i = 1, \ldots, n \end{aligned}\right\}. \tag{12}$$

The above problem is again converted into a simpler form by considering $Y_i = X_i - Z_i$; ($i = 1, 2, \ldots, n$). Equation (12) is reduced to (13) as follows:

$$\left.\begin{aligned} &\text{Max} \sum_{i=1}^{n} \Omega_i = \sum_{i=1}^{n} \overline{N_i}\left(1-e^{-r_i Y_i}\right) \\ &\text{Subject to} \\ &\sum_{i=1}^{n} Y_i \leqslant Z - \sum_{i=1}^{n} Z_i = \overline{Z}\,(Let) \\ &Y_i \geqslant 0,\ i = 1, \ldots, n \\ &\overline{N_i} = N_i - N_{i0},\ i = 1, \ldots, n \end{aligned}\right\}, \tag{13}$$

where we let $Z - \sum_{i=1}^{n} Z_i = \overline{Z}$ such that the obtained (13) can be solved using the theorem proposed by Kapur *et al.* (2003) as its structure is similar to (6). To obtain the optimal solution, we have performed an empirical evaluation in the next section.

3. Numerical illustration

In order to determine the optimal allocation of resources numerically, we have considered a case of software application Mozilla Thunderbird web browser (NVD, 2018). The National Vulnerability Database (NVD) consists of a CVSS label to classify vulnerabilities into different severity levels. For simulating the data based on our requirement, we have categorized the CVSS score into six levels in order to have a sufficient count of vulnerabilities in each level. As mentioned in Table I, the CVSS scores are categorized into four severity level based on their CVSS scores. But, in our case the distribution of vulnerabilities is done such that some least number of vulnerabilities lies in each clubbed category so that the estimation of vulnerabilities makes sense. Hence, the vulnerabilities underlying in the following score distribution: 0.1–3.4, 3.5–4.4, 4.5–5.4, 5.5–7.4, 7.5–8.4 and 8.5–10 are clubbed together for a time period of 2003–2008. As per the criticality of vulnerabilities, the resources are often allocated for the discovery of vulnerabilities that have a higher severity score. The vulnerabilities having a lower severity rating may not trigger any severe risk. Hence, the security manager always tries to capture as much vulnerabilities that have a high risk by allocating as much resources to minimize the exploitation risk. Here, we have assumed total resources that are available for the discovery process as $Z = 1{,}000$ (in thousand dollars). To examine the optimum allocation of resources, we first estimated the parameters N_i and r_i for the software Mozilla Thunderbird pertaining to the considered six severity levels which are given in Table II.

Here, we first solve (6) for generating the optimal values of X_i^* for the various severity levels associated with the vulnerabilities by making the use of recursive Equations (7) and (8) and the obtained values are presented in Table II (Schrage, 1999). The table also provides

information related to the expected number of vulnerabilities discovered, percentage of vulnerability discovered and leftover vulnerabilities in each severity level. It can be observed that out of 636 vulnerabilities only 458 vulnerabilities are discovered through this allocation of resources and contributing around 71 percent discovery of vulnerabilities. Moreover, we also find the proportion of vulnerabilities discovered in each clubbed CVSS score and found that only 67 percent of critical vulnerabilities are discovered as per the current allocation. And the remaining 33 percent of vulnerabilities which are left undetected and can be exploited if discovered by a black-hat. In context to the security of software, management will aim to discovery a maximum number of vulnerabilities corresponding to critical and high severity risk. Hence, for the analysis, we consider at least 90 percent of vulnerabilities for the severity level corresponding to 8.5–10 score are required to be discovered and similarly a minimum of 80, 70, 60, 55 and 50 percent of the vulnerabilities to be discovered for the severity level pertaining to the CVSS score of 7.5–8.4, 5.5–7.4, 4.5–5.4, 3.5–4.4 and 0.1–3.4, respectively.

For the allocation of new resources, we solve (9) in that we consider the discovery of some least proportion of vulnerabilities. The results are summarized in Table III. Here, we allocate the remaining resources for the discovery of vulnerabilities leftover after the allocation of resources as per Equation (11).

It is apparent that as per the initial allocation around 458 vulnerabilities were discovered as a whole and were reduced to a count of 419 vulnerabilities. The reduction in the discovery can be attributed due to the capping of vulnerabilities placed as per the associated risk present in each severity level, thus, leading to a decline in overall discovery percentage of vulnerability to 66 percent whereas the initial percentage was close to 72 percent. In addition to this, if the management wants a certain fraction of vulnerabilities to be discovered then additional resources would be required. In order to maintain the tradeoff between the overall discovery percentage, it is required to allocate extra resources in the discovery. In the present case, if the management desires to allocate an extra of 200 (in thousand dollars) then the current case is summarized in Table IV.

Table II. Parameter estimates and initial resource allocation (6)

CVSS score distribution	N_i	r_i	X_i^*	Ω_i^*	% of vulnerabilities discovered	% of vulnerabilities remaining
0–3.4	203.792	0.005	303.153	160.011	78.517	21.483
3.5–4.4	65.863	0.008	107.878	37.929	57.588	42.412
4.5–5.4	72.918	0.012	113.683	54.745	75.078	24.922
5.5–7.4	34.946	0.029	52.343	27.287	78.084	21.916
7.5–8.4	205.408	0.004	335.186	142.471	69.360	30.640
8.5–10	53.219	0.013	87.757	35.460	66.629	33.371
Total	636.145		1000	457.903	71.981	28.019

Table III. Resource allocation from (9)

CVSS score distribution	N_i	N_{i0}	Z_i^*	Y_i^*	$\Omega_i(Y_i)$	Ω_i^*	% of vulnerabilities discovered	% of vulnerabilities remaining
0–3.4	203.792	101.896	136.635	6.271	3.191	105.086	51.56	48.44
3.5–4.4	65.863	36.225	100.429	0.000	0.000	36.225	55.00	45.00
4.5–5.4	72.918	43.751	74.971	0.000	0.000	43.751	60.00	40.00
5.5–7.4	34.946	24.462	41.516	0.000	0.000	24.462	70.00	30.00
7.5–8.4	205.408	164.326	456.061	0.000	0.000	164.326	80.00	20.00
8.5–10	53.219	47.897	184.118	0.000	0.000	47.897	90.00	10.00
Total	636.145	418.557	993.729	6.271	3.191	421.747	66.29	33.70

CVSS score distribution	N_i	N_{i0}	Z_i^*	Y_i^*	$\Omega_i(Y_i)$	X_i^*	Ω_i^*	% of vulnerabilities discovered	% of vulnerabilities remaining
0–3.4	203.792	101.896	136.635	158.782	56.362	295.417	158.258	77.657	22.343
3.5–4.4	65.863	36.225	100.429	2.514	0.586	102.942	36.811	55.890	44.110
4.5–5.4	72.918	43.751	74.971	35.501	10.267	110.472	54.018	74.081	25.919
5.5–7.4	34.946	24.462	41.516	9.474	2.519	50.990	26.981	77.207	22.793
7.5–8.4	205.408	164.326	456.061	0.000	0.000	456.061	164.326	80.000	20.000
8.5–10	53.219	47.897	184.118	0.000	0.000	184.118	47.897	90.000	10.000
Total	636.145	418.557	993.729	206.271	69.735	1200.000	488.292	76.758	23.242

Table IV.
Allocation after addition of extra resources (13)

From Table IV, we may notice that as a whole about 77 percent of vulnerabilities are discovered and at the same time maintaining the respective percentage of vulnerabilities discovered from each severity level. Hence, as per the above calculations, we can advocate that allocation of resources has a major impact on the discovery of vulnerabilities. Moreover, the decision regarding splitting of resources among different severity levels minimizes the risk of exploitation for a vulnerability belonging to a critical level.

4. Conclusion

This paper considers an important aspect by providing an optimization approach accounting for the allocation of resources in the discovery of vulnerabilities belonging to various severity levels. Here, we have formulated different sets of optimization problems to describe how the resources are allocated in order to attain an aspirational level for vulnerability discovery in each severity level. We consider dynamic programming methodology and utilizing recursive set of conditions to obtain optimal resources that are required to be fulfilled. Allotting more resources often increases the discovery intensity, as a greater number of vulnerabilities are detected. Although firms aim to discovery as many vulnerabilities exist in the software either inhouse or via white hat users but the vulnerabilities having a high severity rating are required to be discover at the earlier and to be allocated sufficient funds.

References

Alhazmi, O.H. and Malaiya, Y.K. (2005), "Modeling the vulnerability discovery process", *Proceedings of the 16th IEEE International Symposium on Software Reliability Engineering, IEEE, Chicago, IL*, pp. 138-147.

Alhazmi, O.H., Malaiya, Y.K. and Ray, I. (2007), "Measuring, analyzing and predicting security vulnerabilities in software systems", *Computers & Security*, Vol. 26 No. 3, pp. 219-228.

Anand, A. and Bhatt, N. (2016), "Vulnerability discovery modeling and weighted criteria based ranking", *Journal of the Indian Society for Probability and Statistics*, Vol. 17 No. 1, pp. 1-10.

Anand, A., Das, S., Singh, O. and Kumar, V. (2019), "Resource allocation problem for multi versions of software system", *2019 Amity International Conference on Artificial Intelligence (AICAI), IEEE, Dubai*, pp. 571-576.

Berman, O. and Ashrafi, N. (1993), "Optimization models for reliability of modular software systems", *IEEE Transactions on Software Engineering*, Vol. 19 No. 11, pp. 1119-1123.

Bhatt, N., Anand, A., Aggrawal, D. and Alhazmi, O.H. (2018), "Categorization of vulnerabilities in a software", in Anand, A. and Ram, M. (Eds), *System Reliability Management: Solutions and Technologies*, CRC Press, Boca Raton, FL, pp. 121-135.

Bhatt, N., Anand, A., Yadavalli, V.S.S. and Kumar, V. (2017), "Modeling and characterizing software vulnerabilities", *International Journal of Mathematical, Engineering and Management Sciences*, Vol. 2 No. 4, pp. 288-299.

Brady, R.M., Anderson, R.J. and Ball, R.C. (1999), "Murphy's law, the fitness of evolving species, and the limits of software reliability (No. UCAM-CL-TR-471)", Computer Laboratory, University of Cambridge, Cambridge.

Cao, Z., Lin, J., Wan, C., Song, Y., Zhang, Y. and Wang, X. (2017), "Optimal cloud computing resource allocation for demand side management in smart grid", *IEEE Transactions on Smart Grid*, Vol. 8 No. 4, pp. 1943-1955.

Finifter, M., Akhawe, D. and Wagner, D. (2013), "An empirical study of vulnerability rewards programs", *Presented as part of the 22nd (USENIX) Security Symposium ((USENIX) Security 13), Washington, DC*, pp. 273-288.

Hadley, G. (1964), *Nonlinear and Dynamic Programming*, Addision-Wesley, Reading, MA.

Inoue, S., Hotta, K. and Yamada, S. (2017), "On estimation of number of detectable software faults under budget constraint", *International Journal of Mathematical, Engineering and Management Sciences*, Vol. 2 No. 3, pp. 135-139.

Kapur, P.K., Chanda, U. and Kumar, V. (2010), "Dynamic allocation of testing effort when testing and debugging are done concurrently", *Communication in Dependability and Quality Management: An International Journal Serbia*, Vol. 13 No. 3, pp. 14-28.

Kapur, P.K., Jha, P.C. and Bardhan, A.K. (2003), "Dynamic programming approach to testing resource allocation problem for modular software", *Ratio Mathematica*, Vol. 14 No. 1, pp. 27-40.

Kapur, P.K., Sachdeva, N. and Khatri, S.K. (2015), "Vulnerability discovery modeling", *International Conference on Quality, Reliability, Infocom Technology and Industrial Technology Management, I. K. International Publishing*, pp. 34-54.

Kapur, P.K., Pham, H., Gupta, A. and Jha, P.C. (2011), *Software Reliability Assessment with OR Applications*, Springer Series in Reliability Engineering, Springer-Verlag, London.

Kaur, J., Anand, A. and Singh, O. (2019), "Modeling software vulnerability correction/fixation process incorporating time lag", in Anand, A. and Ram, M. (Eds), *Recent Advancements in Software Reliability Assurance*, CRC Press, Boca Raton, FL, pp. 39-58.

Keller, J. (2018), "Vulnerability reward program: 2017 year in review", available at: https://security.googleblog.com/2018/02/vulnerability-reward-program-2017-year.html (accessed July 12, 2019).

Khan, M.G., Ahmad, N. and Rafi, L.S. (2008), "Optimal testing resource allocation for modular software based on a software reliability growth model: a dynamic programming approach", *International Conference on Computer Science and Software Engineering, IEEE, Vol. 2, Hubei*, pp. 759-762.

Maillart, T., Zhao, M., Grossklags, J. and Chuang, J. (2017), "Given enough eyeballs, all bugs are shallow? Revisiting Eric Raymond with bug bounty programs", *Journal of Cybersecurity*, Vol. 3 No. 2, pp. 81-90.

NVD (2018), "NVD data feeds", available at: https://nvd.nist.gov/vuln/data-feeds (accessed July 12, 2019).

Ohtera, H. and Yamada, S. (1990), "Optimal allocation and control problems for software-testing resources", *IEEE Transactions on Reliability*, Vol. 39 No. 2, pp. 171-176.

Ozment, A. (2007), "Vulnerability discovery and software security", PhD dissertation, University of Cambridge Computer Laboratory, Cambridge, available at: http://andyozment.com/papers/ozment_dissertation.pdf

Rescorla, E. (2005), "Is finding security holes a good idea?", *IEEE Security & Privacy*, Vol. 3 No. 1, pp. 14-19.

Schrage, L. (1999), *Optimization Modeling with LINGO*, Lindo Systems, North Dayton Street, Chicago, IL.

Sommestad, T., Holm, H. and Ekstedt, M. (2012), "Effort estimates for vulnerability discovery projects", *2012 45th Hawaii International Conference on System Sciences, IEEE, Maui, HI*, pp. 5564-5573.

Wei, W., Fan, X., Song, H., Fan, X. and Yang, J. (2018), "Imperfect information dynamic Stackelberg game based resource allocation using hidden Markov for cloud computing", *IEEE Transactions on Services Computing*, Vol. 11 No. 1, pp. 78-89.

Yamada, S., Ichimori, T. and Nishiwaki, M. (1995), "Optimal allocation policies for testing-resource based on a software reliability growth model", *Mathematical and Computer Modelling*, Vol. 22 No. 10, pp. 295-301.

Zhao, M., Laszka, A. and Grossklags, J. (2017), "Devising effective policies for bug-bounty platforms and security vulnerability discovery", *Journal of Information Policy*, Vol. 7, pp. 372-418.

Further reading

Anand, A., Das, S., Aggrawal, D. and Klochkov, Y. (2017), "Vulnerability discovery modelling for software with multi-versions", in Ram, M. and Davim, J.P. (Eds), *Advances in Reliability and System Engineering*, Springer International Publishing, pp. 255-265.

About the authors

Navneet Bhatt is presently pursuing his PhD Degree from the University of Delhi, Delhi, India. He received his MPhil and Master's Degrees from the Department of Operational Research, University of Delhi, Delhi, India in 2015 and 2013, respectively. His research areas include software reliability, software security and vulnerability assessment through machine learning techniques.

Adarsh Anand did his Doctorate in the area of Innovation Diffusion Modeling in Marketing and Software Reliability Assessment. Presently, he is working as Assistant Professor in the Department of Operational Research, University of Delhi (INDIA). He has been conferred with Young Promising Researcher in the field of Technology Management and Software Reliability by Society for Reliability Engineering, Quality and Operations Management (SREQOM) in 2012. He is a lifetime member of the Society for Reliability Engineering, Quality and Operations Management (SREQOM). He is also on the editorial board of *International Journal of System Assurance and Engineering Management* (Springer). He has Guest edited several Special Issues for Journals of international repute. He has edited two books, namely "System Reliability Management (Solutions and Technologies)" and "Recent Advancements in Software Reliability Assurance" under the banner of Taylor and Francis (CRC Press). He has publications in journals of national and international repute. His research interest includes modeling innovation adoption and successive generations in marketing, software reliability growth modeling and social media analysis. Adarsh Anand is the corresponding author and can be contacted at: adarsh.anand86@gmail.com

Deepti Aggrawal is currently working as Assistant Professor at USME, Delhi Technological University, India. She obtained her PhD Degree from Department of Operational Research, University of Delhi. She was Operations Manager in Axis Bank till she joined as a research scholar in the Department of Operational Research in 2011. Her research areas include marketing and software reliability. She is a life member of SREQOM and has publications in journals of national and international repute.

For instructions on how to order reprints of this article, please visit our website:
www.emeraldgrouppublishing.com/licensing/reprints.htm
Or contact us for further details: **permissions@emeraldinsight.com**

Global Journal of Flexible Systems Management
https://doi.org/10.1007/s40171-019-00220-x

ORIGINAL RESEARCH

Managerial Paradox Toward Flexibility: Emergent Views Using Thematic Analysis of Literature

Sanjai Kumar Shukla[1] · **Sushil**[2] · **Manoj Kumar Sharma**[3]

Received: 3 September 2019 / Accepted: 21 September 2019
© Global Institute of Flexible Systems Management 2019

Abstract *The organizational flexibility finds a substantial place in the literature; however, limited evidence of its institutionalization by managers is noticed. This paper analyzes the managerial paradox toward flexibility in managing business complexity, uncertainty handling, organizational reorientation, and structuring decisions across organizational functions. Thematic analysis is used to investigate the managerial flexibility paradox. The literature is identified in three categories related to workplace, frameworks or concepts, and business environment. It further shows two kinds of orientations: the first one deals with concerns in the implementation of flexibility, and the second showcases the concomitant benefits or performance. The literature related to concerns deals with uncertainty and risk management, while the performance-related literature deals with the augmentation of flexibility theories. The practicing manager seeks advice in the literature to implement the concepts of flexibility and finds its limited availability. This is the prime reason for the emergence of managerial paradox. This paper recommends that future contributions should emphasize the guidelines for flexibility adoption while expounding the flexibility theories. The findings presented in this paper can potentially draw the attention of academicians and practitioners to devise the ways to implement and enhance organizational flexibility.*

Keywords Flexibility paradox · Flexibility practices · Flexibility implementation · Managerial flexibility · Organizational flexibility · Workplace flexibility

✉ Sanjai Kumar Shukla
shuklasanjai@rediffmail.com;
shuklasanjaikumar@gmail.com

Sushil
sushil@dms.iitd.ac.in; profsushil@gmail.com

Manoj Kumar Sharma
osd.eastcampus@dtu.ac.in

[1] Bharat Electronics Ltd, Ghaziabad, UP 201011, India

[2] Department of Management Studies, Indian Institute of Technology Delhi, New Delhi 110016, India

[3] Delhi Technological University, Rohini, India

Introduction

Managerial flexibility is valued as a strategic asset (Ten-Dam 1987). The importance of managerial flexibility resulted in "Managerial Flexibility Act 2001" introduction in the US senate (Managerial Flexibility Act 2001). This is legislator visualization flexibility that provides freedom to manage and influence the employee's management, retirement, personal property, separation, and retirement provisions. The literature suggests the need for flexibility for organic structure, innovation, planning, and entrepreneurial organizations in the dynamic environment (Eppink 1978; Mintzberg 1980, 1989). Managerial flexibility institutes positive changes in top and bottom line performers (Sullivan 1994).

Two distinct types of contributions are observed on managerial flexibility. The first deals with how managerial flexibility helps the organization performance, and the second deals with how the managerial flexibility is impacted by organizational policies and resources. Flexibility is a capability nurtured systematically over a period of time through various initiatives. The flexibility enhances the organizational efficiency and performance which attracts managerial attention. The convergence of efficiency, agility, systems thinking, strategy, and performance

Published online: 12 October 2019

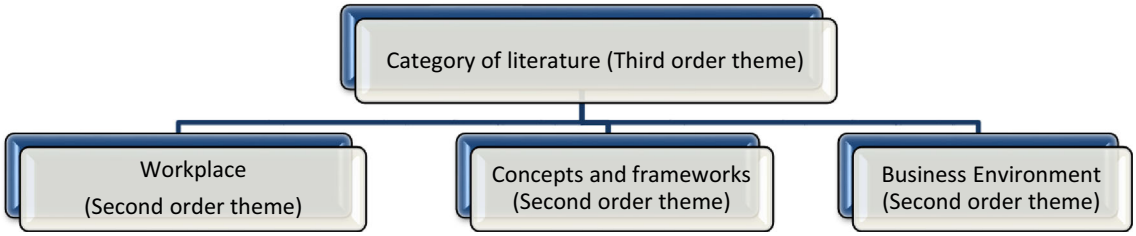


Fig. 1 Organization of the literature

for flexibility is observed (Ebben and Johnson 2005; Nadkarni and Narayanan 2007; Sharma et al. 2010a; Subramanyam et al. 2012; Chowdhury and Jangle 2018; Khorasani 2018; Shukla and Sushil 2018, 2020a, b). The flexible systems focus on core activities, avoid wastage of effort, and institute reconfiguration capabilities enabling organizations to adapt to a new situation for strategic goals (Melnik et al. 2019). The process-based flexibilities result in well-organized manufacturing, productive R&D effort, and financial performance (Narasimhan et al. 2004; White et al. 2005; Sherehiy et al. 2007; Deng et al. 2014). The sustenance and vitality are other predominant upshots of flexibility facilitated by organizational change (Sushil 2012a, b, 2013). The flexibility helps to hedge against the uncertain business environment and build performance (Marschak and Nelson 1962; McKeown 2012). Competitiveness and organizational design endorse viable actions for flexibility (Volberda 1992, 1996, 1997). The supply chain integration, flexibility, and barriers are a special concern of uncertainty management (Chowdhury et al. 2018; Lu et al. 2018). Incorporating flexibility in strategy has been advised by Mintzberg et al. (1995). The literature has also put considerable focus on enablers and inhibitors of flexibility (Chow and Cao 2008; Bottani 2009; Avazpour et al. 2014; Sopelana et al. 2014).

The benefits offered by flexibility stimulate its institutionalization into organizational practices. The limited availability of work practices, guidelines, operating procedures, and strategies for effective implementation of flexibility causes its ignorance leading to paradoxical behavior. Some literary debates show the contradictions and paradox toward flexibility. For example, the "Flowing stream strategy framework" by Sushil (2012a) explains the institutionalization of flexibility by maintaining vital factors of organizational continuity while proceeding for desired change. However, De Leeuw and Volberda (1996) noted the paradox of flexibility results from change and continuity. Few authors suggest flexibility as mediator or driver of performance rather than considering its direct impact on the performance (Sharma et al. 2010; Acharya 2019; Shalender and Yadav 2019). There are few studies related to the paradoxes in the context of risk and inter-organizational collaborations, covert resistance, convergent policy and divergent practice, new product development, cognitive functioning, and high-variety and low-volume manufacturing (Leonard-Barton 1992; Clarke 2002; Couchman and Fulop 2002; Fleming and Spicer 2002; Katic and Agarwal 2018; Ganuthula and Sinha 2019). However, the managerial paradox toward flexibility is largely ignored which this paper addresses. This study has been done in two parts: in the first part, literary gap between practice and theory, intention and execution of flexibility is analyzed, and the second part synthesizes the findings and builds recommendation to handle the managerial paradox of flexibility.

The selection criteria emphasize the theory- and practice-related contributions for managerial action within the boundary of organizational control for building flexibility. Three categories of literature, i.e., workplace, business environment, and concepts and frameworks, have emerged from thematic analysis. These categories capture all the dimension of flexibility relevant to managerial attention and depicted in Fig. 1 and examined further.

The contributions related to performance and concerns of flexibility are specially searched. The performance and concern are used as codes in the thematic analysis. Three types of managerial paradoxes toward flexibility have been identified during analysis.

First Managerial Paradox Toward Flexibility

The first managerial paradox toward flexibility has emerged after analyses of the gap between the literature dealing with the benefits of flexibility and advice to obtain it.

Second Managerial Paradox Toward Flexibility

The second paradox reflects the attitude of lower and higher management toward flexibility. The top management decides the strategy where flexibility finds its place. The lower and middle managerial strategies seek workable advice and have a different perception than top management. The difference is primarily in the approach related to strategic and operational aspects from where this paradox emerges.

Third Managerial Paradox Toward Flexibility

As the flexibility initiatives require resources, the extent and scope of initiatives create a third kind of paradox. The management has to identify relevant flexibility and define optimal flexibility between total flexibility at a much higher cost and rigidity at minimal cost in the absence of literary support to define such optimality (Sushil 2015b).

Methodology

Thematic analysis is used to investigate the managerial paradox. The paradox is identified by finding themes associated with managerial functions. The explicit meaning associated with a given context can be easily identified using this method than other methods such as content analysis (Saldana 2009). Thematic analysis helps to build the perspectives and social construction of representation (Braun and Clarke 2013). The exploration of the benefit of flexibility initiatives in organizations and impediments or concerns in the implementation of these management actions is identified using thematic analysis to build the perspective. Following research questions analyze the managerial paradox toward flexibility using a thematic analysis.

- What are the benefits of flexibility initiatives undertaken by management in the sphere of organizational control?
- What are concerns, hurdles, and mechanisms in implementing flexibility initiatives?

Thematic Analysis

Six stages of thematic analysis have been used to find themes in the literature in qualitative research pertaining to research questions. The approaches demonstrated by Braun and Clarke (2006), Fereday and Muir-Cochrane (2006), and Gibbs (2007) have been used for the identification of the common pattern of managerial interest toward flexibility in different contributions.

Stage 1: Familiarity with the Data

Managerial flexibility strongly relates to planning and execution. The criterion of identifying the relevant literature is defined in light of managerial paradox. The keywords, i.e., flexible, flexibility, agile, agility in title, abstract, and body of papers, are used to search the literature. The flexibility in contexts like integration, agility, cost-effectiveness, responsiveness, quickness, sustainability, speed, leanness, aspects of continuity, change, dynamic capabilities, and organizational control are used to get comprehensive set of literary contributions (Volberda 1997; Qumer and Henderson-Sellers 2008; Sushil 2013; Avazpour et al. 2014; Borland et al. 2016). The literature related to agile methodology is used to identify the management concerns in agility adoption. Further to it, only those contributions are considered which make recommendations within organizational controls. Flexibility evaluation papers are considered if they recommend workable practices. The literature related to the measurement of specific flexibilities is grouped in the performance category since it uses enhancement of independent variables, i.e., organizational effort/investments, to increase the dependent variable, i.e., flexibility performance.

Stage 2: Generating Initial Codes

The "concern" and "performance" are used as the codes (Table 1). The papers in the performance category put less focus on action and elaborate niche flexibility benefit which invites management attention. The performance-related papers also give a broad recommendation for flexibility based on simulation, results, and reviews. The decision to put the literature into concern orientation is made if the recommendation of papers significantly deals with advice validated by example or field data that can be potentially utilized by managers as the guideline to enhance organizational flexibility. These papers also deal

Table 1 Code used for thematic analysis

Code 1	Label	Performance
	Definition	Organizational performance due to flexibility
	Description	1. The context in the literature explicitly or implicitly mentioning situations where flexibility provides organizational performance or benefit
Code 2	Label	Concern
	Definition	Organizational concern in implementing flexibility initiatives
	Description	1. The context in the literature explicitly or implicitly mentioning situations where management deliberates on resource requirement and feasibility of such initiatives. 2. The literary context exploring ways and methods of implementing such initiatives

with flexibility practices that potentially address the management concern and primarily action-oriented.

Stage 3: Searching For Themes

The relationship with codes "concern" and "performance" has been further looked in the context of flexibility for identifying the first-order, clustered, second-order, and third-order themes.

Stage 4: Reviewing Themes

The first-order, clustered, second-order, and third-order themes provide "Focus of literature," "Orientation of literature," "Subcategory of literature," and "Category of literature," respectively.

Stage 5: Defining and Naming Themes

Themes are named and grouped after iterative review and summarized in Figs. 2, 3, and 4.

Stage 6: Production of Report

The subsequent analysis is based on the commonality and discrepancies across themes in all the three categories of the literature.

Results of Thematic Analysis

Thematic analysis has been performed iteratively, and the results are summarized under various themes. The third-order themes represent broad categories of the literature, i.e., workplace, concepts and frameworks, and business environment. The context of flexibility is noted in first-order themes and aggregated in second-order themes, and clustered themes as per descriptions and tabulation of data in this section. The trends of literary work for managers are focused on risk and uncertainty mitigation strategies while substantial work is underway for advancing flexibility concepts. The detailed analysis shows that the gap between practice and theory is more prevalent. The literature in the area of strategy, practice, and operations in flexibility context is identified under "workplace." The literature dealing with concepts, frameworks, relationships, performance analysis of systems in the context of flexibility is incorporated under "concept and framework" category. The context outside organization refers to the volatile "business environment" with the recommendation of managerial action from within the organizations. The comparative distribution of contributions utilized for setting the context of study, analysis, and validation of findings is summarized in the "Appendix".

Third-Order Theme: Workplace

This category of the literature refers to the people and organizational processes to improve performance. The second-order themes are related to business processes, sustainability, strategy, network/supply chain, and actors. Figure 2 provides a summary of observations of workplace flexibility related to all themes.

Second-Order Theme: Business Process

Organizational communication is an important business process. The technological aid in this process results in flexible and adaptable communication across the organization (Moore 2001). Different communication modes in the integrated form are crucial for interactivity and strategic intent (Struweg 2014). Organizational communication is an important factor considered to examine managerial flexibility by Bamel et al. (2013). The impact of information systems (IS) on flexible work plans and business processes finds extensive focus in the literature. Earlier work envisaged the planning for flexibility in manufacturing systems using optimal cost and capacity analysis (Tang 1991). The literature also addressed the challenges in the enablement of flexible work plans using information systems in functional, behavioral, and operational perspectives (Reichert and Weber 2012). The substantial use of information and communication technologies enhances business process flexibility which also impacts business alignment, communication, and collaboration (Steinbring et al. 2013). The reuse of objects and first-time-right design in the information technology industry is emphasized for instituting the flexibility and speeding up the development process (Favaro et al. 1998; Tseng and Lin 2011). For a software product, different scenarios are to be considered to meet the risk and uncertainty in the final usage, and incorporate flexibility (Chastek et al. 2009). In the information technology organizations, the service orientation in design process enables flexible products (Gulledge and Deller 2009).

The relationship of business processes with flexibility is noticed by many authors. The flexibility in business processes handles market dynamics and builds a value proposition in the IT sector (Raghuveer et al. 2014). The innovation performance results from interaction among multiple resources facilitated by coordination flexibility (Li et al. 2017). There exists a strong relationship between knowledge management, agile methodology acceptance, and strategic flexibility that facilitates innovation (Chan and Thong 2009; Kamasak et al. 2016). The vertical

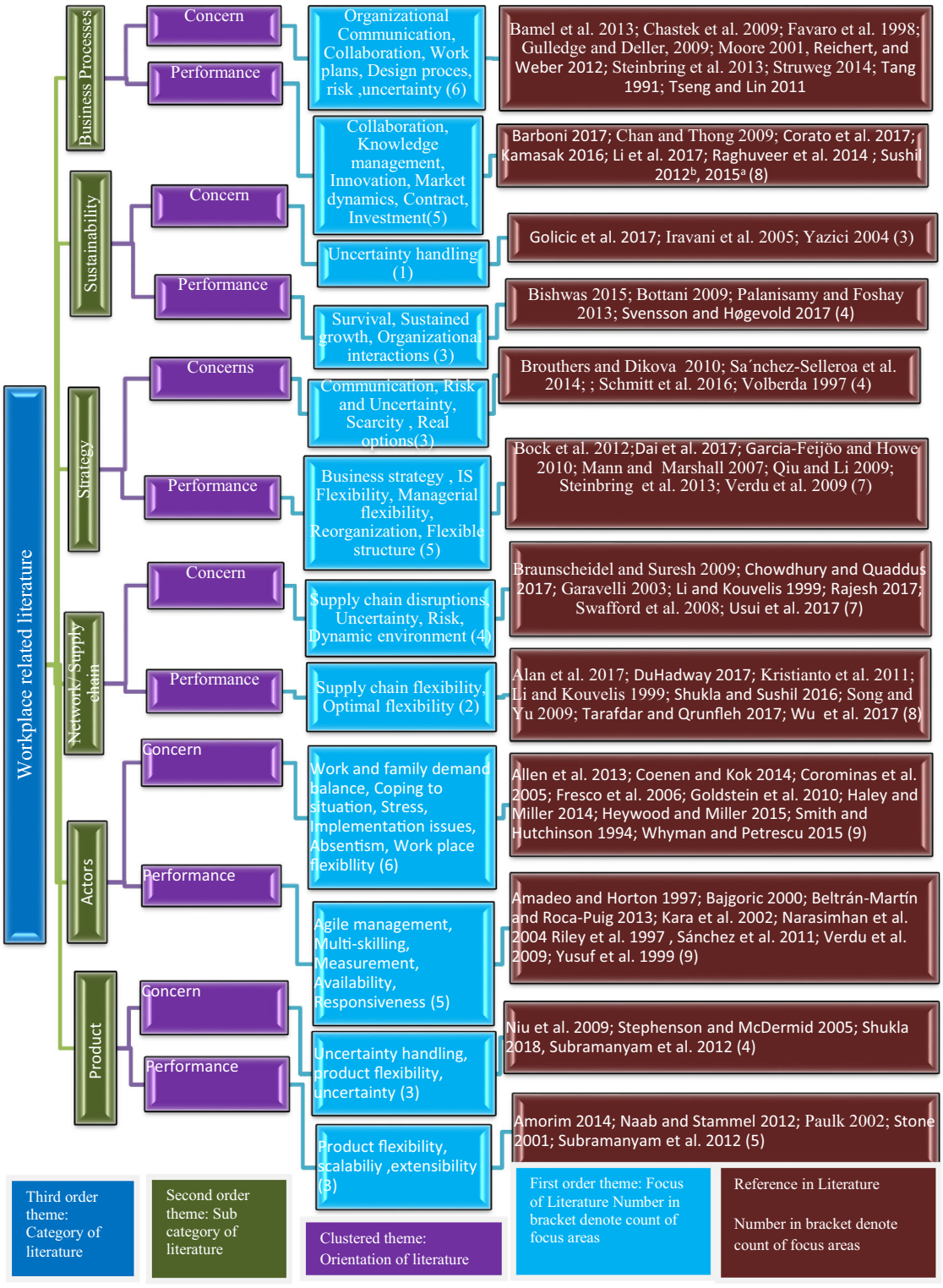


Fig. 2 Summary of the literature for flexibility in the workplace

flexibility enables multiple ways of investment for hedging strategies (Corato et al. 2017). The flexible contract enables the timely decision (Barboni 2017). The flexibilities in business processes, value network, and entire ecosystem impact the organizational performance (Sushil 2012a, 2015a, 2016).

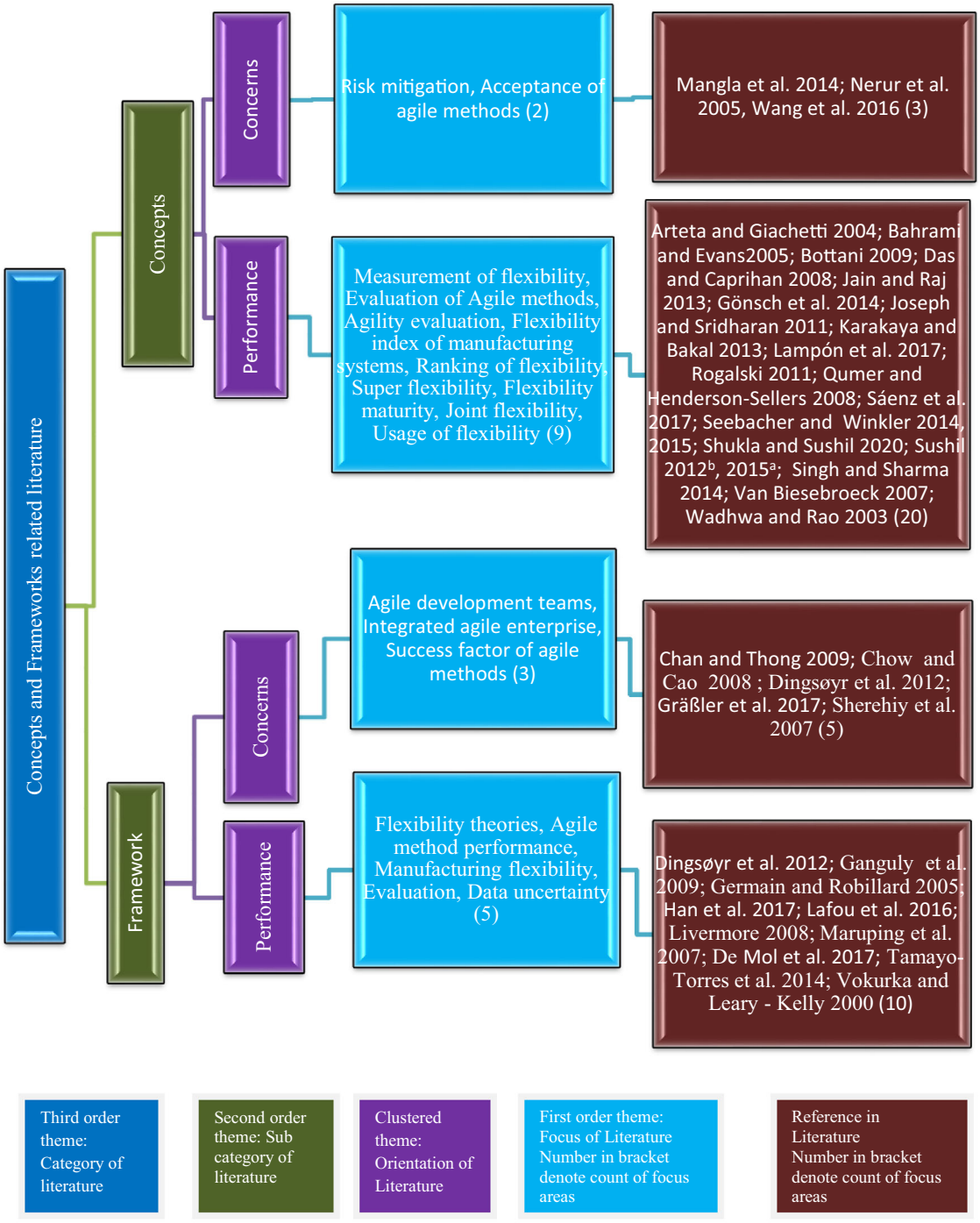


Fig. 3 Summary of the literature for flexibility in framework and concepts

Second-Order Theme: Sustainability

The sustained growth is the long-term objective of flexibility. During the volatility of demand, the sustenance of supplies requires the organization to have volume, mix, routing, and labor flexibilities (Yazici 2005). The proper utilization of capacity and suitable system structure is recommended for cushioning against uncertainty (Iravani et al. 2005). Sustainability through experimentation and resources allocation is advised by Golicic et al. (2017). The flexibility provides strength to an organization in the form of vitality, survival, and growth (Bishwas 2015). The result-oriented interactions among organizational functions facilitate sustained growth. For the flexible enterprise, the

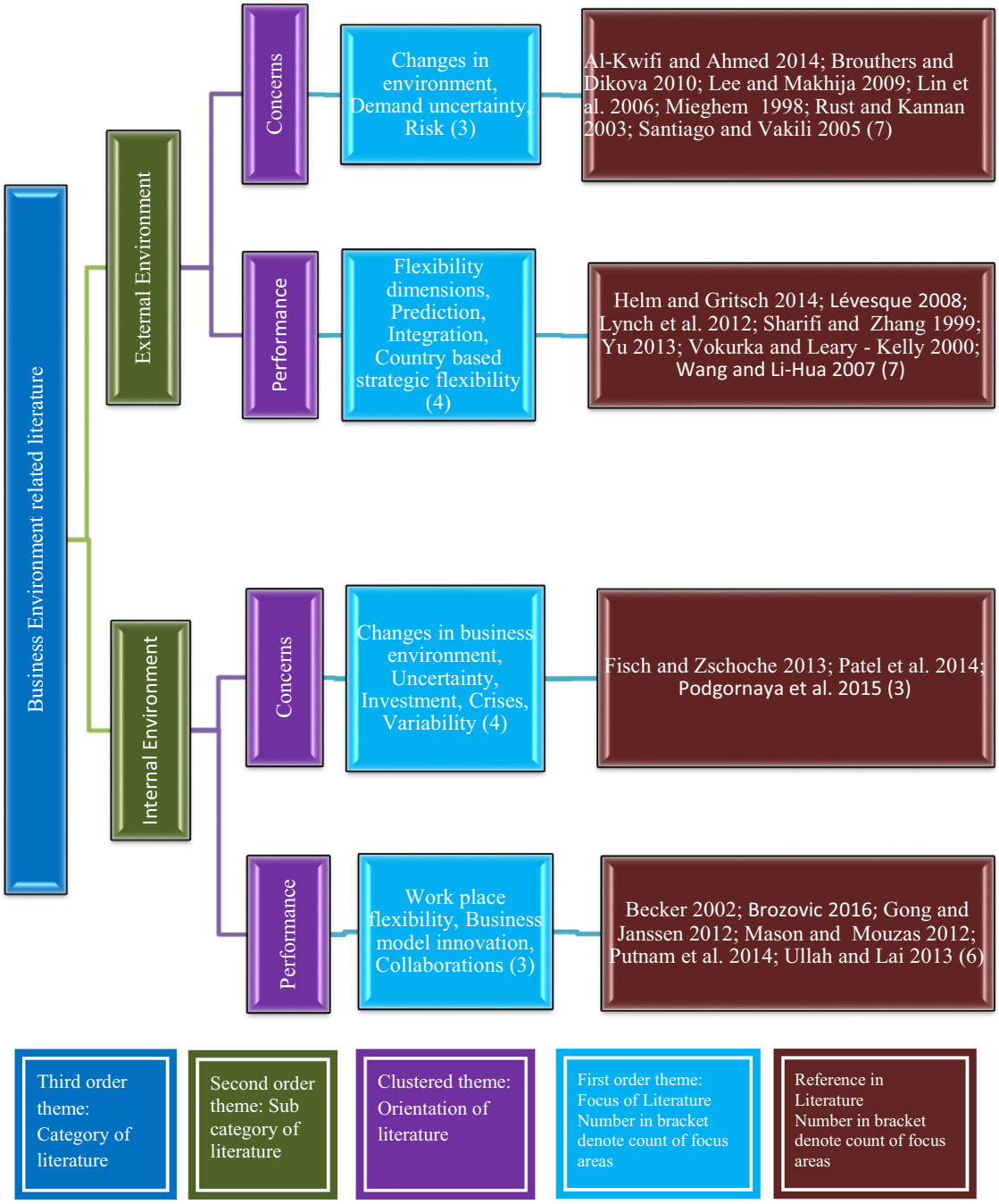


Fig. 4 Summary of the literature for flexibility in a business environment

interdependence of quality parameters, competitive bases, several types of flexibilities, and agility attributes has been noticed (Bottani 2009). Palanisamy and Foshay (2013) noticed that internal flexibility is necessary for information technology infrastructure exploitation. The business sustainability effort evolves over a period of time through adaptation to change (Svensson and Høgevold 2017).

Second-Order Theme: Strategy

The strategy and flexibility are jointly addressed in the literature under strategic flexibility and alignment. The strategic flexibility enhances the ability of the organizations to cope up with the business environment by changing the strategies. The flexible structure emerges from strategic flexibility. The flexible structure is facilitated by

good communication (Sánchez-Sellero et al. 2014). The new product and market combinations are strongly recommended for uncertainty and strategic flexibility (Volberda 1997). The reorganizational capability in resource allocations in the value network is required for showing strategic renewal behavior under scarcity (Schmitt et al. 2016). The investment in real options is required to mitigate the high level of uncertainty for strategic flexibility (Brouthers and Dikova 2010).

The organizational and system structures provide significant support to strategic flexibility. The information system flexibility and visibility positively impact business strategic factors (Qiu and Li 2009). The managerial flexibility is the reasonably good indicator of overall flexibility which top management strives to achieve. The contractual provisions enhance managerial flexibility that benefits the organization (García-Feijóo and Howe 2010). The system structure, formalization, and strategic flexibility are important for uncertainty mitigation (Steinbring et al. 2013). There are encouraging effects of structure and culture on the business model and strategic flexibility (Bock et al. 2012). The flexible organizations show the resource reorganization capability in value network. The benefit of team dynamics and reorganization has been noticed at national level also (Mann and Marshall 2007). Flexible structures are less formal. The organic structure and less formalization are crucial to organizational flexibility (Verdu et al. 2009). The utilization of loosely coupled resources is recommended for strategic flexibility (Dai et al. 2017).

Second-Order Theme: Network/Supply Chain

In the supply chain area, there is ample literature to visualize multiple facets of flexibility. The integration of information systems in supply chain positively impacts flexibility (Swafford et al. 2008). The cultivation of agility for risk management allows a firm to answer rapidly to marketplace variations as well as to anticipate the interruptions in the supply network (Braunscheidel and Suresh 2009). The contract for firms in terms of a number of units without specifying an exact time to purchase provides benefit in the uncertain situation (Li and Kouvelis 1999). The supply chain flexibility is advised in three stages for start-up firms by building relational governance with the suppliers through exercising dynamic economic power concentration, secondly by providing technical support to suppliers, and thirdly by enforcing non-exclusivity arrangements (Usui et al. 2017). The supply chain agility has been advised for competitive advantage through the integration of processes, information, and strategic alliances (Wu et al. 2017). The risk and vulnerability are emerging as a major concern which is mitigated through robustness, resilience, and supplier development (Behzadi et al. 2017; Kurniawan et al. 2017). The resilience can be improved by instituting planning and design capability proactively and reactively (Chowdhury and Quaddus 2017; Rajesh 2017).

The flexibility helps the industry sectors and organizations for business performance. Multiple options in sourcing demand flexibility in network for large industry sector (Shukla and Sushil 2016a, b). There is a positive influence of advance planning on the flexibility of supply chain (Kristianto et al. 2011). The supply network elements consist of logistics, distributor, integration platforms, and sourcing activities. Different supply chain configurations are possible with available suppliers. The same configurations can cost differently for different organizations. The organization may exercise limited, total flexibility or optimal flexibility by including needed supply chain elements (Garavelli 2003). The trust among supply chain partners also plays a significant role in enabling flexibility in the supply chain. There is a positive association between flexibility, trust of distributors, and legitimacy in the supply network (Song and Yu 2009). The supply chain flexibility has a direct and noteworthy impact on strategic and manufacturing flexibilities which affect the firm performance (Alan et al. 2017). The flexibility in the supply network helps to improve the information systems capability (Tarafdar and Qrunfleh 2017). DuHadway et al. (2017) analyzed the source of risk mitigation strategies and supply chain interruptions for recovery of operations.

Second-Order Theme: Actor

The flexibility context in actors includes its enabler and associated benefits. Actor flexibility is the output of many organizational endeavors and integral to managerial flexibility. The main actors for the organization are the workforce along with suppliers, partners, and customers. The work interference with the family has an impact on employees in the form of flexitime, flexiplace, and availability (Allen et al. 2013). On gender equality in management, Smith and Hutchinson (1994) made the observations regarding the participation of women and their concerns for flexibility of accommodation of family demands along with the ongoing job. The human resource policies affect the individual flexibility which further affects other flexibilities of the organization. The coping and explanatory flexibilities affect the depression and anxiety symptoms of employees (Fresco et al. 2006). The flexibility of individuals and stress is impacted by the provision of short leave to the employee which relieves the stress and sleep difficulties (Haley and Miller 2014). The impact of human resource in agile methods finds considerable attention. The ground management strategies for agile execution are

especially encouraged (Goldstein et al. 2010). The workforce flexibility is related to the adaptability to move between different jobs necessitating skill in different domains. For workers, family-oriented practices, flexitime, and job sharing practices result in low absenteeism (Whyman and Petrescu 2015). The family-friendly policies and flexible schedules especially for women employees enhance productivity (Heywood and Miller 2015). Coenen and Kok (2014) noted the importance of flexibility in the workplace resulting in new product development. The concept of annualized hours can effectively be used for operative flexibility and work hours planning (Corominas et al. 2005).

The significance of flexibility in the context of the actor is noted by multiple authors. Sound human resource practices are required for building skill and behavioral flexibility (Beltrán-Martín and Roca-Puig 2013). The other important actors for organizations are suppliers, partners, and customers. Bajgoric (2000) emphasized the efficacy of information sharing among stakeholders for agile management. The participation of outside actors and their responsiveness are required for flexible enterprises (Narasimhan et al. 2004). For flexibility in actors, Verdu et al. (2009) analyzed the measurement of responsiveness by defining the fit between organizational and contextual variables. They identified key elements to determine flexibility; these are time, cost, and intention. The measurement of flexibility is based upon these factors and their derivatives like speed, rapid deployment, cost involved, and defensive or proactive measures for intention. The attributes of agility, i.e., multi-skilling and concurrent execution, are attributed to workforce flexibility (Yusuf et al. 1999). The evaluation of functional flexibility is related to the adaptability of employees to handle different tasks and perform successfully (Riley and Lockwood 1997). The flexible human resource practices promote innovativeness under environmental dynamism (Martínez-Sánchez et al. 2011). The actor's flexibility is primarily envisaged for productivity (Amadeo and Horton 1997). The human flexibility is impacted by external factors resulting from turbulent trading conditions. The human resource practices impact the inter flexibility elements, i.e., organizational structure, technology, and information system (Kara et al. 2002).

Second-Order Theme: Product

The flexibility associated with the product helps the organization to cater for different kinds of markets and customers. The software products are evolving at a fast pace as compared with other sectors. The requirement aiding the product flexibility can be extracted by deriving indicators of customer uncertainty (Stephenson and McDermid 2005). Product flexibility helps common architecture exploitation to generate multiple products (Shukla 2018). The requirement of variety of users enhances the product flexibility and handles the uncertainty (Niu et al. 2009). The granularity of software promotes efficiency and flexibility (Subramanyam et al. 2012). The flexibility in products can be defined and measured using a mathematical framework by utilizing the parameters like extensiveness and scalability which enhances the functional flexibility (Stone 2001). The product flexibility measures extensibility and scalability that further impacts the business issues in ecosystem (Amorim et al. 2014). The flexibility-aware plans with real-time feedback provide faster improvement opportunities (Naab and Stammel 2012). The technical design parameter also improves the agility (Paulk 2002).

Third-Order Theme: Concepts and Frameworks of Flexibility

The large chunk of the literature is devoted to flexibility concepts and frameworks. It is categorized into two second-order themes and summarized in Fig. 3. The second-order theme, i.e., concept, enhances flexibility theories. Another second-order theme, i.e., framework, deals with strategic advice and operational strategies to enhance agility and overcome hurdles in organizational adoption. The clustered themes deal with performance and concerns.

Second-Order Theme: Concepts

The theories explore the relationship between organizational performance parameters and motives of flexibility. The conceptual papers explore the origin of organizational flexibility. The concern in conceptual papers points to need for an operational method for enhancing flexibility. Mangla et al. (2014) noticed that a flexible framework is required for risk mitigation strategy in the supply chain. The limited bureaucracy and formalism along with organization cultures are conducive to innovation and required agile methodology acceptance in IT organizations (Nerur et al. 2005). The flexibility analysis of process operation and supply chain requires a mechanism to measure several types of flexibilities at a conceptual level (Wang et al. 2016).

The application of agile concepts in the organization enables responsiveness; however, the cost implication for such initiatives is also noticed (Van Biesebroeck 2007). The choice of agile methods among available options for organizations shall be properly evaluated using criteria and tools for adoption (Qumer and Henderson-Sellers 2008). There are mechanisms to evaluate the agility in the organization (Arteta and Giachetti 2004). The measurement of specific flexibilities also finds a reasonable place in the

literature. Das and Caprihan (2008) formulated a mechanism to evaluate manufacturing flexibility by quantifying underlying factors and proposed flexibility index of manufacturing systems. A generic mechanism to evaluate the organization-wide flexibility is also feasible in different maturity levels (Shukla and Sushil 2020a, b). The comprehensive framework of evaluation and measurement of select flexibility for the production system is provided by Rogalski (2011). Singh and Sharma (2014) identified the flexibility factors and used them for developing a framework to measure supply chain flexibility.

The benefit of flexibility initiatives, their combined impact, and interrelationship are conceptualized by authors. Al-Kwifi and Ahmed (2014) described the simulation scenario for the impact of a pallet, and routing flexibilities and their combined impact in the production environment. The concept is enhanced to accommodate the interactions of various types of flexibilities and other organizational functions. The quality parameters, competitive bases, agile attributes, and enablers show mutual interdependence (Bottani 2009). The concept of joint flexibility of the product is elaborated by Karakaya and Bakal (2013) for sourcing of components. Joseph and Sridharan (2011) examined the impact of sequencing and routing flexibilities, and sequencing rules on manufacturing systems and concluded significant relationship among them. The prioritization of initiatives is natural phenomena where cost is expected to be incurred. It has been noticed that program flexibility ranks the highest and production flexibility ranks the lowest in flexible manufacturing system (Jain and Raj 2013).

The flexibility maturity model further enhances concepts. It categorizes organizational flexibility into maturity levels for comprehensive analysis (Wadhwa and Rao 2003; Sushil 2012a, 2015a). The more intense concept of super-flexible organization has been put forward to mitigate any kind of uncertainty by changing the business model of organizations within a very short time (Bahrami and Evans 2005). The contingency view of flexibility analyzes the upstream and downstream relationship with suppliers and customers (Sáenz et al. 2017). The flexibility relates to an observable aspect of organizational indicators of performance. The lead time, cost deviations in process chain, efficiency, and performance have been utilized to measure the dyadic supply chain and manufacturing flexibilities (Seebacher and Winkler 2014, 2015). Gönsch et al. (2014) advised usage of product flexibility for revenue management using the deterministic linear program. The relocation theories about operational flexibility and decision making have been given by Lampón et al. (2017).

Second-Order Theme: Framework

The framework-oriented literature relates to the application and acceptance of agile methods. It also addresses the managerial concern in agile method adoption along with advice to management. Dingsøyr et al. (2012) studied the performance of various agile methodologies in a decade starting in the year 2000 that are mainly related to concerns and acceptance of different agile methods. Information technology and manufacturing organizations emphasize on agile methodologies. The issues related to acceptance of agile methods for software development is a special concern in the literature (Chan and Thong 2009). The integrated view of agile enterprise has been the major concern due to multiple agile methodologies and characteristics of organizations (Sherehiy et al. 2007). A large chunk of the literature identifies the success factors of agile methodology. It covers the wide range of performance factors that include quality, scope, time, engineering techniques, cost impacting delivery strategy, project management process, team capability, customer involvement, and team environment (Chow and Cao 2008). The decoupling of production system environment with products characteristics leads to flexibility in entire systems (Gräßler et al. 2017).

The agile methods are further enhanced to manage uncertainty, risk, and performance. Managing the changing user requirements using controls in agile methods, building effective capability, and technology exploration for manufacturing flexibility have been addressed in the literature (Vokurka and Leary Kelly 2000; Maruping et al. 2009; Tamayo-Torres et al. 2014). All the flavors of agile methods cause the organizations to focus on core and productive work and enhance the organizational size (Germain and Robillard 2005, Livermore 2008). These benefits can potentially attract higher management attention. The flexible way of delivery and investments leads to cost-effectiveness, increased market share, responsiveness, better products, flexible interface, and good relationship (Ganguly et al. 2009, Lafou et al. 2016). The frameworks are extended to identify the flexibility aspects emerging due to uncertain and probabilistic data for extracting the information for decision making (De Mol et al. 2017). The theoretical model also includes operational, strategic, and transactional IT flexibility in the supply chain (Han et al. 2017).

Third-Order Theme: Business Environment

The business model innovation is the result of careful scanning of the environment. The organizations often do investments in other countries, which is influenced by organizational strategy in a given business environment. Some literary contributions suggest the use of crises management methods in the highly turbulent environment. The collaborations, use of network, and entrepreneurship are an important part of crises management endeavor. Collaboration and integration are compulsory for

predictable and effective strategy under uncertainty. The second-order themes related to the business environment are categorized into the internal and external environment impacting organizational strategy and flexibility together. Figure 4 summarizes all the themes pertaining to the business environment.

Second-Order Theme: External Environment

The investments are required to institute flexibility and hedge against uncertainty (Mieghem 1998). The strategic flexibility is associated with the acquisition of real option under demand uncertainty (Brouthers and Dikova 2010). The investment in R&D projects provides increased payoff and value under uncertainty (Santiago and Vakili 2005). The e-service is emphasized to stay competitive in the turbulent business environment (Rust and Kannan 2003). The organizational resources and decision making mandate the integration process to be very effective for incorporating flexibility in the strategy process. Lin et al. (2006) recommended a collaborative relationship as a strategy for the agile enterprise. There are case studies on global brands for being successful in crises (Al-Kwifi and Ahmed 2014). The investment to enhance the agility is especially noted in the literature. The foreign and export investments have an optimistic impact on organization flexibility and help in economic crises (Lee and Makhija 2009).

The business environment impacts the organizational processes for flexibility assimilation. There is a significant relationship between the external environment and flexibility (Yu 2013). The agile manufacturing recommends management of change in the business environment for a strategic response (Sharifi and Zhang 1999). There is a strong association among manufacturing flexibility with strategy, organizational structure, environmental uncertainty, and technology development (Vokurka and Leary Kelly 2000). The business aspects indicated by markets and products have a strong association in steady and unstipulated business environment (Lynch et al. 2012). The international entrepreneurship provides better ways to deal with uncertain environment than the use of networks (Helm and Gritsch 2014). The country-specific measures of strategic flexibility in an auto part industry in China are discussed by authors (Wang and Li-Hua 2007; Lévesque 2008).

Second-Order Theme: Internal Environment

The internal environment-related literature deals with options available to management that impacts the organization in whole or in parts. Risk management, uncertainty mitigation, prediction, and effective entrepreneurship are important for organizational flexibility. Podgornaya et al. (2015) studied the flexibility with a choice of business models and recommended strategies for risk reduction. Under increasing erraticism, entrepreneurship provides better performance (Patel et al. 2015). The organizations can predict the establishment of a new site based upon the net present value, growth option value, operational flexibility, and value of the existing production network (Fisch and Zschoche 2012).

There is a positive impact of initiatives for flexibility on factors like workplace planning and design (Becker 2002), workplace flexibility initiatives (Putnam et al. 2014), business model innovation (Mason and Mouzas 2012), organization's resources (Gong and Janssen 2012), etc. The utilization of information systems for the changed business scenario is highly recommended (Ullah and Lai 2013). Brozovic (2016) analyzed strategic flexibility in the context of uncertainty through a literature review and given a conceptual framework and means of evaluation.

Interpretive Synthesis

The concern and performance of all three categories of the literature have been analyzed, and common focus areas are noted. The process and issues in the implementation of flexibility initiatives denote the concerns and benefits due to flexibility initiatives which denote the performance. The concepts and framework-related papers are dominated by agile methodology. In the business environment category, it is observed that concerns are mostly related to uncertainty, crises, and risk management in first-order themes, i.e., focus of the literature. The mitigation of risk, uncertainty, and overcoming of hurdles in the adoption of the agile methodology are primarily being dealt with concerns in "concepts and frameworks." In the workplace category and concern-related literature, a large chunk is devoted to handling uncertainty, stress, risk, supply chain disruption, and organizational communication using flexibility. The common concerns are uncertainty and risk mitigation in "business environment," "concepts and framework," and "workplace" categories. Figure 5 represents the views emerged from an analysis of concern-related contributions.

The performance-related literature of business environment encompasses the flexibility aspects of business model innovation, integration, and collaboration. The success factors of agile methodology acceptance, performance, expansion of framework, evaluation, and raking of flexibility are dealt in the performance category of "concepts and frameworks." The inclusive performance of the organizational function, collaboration, design process, knowledge management, reorganization, flexible strategies, managerial flexibility, supply chain flexibility, agile methodology, multi-skilling, and measurement of

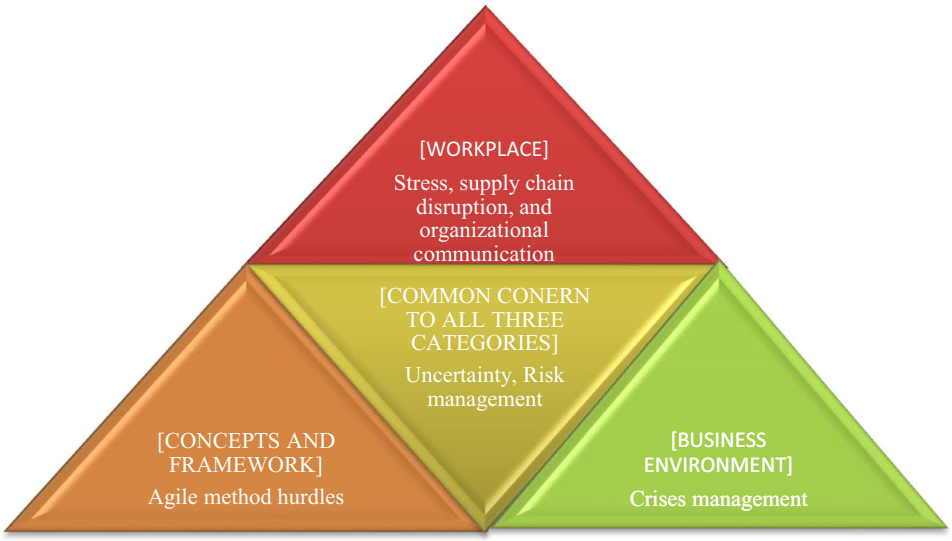


Fig. 5 Consolidation of the literature pertaining to concerns for all three categories

flexibility find a place in workplace-related literature. The enhancement of flexibility theories, measurement, and ranking is common to all three categories of the literature in performance orientation. Figure 6 offers consolidated view of common performance under all three categories.

The commonality and differences are also analyzed quantitatively. The number of contributions and context in different order themes is visually depicted in Fig. 7. The inferences are made for validating the flexibility paradox.

Discussion on Managerial Paradox

The managerial flexibility touches all spheres of operations and strategy and theoretical and practical knowledge base. These are comprehensively captured under various first-order themes. The first-order theme in category "Workplace" is organizational communication, collaboration, work plans, design process, knowledge management, innovation, market dynamics, contract, investment, survival, sustained growth, organizational interactions, communication, risk, scarcity, business strategy, information systems (IS) flexibility, managerial flexibility, reorganization, flexible structure, supply chain disruptions, risk, dynamic environment, supply chain flexibility, optimal flexibility, work and family demand balance, coping with situation, stress, implementation issues, absenteeism, workplace flexibility, agile management, multi-skilling, measurement, availability, responsiveness, uncertainty handling, product flexibility, scalability, and extensibility. Clustered themes pertain to performance and concerns. The

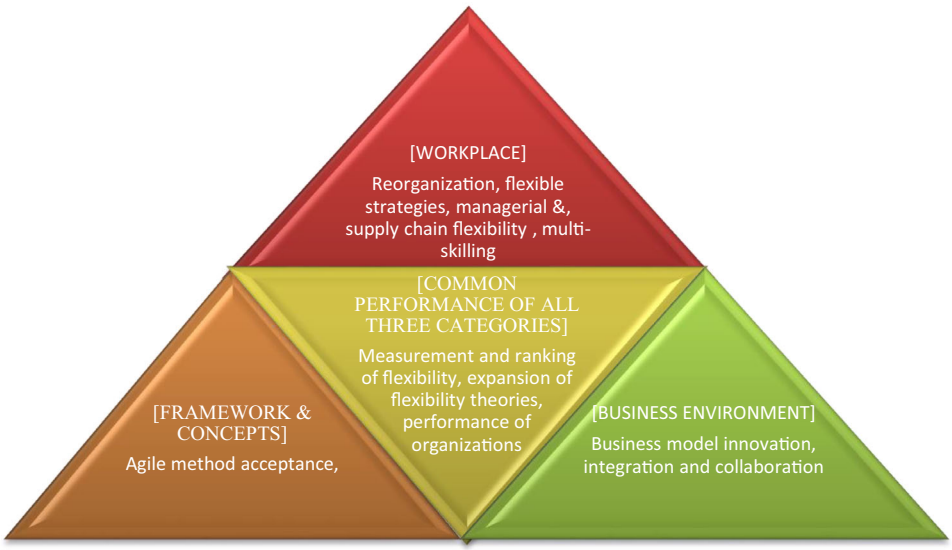


Fig. 6 Consolidation of the literature pertaining to performance for all three categories

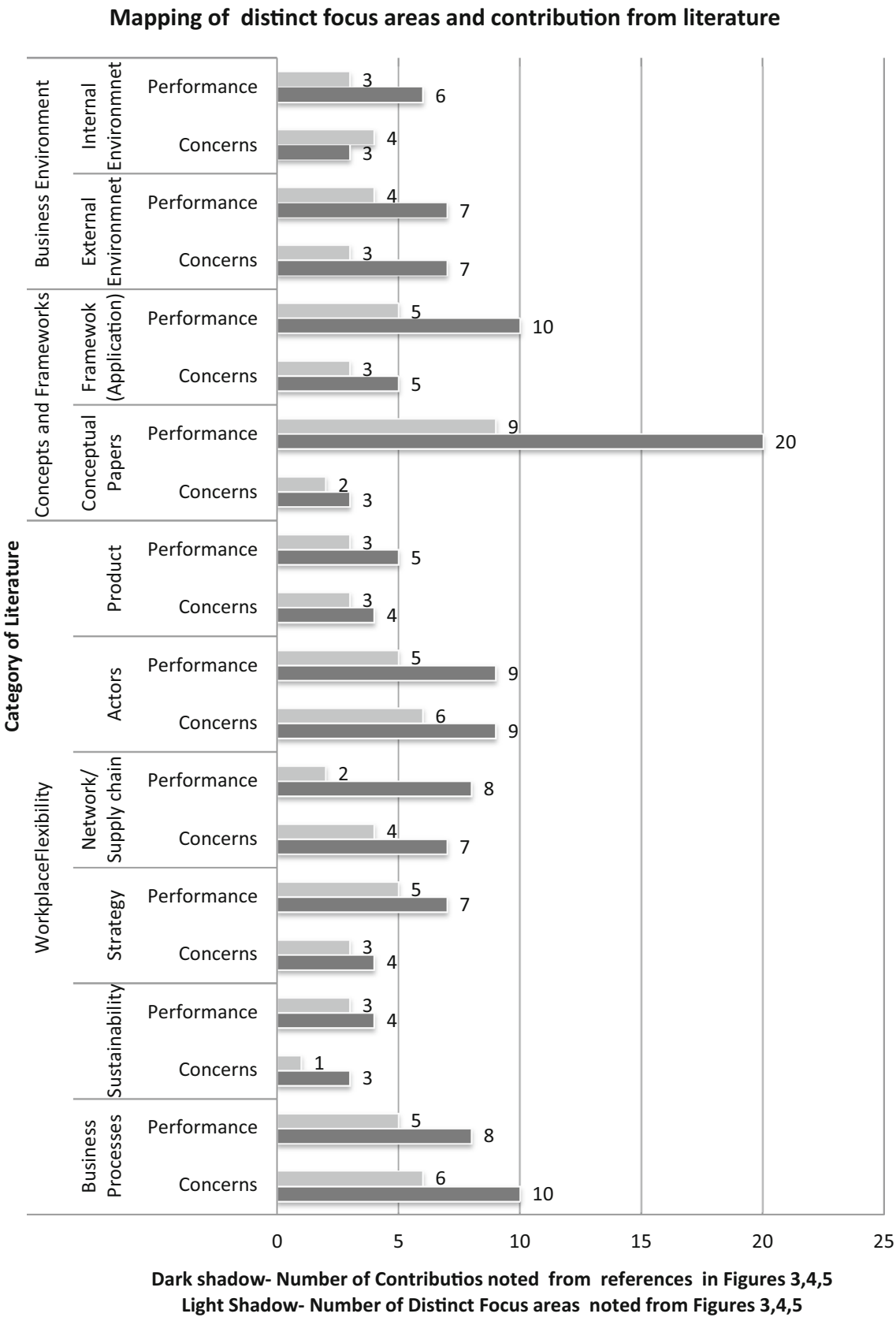


Fig. 7 Mapping of distinct focus areas and contribution in the literature

second-order themes are business processes, sustainability, strategy, network/supply chain, and actors.

The first-order themes of "Concepts and Framework" are risk mitigation, acceptance of agile methods, measurement of flexibility, evaluation of agile methods, agility evaluation, flexibility index of manufacturing systems, ranking of flexibility, super-flexibility, flexibility maturity, joint flexibility, usage of flexibility, agile development teams, integrated agile enterprise, success factor of agile methods, flexibility theories, agile method performance, manufacturing flexibility, and data uncertainty. The clustered themes examine the relationship with performance and concern of flexibility. Two themes (second-order) are identified in this category.

The first-order themes in category "Business Environment" are changes in the environment, demand uncertainty, risk, flexibility dimensions, prediction, integration, country, country-based strategic flexibility, changes in the business environment, uncertainty, investment, crises, variability, workplace flexibility, business model innovation, and collaborations. The clustered themes point to the relationship of flexibility reported in the literature w.r.t performance and concern. The second-order themes relate to the internal and external perspectives of organizations.

The common concerns and performance for all three categories are used to substantiate the flexibility paradoxes. The quantitative and qualitative analyses of commonalities lead to inferences about first and second managerial paradoxes, respectively.

First Managerial Flexibility Paradox

The observations are noted from Fig. 7. The contributions in performance orientation are significantly higher as compared with concern in "Concept and Framework" category. In subcategory "Actors" and "External environment," an equal number of contributions are observed in both performance and concern orientations because the contributions are identified from human resource and business environment literature. In subcategory, i.e., "Business process," the number of contributions related to concern is higher. The contributions in performance orientations are asymmetrically higher as compared with concern in most of the subcategories. It is observed that these subcategories primarily relate to operations and strategy. It is also noted that subcategories, where contributions in concerns are equal or higher, require the involvement of senior management for investment toward flexibility. The generic behavior of a lower number of contributions in concern generates the first flexibility paradox which is built on the assumption of limited advice to implement the advanced flexibility concepts.

Further investigation reveals that risk and uncertainty-related literature find considerable support in non-flexibility and practice-oriented literature. The widely utilized methodologies such as capability maturity model and project management frameworks put significant emphasis on risk and uncertainty mitigation through standard operating procedures (PMI 2009; CMMI-DEV 2010; Hopkin 2017). For risk and uncertainty management, the available literature points to usage of tools and techniques (Raz and Michael 2001). The context in which the risk and uncertainty management to be used is also echoed in the literature (Ward and Chapman 2003; Cicmil et al. 2017). Enough literature is available to showcase the benefit of risk management for project success (De Bakker et al 2010). The impact of risk on decision process is elaborated in the literature (Froot and Stein 1998; Cole et al. 2017). The extreme form of the use of risk management in case of failures is also noted (Chopra and Sodhi 2004). This shows that integrated view has emerged in area of risk management only that expressively widens the first managerial paradox.

Second Managerial Flexibility Paradox

Second paradox toward flexibility related to the difference in approach toward flexibility in higher and lower management. The different topics covered in the literature are identified as focus areas. The number of focus areas in performance orientation is more than concern orientation in most of the subcategories. This creates a different perception of flexibility to top and lower management and is evident in Fig. 7. The multiple contributions deliberate the same flexibility aspects which narrow the scope of discussion. The central focus of performance and concerns noted from Figs. 5 and 6 is the enhancement of flexibility theories and risk management. This denotes a wide gap between theory and practice of flexibility hinting at the second managerial paradox.

The context associated with the product, people, and processes for managerial direction for flexibility has unique benefits. The strong relationship of business strategy, flexibility, and performance has relevance to top management (Gupta and Somers 1996). The requirement of organizational flexibility is recognized by top management firstly by scanning the business environment and secondly through literature support. The flexibility adoption requires standard procedures to be followed up. Pailles et al. (1997) noted that decision on flexibility is taken at higher levels,

while it is promulgated in parts to the working level where the empirical evidence exists for implementation. The need for flexibility from lower managerial functions is mandated if an organization is operating efficient value network. Manders et al. (2016) found that in supply chain context the organization's endeavors for flexibility cater to direct need of next member. The non-availability of clear guidelines will create indifference to top management advice for flexibility. Wang et al. (2015) noted that top management's shared vision neither facilitates nor impedes the organizational flexibility; however, it positively influences the technological flexibility. The discussion points to indifference of managerial endeavor for implementing the concepts of flexibility understood by top management and supports second managerial paradox.

Third Managerial Flexibility Paradox

The third paradox is related to the extent and scope of flexibility initiatives and finding optimality. There are various modes of flexibility initiatives; some are emergent and others are pre-planned strategic measures. The limit of flexibility measures is decided in terms of how far and a fast current system can cope (Slack and Correa 1992). Treville et al. (2007) found that 40% of flexibility improvement initiatives are unsuccessful. This required ample resources in hand for organizations to venture for flexibility initiatives. But once the initiatives are successful, the competitive advantage comes in. Even the flexibility as capability to handle risk mitigation and innovation is also questioned in the literature (Sushil 2017; Jena and Memon 2018). The strong focus on the flexibility literature toward the benefits and lower success rate of flexibility initiative will create the paradoxical situation for management practitioners who will look for optimality of resource allocation for this purpose. This creates the third managerial paradox and holds back the flexibility initiatives.

Conclusion, Limitations, and Future Research

Managerial flexibility is the ability to adapt to situations w.r.t time and scale to take advantage of business decisions (Business Directory n.d). The most cited benefit is investment and cost decision. It impacts the relationship between investment and uncertainty (Xie 2009). The clearly defined construct of managerial flexibility is largely missing in the literature. The uncertainty and risk are other focus areas that eventually relate to cost optimization. The manifold options reflect the managerial flexibility and evident from numerous contexts identified in first-order themes. Still these options lead to option pricing analysis and real options investment. The important point is ignoring long-term goals enabled by managerial flexibility in the literature. Flexibility is envisaged for vitality and sustainability (Sushil 2011). Very few contexts in the first-order themes relate to long-term agenda. This necessitates broadening the ambit of managerial flexibility through setting the research agenda. A highly relevant subject pertains to the execution of project orders where impact of managerial flexibility is ignored. This is clearly visualized by absence in first-order themes. For managerial action and attention toward flexibility, it is advised to look in concern-related contribution. This will help them to exploit the benefits of flexibility by deploying envisaged resources. The dealing of uncertainty with flexibility has resulted in many contributions using metric-based approaches; these are approaching maturity. The contributions relating managerial flexibility with real options decisions are evolving rapidly with potential of maturity in the near future. The real options evaluations along with valuations of acquisitions and returns are noted as important indicator of managerial flexibility (Jain et al. 2018). A closed-loop behavior is seen for flexibility benefits and its enablers. The commonly understood benefit of managerial flexibility is increased motivation, concentration, and productivity derived from flexible work options (Sweet and James 2013). However, the same contributes to workforce flexibility (Patil and Suresh 2019). The study of the literature suggests that flexibility has been used along with other strategic tools and objectives. The social and organizational theories put flexibility along with structure and culture for performance (Birasnav et al. 2019). This leads to a broad conclusion that it has been used as a mediator of organizational performance. The direct influence of managerial flexibility on performance is largely ignored in the extant literature.

The discussion on paradoxes reveals that two categories of contributions are expected from future work: one is concerned with advice for the strategic purpose for top management, and the other is guidance for operational aspects dealing with flexibility. In this paper, the reason that drives the flexibility has been investigated along with deliberation of future scenarios. The paradoxical behavior of management is built upon this finding where the practicing manager seeks advice from the literature to implement the advanced concepts of flexibility, while the

practices of implementation are focused on risk and, uncertainty mitigation strategies. The non-availability of literary advice to management will result in ignoring the flexibility benefits. The analysis of focus of the literature provides broad dimension of paradox resulting from concentration of contributions in different focus areas. The analysis shows that there is ample literature dealing with uncertainty and risk, while there is limited contribution on severe form of uncertainty in the form of crises, supply chain disruption, and flexibility in organizational communication. Future researchers can focus on these areas. This paper has explored all the aspects of managerial flexibility falling within ambit of organizational boundaries. However, the influence of outside factors impacting the managerial flexibility can be explored by future researchers using thematic analysis.

This paper also suggests future research direction in an area of flexibility adoption leading to organizational competitiveness and performance. The flexibility-related theories start from measuring certain types of flexibility and end up with comprehensive treatment toward maturity models. Some of theories and practices are widely used, while others are still evolving. It is also observed that process definition and application methodology supported by empirical studies for practicing managers are primarily available for risk and uncertainty management functions. The same is lacking in advanced concepts and frameworks of flexibility which make it difficult for practicing managers. The substantial gap opens up a new research area for defining workable methods to practice flexibility. There is evidence of a methodological gap in the application of flexibility and agility in the supply chain (Fayezi et al. 2016). Yu et al. (2015) dealt with operational flexibility and performance outcomes aspects simultaneously. The three paradoxes provide the avenues to prospective researchers to fill the gap between operation and concepts through operational means. The flexibility in business model innovation, integration, collaboration, flexible strategies, and agile method acceptance is identified for further exploration by researchers as limited contributions are seen in these areas. This paper has made inferences instead of conclusions regarding the managerial paradoxes. These inferences can be changed in light of more advanced contributions on the subject. This paper has made an attempt to highlight various aspects of managerial flexibility which is central to overall organizational flexibility but still ignored in practice and theory.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix

Number of contributions and focus areas

Category of the identified literature	Context of the literature	Relevant contributions	Distinct focus areas
Contributions used for finding managerial flexibility paradox			
Workplace			
Business processes	Concerns	10	6
	Performance	8	5
Sustainability	Concerns	3	1
	Performance	4	3
Strategy	Concerns	4	3
	Performance	7	5
Network/supply chain	Concerns	7	4
	Performance	8	2
Actors	Concerns	9	6
	Performance	9	5
Product	Concerns	4	3
	Performance	5	3
Concepts and frameworks			
Concepts papers	Concerns	3	2
	Performance	20	9
Framework (application oriented)	Concerns	5	3
	Performance	10	5
Business environment			
External environment	Concerns	7	3
	Performance	7	4
Internal environment	Concerns	3	4
	Performance	6	3

References

Acharya, S. (2019). Beyond learning outcomes: Impact of organizational flexibility on strategic performance measures of commercial e-learning providers. *Global Journal of Flexible Systems Management, 20*(1), 31–41.

Al-Kwifi, O. S., & Ahmed, Z. U. (2014). CHALCO: Building a global brand while passing industry crisis. *Journal of Technology Management in China, 9*(1), 75–90.

Allen, T. D., Johnson, R. C., Kiburz, K. M., & Shockley, K. M. (2013). Work–family conflict and flexible work arrangements: Deconstructing Flexibility. *Personnel Psychology, 66*(2), 345–376.

Amadeo, E. J., & Horton, S. (1997). Labour flexibility and productivity: An Overview. In E. J. Amadeo & S. Horton (Eds.), *Labour productivity and flexibility* (pp. 1–35). London: Palgrave Macmillan.

Amorim, S. D. S., Almeida, E. S. D., & McGregor, J. D. (2014). Scalability of ecosystem architectures. In *IEEE/IFIP conference on software architecture*, Sydney, NSW, pp. 49–52. https://doi.org/10.1109/wicsa.2014.36.

Arteta, B. M., & Giachetti, R. E. (2004). A measure of agility as the complexity of the enterprise system. *Robotics and Computer-Integrated Manufacturing, 20,* 495–503.

Avazpour, R., Ebrahimi, E., & Fathi, M. R. (2014). Prioritizing agility enablers based on agility attributes using fuzzy prioritization method and similarity-based approach. *International Journal of Economy Management and Social Sciences, 3,* 143–153.

Bahrami, H., & Evans, S. (2005). The need for super-flexibility. In *Super-flexibility for knowledge enterprises*, Springer, Berlin, pp. 1–10. https://doi.org/10.1007/b138009.

Bajgoric, N. (2000). Web-based information access for agile management. *International Journal of Agile Management Systems, 2*(2), 121–129.

Bamel, U. K., Rangnekar, S., Rastogi, R., & Kumar, S. (2013). Organizational process as antecedent of managerial flexibility. *Global Journal of Flexible Systems Management, 14*(1), 3–15.

Barboni, G. (2017). Repayment flexibility in microfinance contracts: Theory and experimental evidence on take-up and selection. *Journal of Economic Behavior & Organization, 142,* 425–450.

Becker, F. (2002). Improving organizational performance by exploiting workplace flexibility. *Journal of Facilities Management, 1*(2), 154–162.

Behzadi, G., O'Sullivan, M., Olsen, T. L., & Zhang, A. (2017). Agribusiness supply chain risk management: A review of quantitative decision models. *Omega, 79,* 21–42.

Beltrán-Martín, I., & Roca-Puig, V. (2013). Promoting employee flexibility through HR practices. *Human Resource Management, 52,* 645–674.

Birasnav, M., Chaudhary, R., & Scillitoe, J. (2019). Integration of social capital and organizational learning theories to improve operational performance. *Global Journal of Flexible Systems Management, 20*(2), 141–155.

Bishwas, S. K. (2015). Achieving organization vitality through innovation and flexibility: An empirical study. *Global Journal of Flexible Systems Management, 16*(2), 145–156.

Bock, A. J., Opsahl, T., George, G., & Gann, D. M. (2012). The effects of culture and structure on strategic flexibility during business model innovation. *Journal of Management Studies, 49*(2), 279–305.

Borland, H., Ambrosini, V., Lindgreen, A., & Vanhamme, J. (2016). Building theory at the intersection of ecological sustainability and strategic management. *Journal of Business Ethics, 135*(2), 293–307.

Bottani, E. (2009). A fuzzy QFD approach to achieve agility. *International Journal of Production Economics, 119*(2), 380–391.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101.

Braun, V., & Clarke, V. (2013). *Successful qualitative research: A practical guide for beginners*. Thousand Oaks: Sage.

Braunscheidel, M. J., & Suresh, N. C. (2009). The organizational antecedents of a firm's supply chain agility for risk mitigation and response. *Journal ofOperations Management*, *27*(2), 119–140.

Brouthers, K. D., & Dikova, D. (2010). Acquisitions and real options: The greenfield alternative. *Journal of Management Studies, 47*(6), 1049–1071.

Brozovic, D. (2016). Strategic flexibility: A review of the literature. *International Journal of Management Reviews., 20*(1), 3–33.

Business Directory. (n.d). Retrieved from http://www.businessdictionary.com/definition/managerial-flexibility.html.

Chan, A. T. L., Ngai, E. W. T., & Moon, K. K. L. (2017). The effects of strategic and manufacturing flexibilities and supply chain agility on firm performance in the fashion industry. *European Journal of Operational Research, 259*(2), 486–499.

Chan, F. K. Y., & Thong, J. Y. L. (2009). Acceptance of agile methodologies: A critical review and conceptual framework. *Decision Support Systems, 46*(4), 803–814.

Chastek, G. J., Donohoe, P., & McGregor J. D. (2009). Formulation of a production strategy for a software product line. *CMU/SEI TN-025*, 1-25, Retrieved from https://resources.sei.cmu.edu/asset_files/TechnicalNote/2009_004_001_15080.pdf.

Chopra, S., & Sodhi, M. S. (2004). Managing risk to avoid supply-chain breakdown. *MIT Sloan Management Review, 46*(1), 53–61.

Chow, T., & Cao, D. B. (2008). A survey study of critical success factors in agile software projects. *The Journal of Systems and Software, 81*(6), 961–971.

Chowdhury, M. H., & Quaddus, M. (2017). Supply chain resilience: Conceptualization and scale development using dynamic capability theory. *International Journal of Production Economics, 188*(C), 185–204.

Chowdhury, M. M. H., Umme, N. J., & Nuruzzaman, M. (2018). Strategies for mitigating supply-side barriers in the apparel supply chain: A study on the apparel industry of Bangladesh. *Global Journal of Flexible Systems Management, 19*(Suppl 1), 41–52.

Chowdhury, R., & Jangle, N. (2018). Critical systems thinking towards enhancing community engagement in micro-insurance. *Global Journal of Flexible Systems Management, 19*(3), 209–224.

Cicmil, S., Cooke-Davies, T., Crawford, L., & Richardson, K. (2017). Exploring the complexity of projects: Implications of complexity theory for project management practice. *Project Management Institute*. Retrived from: https://www.pmi.org/learning/academic-research/exploring-the-complexity-of-projects-implications-of-complexity-theory.

Clarke, T. (2002). A paradox of governance convergent policy and divergent practice in corporate governance in Asia. In S. R. Clegg (Ed.), *Management and organization paradoxes, advances in organization studies*, vol. 9, pp. 247–273.

CMMI-DEV V1.3. (2010). Capability maturity model integration for development version 1.3., 2010. Improving processes for developing better products and services. Technical report.CMU/SEI-2010-TR-033, Software Engineering Institute.

Coenen, M., & Kok, R. A. W. (2014). Workplace flexibility and new product development performance: The role of telework and flexible work schedules. *European Management Journal, 32*(4), 564–576.

Cole, S., Giné, X., & Vickery, J. (2017). How does risk management influence production decisions? Evidence from a field experiment. *The Review of Financial Studies, 30*(6), 1935–1970.

Corato, L. D., Moretto, M., & Rossini, G. (2017). Financing flexibility: The case of outsourcing. *Journal of Economic Dynamics and Control*, *76*(C), 35–65.

Corominas, A., Lusa, A., & Pastor, R. (2005). Annualised hours: A real flexibility tool. *OR Insight, 18*(1), 10–14.

Couchman, P. K., & Fulop, L. (2002). The meanings of risk and interorganizational collaboration. In S. R. Clegg S.R. (Ed.), *Management and organization paradoxes*, *advances in organization studies*, Vol. 9, pp. 41–64.

Dai, Y., Goodale, J. C., Byun, G., & Ding, F. (2017). Strategic flexibility in new high-technology ventures. *Journal of Management Studies, 55*(2), 265–294.

Das, A., & Caprihan, R. (2008). A rule-based fuzzy-logic approach for the measurement of manufacturing flexibility. *The International Journal of Advanced Manufacturing Technology, 38*(11–12), 1098–1113.

De Bakker, K., Boonstra, A., & Wortmann, H. (2010). Does risk management contribute to IT project success? A meta-analysis of empirical evidence. *International Journal of Project Management, 28*(5), 493–503.

De Leeuw, A. C. J., & Volberda, H. W. (1996). On the concept of flexibility: A dual control perspective. *Omega, 24*(2), 121–139.

De Mol, R., Bronselaer, A., & De Tré, G. (2017). Evaluating flexible criteria on uncertain data. *Fuzzy Sets and Systems, 328,* 122–140.

Deng, Z., Guo, H., Zhang, W., & Wang, C. (2014). Innovation and survival of exporters: A contingency perspective. *International Business Review, 23*(2), 396–406.

Dingsøyr, T., Nerur, S., Balijepally, V., & Moea, N. B. (2012). A decade of agile methodologies: Towards explaining agile software development. *The Journal of Systems and Software, 85*(6), 1213–1221.

DuHadway, S., Carnovale, S., & Hazen, B. (2017). Understanding risk management for intentional supply chain disruptions: Risk detection, risk mitigation, and risk recovery. *Annals of Operations Research, 2017,* 1–20.

Ebben, J. J., & Johnson, A. C. (2005). Efficiency, flexibility, or both? Evidence linking strategy to performance in small firms. *Strategic Management Journal, 26*(13), 1249–1259.

Eppink, D. J. (1978). Planning for strategic flexibility. *Long Range Planning, 11*(4), 9–15.

Favaro, J. M., Favaro, K. R., & Favaro, P. F. (1998). Value based software reuse investment. *Annals of Software Engineering, 5*(1), 5–52.

Fayezi, S., Zutshi, A., & O'Loughlin, A. (2016). Understanding and development of supply chain agility and flexibility: A structured literature review. *International Journal of Management Reviews, 19*(4), 379–407.

Fereday, J., & Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. *International Journal of Qualitative Methods, 5*(1), 80–92.

Fisch, J. H., & Zschoche, M. (2012). The role of operational flexibility in the expansion of international production networks. *Strategic Management Journal, 33*(13), 1540–1556.

Fleming, P., & Spicer, A. (2002). Unravelling the paradox of covert resistance in organizations. In S. R. Clegg (Ed.), *Management and Organization Paradoxes*. Sydney: University of Technology.

Fresco, D. M., Williams, N. L., & Nugent, N. R. (2006). Flexibility and negative affect: examining the associations of explanatory flexibility and coping flexibility to each other and to depression and anxiety. *Cognitive Therapy and Research, 30,* 201–210.

Froot, K. A., & Stein, J. C. (1998). Risk management, capital budgeting, and capital structure policy for financial institutions: An integrated approach. *Journal of Financial Economics, 47*(1), 55–82.

Ganguly, A., Nilchiani, R., & Farr, J. V. (2009). Evaluating agility in corporate enterprises. *International Journal of Production Economics, 118*(2), 410–423.

Ganuthula, V. R. R., & Sinha, S. (2019). Flexibility in cognitive functioning: The play of paradox. *Global Journal of Flexible Systems Management*. https://doi.org/10.1007/s40171-019-00216-7.

Garavelli, A. C. (2003). Flexibility configurations for the supply chain management. *International Journal of Production Economics, 85*(2), 141–153.

García-Feijóo, L., & Howe, J. S. (2010). Financial contracting and managerial flexibility. *Quarterly Journal of Finance and Accounting, 49*(2), 5–26.

Germain, E., & Robillard, P. N. (2005). Engineering-based processes and agile methodologies for software development: A comparative case study. *The Journal of Systems and Software, 75*(1–2), 17–27.

Gibbs, G. R. (2007). *Thematic coding and categorizing. Analyzing Qualitative Data*. London: SAGE Publications Ltd.

Goldstein, Z., Petrie, D., & Sherif, Y. (2010). Finding middle ground management strategy for software development. *Information Management & Computer Security, 18*(3), 185–197.

Golicic, S. L., Flint, D. J., & Signori, P. (2017). Building business sustainability through resilience in the wine industry. *International Journal of Wine Business Research, 29*(1), 74–97.

Gong, Y., & Janssen, M. (2012). From policy implementation to business process management: Principles for creating flexibility and agility. *Government Information Quarterly, 29*(Supplement 1), S61–S71.

Gönsch, J., Koch, S., & Steinhardt, C. (2014). Revenue management with flexible products: The value of flexibility and its incorporation into DLP-based approaches. *International Journal of Production Economics, 153,* 280–294.

Gräßler, I., Pöhler, A., & Hentze, J. (2017). Decoupling of product and production development in flexible production environments. *Procedia CIRP, 60,* 548–553.

Gulledge, T., & Deller, G. (2009). Service-oriented concepts: Bridging between managers and technologists. *Industrial Management & Data Systems, 109*(1), 5–15.

Gupta, Y. P., & Somers, T. M. (1996). Business strategy, manufacturing flexibility, and organizational performance relationships: A path analysis approach. *Production and Operations Management, 5*(3), 204–233.

Haley, M. R., & Miller, L. A. (2014). Correlates of flexible working arrangements, stress, and sleep difficulties in the US workforce: Does the flexibility of the flexibility matter? *Empirical Economics, 48*(4), 1395–1418.

Han, J. H., Wang, Y., & Naim, M. (2017). Reconceptualization of information technology flexibility for supply chain management: An empirical study. *International Journal of Production Economics, 187,* 196–215.

Helm, R., & Gritsch, S. (2014). Examining the influence of uncertainty on marketing mix strategy elements in emerging business to business export-markets. *International Business Review, 23*(2), 418–428.

Heywood, J. S., & Miller, L. A. (2015). Schedule flexibility, family friendly policies and absence. *The Manchester School, 83*(6), 652–675.

Hopkin, P. (2017). *Fundamentals of risk management: understanding, evaluating and implementing effective risk management*. London: Kogan Page Publishers.

Iravani, S. M., Van Oyen, M. P., & Sims, K. T. (2005). Structural flexibility: A new perspective on the design of manufacturing and service operations. *Management Science*, *51*(2), 151–166.

Jain, S., Kashiramka, S., & Jain, P. K. (2018). Impact of organizational learning and absorptive capacity on the abnormal returns of acquirers: Evidence from cross-border acquisitions by indian companies. *Global Journal of Flexible Systems Management, 19*(4), 289–303.

Jain, V., & Raj, T. (2013). Ranking of flexibility in flexible manufacturing system by using a combined multiple attribute decision making method. *Global Journal of Flexible Systems Management, 14*(3), 125–141.

Jena, L. K., & Memon, N. Z. (2018). Does workplace flexibility usher innovation? A moderated mediation model on the enablers of innovative workplace behavior. *Global Journal of Flexible Systems Management, 19*(1), 5–17.

Joseph, O. A., & Sridharan, R. (2011). Effects of routing flexibility, sequencing flexibility and scheduling decision rules on the performance of a flexible manufacturing system. *The*

International Journal of Advanced Manufacturing Technology, 56(1–4), 291–306.

Kamasak, R., Yavuz, M., Karagulle, A. O., & Agca, T. (2016). Importance of strategic flexibility on the knowledge and innovation relationship: An emerging market study. *Procedia: Social and Behavioral Sciences, 229,* 126–132.

Kara, S., Kayis, B., & O'Kane, S. (2002). The role of human factors in flexibility management: A survey. *Human Factors and Ergonomics in Manufacturing & Service Industries, 12*(1), 75–119.

Karakaya, S., & Bakal, İ. S. (2013). Joint quantity flexibility for multiple products in a decentralized supply chain. *Computers & Industrial Engineering, 64*(2), 696–707.

Katic, M., & Agarwal, R. (2018). The flexibility paradox: Achieving ambidexterity in high-variety, low-volume manufacturing. *Global Journal of Flexible Systems Management, 19*(Suppl 1), 69–86.

Khorasani, S. T. (2018). A robust optimization model for supply chain in agile and flexible mode based on variables of uncertainty. *Global Journal of Flexible Systems Management, 19*(3), 239–253.

Kristianto, Y., Ajmal, M. M., & Helo, (2011). Advanced planning and scheduling with collaboration processes in agile supply and demand networks. *Business Process Management Journal, 17*(1), 107–126.

Kurniawan, R., Zailani, S. H., Iranmanesh, M., & Rajagopal, P. (2017). The effects of vulnerability mitigation strategies on supply chain effectiveness: Risk culture as moderator. *Supply Chain Management: An International Journal., 22*(1), 1–15.

Lafou, M., Mathieu, L., Pois, S., & Alochet, M. (2016). Manufacturing system flexibility: Product flexibility assessment. *Procedia CIRP, 41,* 99–104.

Lampón, J. F., Cabanelas, P., & Carballo-Cruz, F. (2017). A model for international production relocation: Multinationals' operational flexibility and requirements at production plant level. *Journal of Business Research, 77*(C), 95–101.

Lee, S. H., & Makhija, M. (2009). Flexibility in internationalization: Is it valuable during an economic crisis? *Strategic Management Journal, 30,* 537–555.

Leonard-Barton, D. (1992). Core capabilities and core rigidities: A paradox in managing new product development. *Strategic Management Journal, 13*(S1), 111–125.

Lévesque, C. (2008). The quest for flexibility in the mexican auto parts industry: Three tales from a multinational company. In *Flexibility at work*, Palgrave Macmillan, UK, pp. 173–193.

Li, C.-L., & Kouvelis, P. (1999). Flexible and risk-sharing supply contracts under price uncertainty. *Management Science, 45*(10), 1289–1462.

Li, Y., Li, P. P., Wang, H., & Ma, Y. (2017). How do resource structuring and strategic flexibility interact to shape radical innovation*?*. *Journal of Product Innovation Management, 34*(4), 471–491.

Lin, C.-T., Chiu, H., & Tseng, Y.-H. (2006). Agility evaluation using fuzzy logic. *International Journal of Production Economics, 101*(2), 353–368.

Livermore, J. A. (2008). Factors that significantly impact the implementation of an agile software development methodology. *Journal of Software, 3*(4), 31–36.

Lu, D., Ding, Y., Asian, S., & Paul, S. K. (2018). From supply chain integration to operational performance: The moderating effect of market uncertainty. *Global Journal of Flexible Systems Management, 19*(Suppl 1), 3–20.

Lynch, J., Mason, R. J., Beresford, A. K. C., & Found, P. A. (2012). An examination of the role for Business Orientation in an uncertain business environment. *International Journal of Production Economics, 137*(1), 145–156.

Managerial Flexibility Act. (2001). S.1612. 107th congress (2001–2002). Retrieved from https://www.congress.gov/bill/107th-congress/senate-bill/1612?overview=closed.

Manders, J. H. M., Caniëls, M. C. J., & Ghijsen, P. (2016). Exploring supply chain flexibility in a FMCG food supply chain. *Journal of Purchasing and Supply Management, 22,* 181–195.

Mangla, S. K., Kumar, P., & Barua, M. K. (2014). A flexible decision framework for building risk mitigation strategies in green supply chain using SAP–LAP and IRP approaches. *Global Journal of Flexible Systems Management, 15*(3), 203–218.

Mann, L., & Marshall, R. J. (2007). Teams in CSIRO: Reorganising for national research imperatives, Innovation: Management. *Policy & Practice, 9,* 136–145.

Marschak, T., & Nelson, R. (1962). Flexibility uncertainty and economic theory. *Metroeconomica, 14*(1-2-3), 42–58.

Martínez-Sánchez, A., Vela-Jiménez, M.-J., Pérez-Pérez, M., & de-Luis-Carnicer, P. (2011). The dynamics of labour flexibility: Relationships between employment type and innovativeness. *Journal of Management Studies, 48*(4), 715–736.

Maruping, L. M., Venkatesh, V., & Agarwal, R. (2009). An control theory perspective on agile methodology use and changing user requirements. *Information Systems Research, 20*(3), 377–399.

Mason, K., & Mouzas, S. (2012). Flexible business models. *European Journal of Marketing, 46*(10), 1340–1367.

McKeown, M. (2012). *Adaptability: The art of winning in an age of uncertainty*. Philadelphia, PA: Kogan Page Publishers. ISBN: 978 0 7494 6460 8

Melnik, A. N., Ermolaev, K. A., & Kuzmin, M. S. (2019). Mechanism for adjustment of the companies innovative activity control indicators to their strategic development goals. *Global Journal of Flexible Systems Management, 20*(3), 189–218.

Mieghem, V. (1998). Investment strategies for flexible resources. *Management Science, 44*(8), 1071–1078.

Mintzberg, H. (1980). Structure in 5's: A synthesis of the research on organization design. *Management Science, 26*(3), 322–341.

Mintzberg, H. (1989). The structuring of organizations. In D. Asch & C. Bowman (Eds.), *Readings in strategic management* (pp. 322–352). London: Palgrave.

Mintzberg, H., Quinn, J. B., & Ghoshal, S. (1995). *The strategy process: Concepts contexts and cases*. Prentice Hall, NJ: Englewood Cliffs.

Moore, S. A. (2001). A foundation for flexible automated electronic communication. *Information Systems Research, 12*(1), 34–62.

Naab, M., & Stammel, J. (2012). Architectural flexibility in a software-system's life-cycle: Systematic construction and exploitation of flexibility. In *Proceedings of the 8th international ACM SIGSOFT conference on Quality of Software Architectures*. https://doi.org/10.1145/2304696.2304701.

Nadkarni, S., & Narayanan, V. K. (2007). Strategic schemas, strategic flexibility, and firm performance: The moderating role of industry clock-speed. *Strategic Management Journal, 28*(3), 243–270.

Narasimhan, R., Talluri, S., & Das, A. (2004). Exploring flexibility and execution competencies of manufacturing firms. *Journal of Operations Management, 22*(1), 91–106.

Nerur, S., Mahapatra, R., & Mangalaraj, G. (2005). Challenges of migrating to agile methodologies. *Communications of the ACM, 48*(5), 73–78.

Niu, j., Shen, L., Peng, S., & Li, F. (2009). A measurement method of software flexibility based on bp network. In *International Workshop on Intelligent Systems and Applications*. https://doi.org/10.1109/IWISA.2009.5072635.

Pailles, N., Yannou, B., & Bocquet, J. C. (1997). Manufacturing flexibility: A new evaluation. In P. Chedmail, J. C. Bocquet, & D. Dornfeld (Eds.), *Integrated design and manufacturing in mechanical engineering*. Dordrecht: Springer.

Palanisamy, R., & Foshay, N. (2013). Impact of user's internal flexibility and participation on usage and information systems flexibility. *Global Journal of Flexible Systems Management, 14*(4), 195–209.

Patel, P. C., Kohtamäki, M., Parida, V., & Wincent, J. (2015). Entrepreneurial orientation-as-experimentation and firm performance: The enabling role of absorptive capacity. *Strategic Management Journal*, *36*(11), 1739–1749.

Patil, M., & Suresh, M. (2019). Modelling the enablers of workforce agility in IoT projects: A TISM approach. *Global Journal of Flexible Systems Management, 20*(2), 157–175.

Paulk, M. C. (2002). *Agile methodologies and process discipline*. Carnegie Mellon University. Retrieved from: http://repository.cmu.edu/isr.

PMI. (2009). *Practice standard for project risk management*. Project Management Institute Retrieved from: https://www.pmi.org/pmbok-guide-standards/framework/practice-standard-project-risk-management.

Podgornaya, A. I., Grudina, S. I., & Avdonina, S. G. (2015). An enterprise flexible development model. *Procedia Economics and Finance, 24,* 519–522.

Putnam, L. L., Myers, K. K., & Gailliard, B. M. (2014). Examining the tensions in workplace flexibility and exploring options for new directions. *Human Relations, 67*(4), 413–440.

Qiu, W., & Li, D. (2009). A study for end users' perceptions of business strategic factors among different IS/IT contexts. *The DATA BASE for Advances in Information Systems, 40*(1), 52–61.

Qumer, A., & Henderson-Sellers, B. (2008). An evaluation of the degree of agility in six agile methods and its applicability for method engineering. *Information and Software Technology, 50*(4), 280–295.

Raghuveer, K. S., Kuppili, R., & Nikhil, P. (2014). Improving competitiveness of IT companies by leveraging flexibility. In *Organisational flexibility and competitiveness*, Springer, India, pp. 213–222.

Rajesh, R. (2017). Technological capabilities and supply chain resilience of firms: A relational analysis using total interpretive structural modeling (TISM). *Technological Forecasting and Social Change, 118*(C), 161–169.

Raz, T., & Michael, E. (2001). Use and benefits of tools for project risk management. *International Journal of Project Management, 19*(1), 9–17.

Reichert, M., & Weber, B. (2012). *Enabling flexibility in process-aware information systems challenges, methods, technologies,* Springer. ISBN: 978-3-642-30409-5 (eBook).

Riley, M., & Lockwood, A. (1997). Strategies and measurement for workforce flexibility: An application of functional flexibility in a service setting. *International Journal of Operations & Production Management, 17*(4), 413–419.

Rogalski, S. (2011). *Flexibility measurement in production systems: Handling uncertainties in industrial production*. Berlin: Springer. ISBN 978-3-642-18117-7.

Rust, R. T., & Kannan, P. K. (2003). E-service: A new paradigm for business in the electronic environment. *Communications of the ACM, 46*(6), 36–42.

Sáenz, M. J., Knoppen, D., & Tachizawa, E. M. (2017). Building manufacturing flexibility with strategic suppliers and contingent effect of product dynamism on customer satisfaction. *Journal of Purchasing and Supply Management., 23*(2), 238–246.

Saldana, J. (2009). *The coding manual for qualitative researchers*. Thousand Oaks, CA: Sage Publications.

Sánchez-Sellero, P., Rosell-Martínez, J. M., & García-Vázquez, J. (2014). Absorptive capacity from foreign direct investment in Spanish manufacturing firms. *International Business Review, 23*(2), 429–439.

Santiago, L. P., & Vakili, P. (2005). On the value of flexibility in R&D projects. *Management Science, 51*(8), 1206–1218.

Schmitt, A., Raisch, S., & Volberda, H. W. (2016). Strategic renewal: Past research, theoretical tensions and future challenges. *International Journal of Management Reviews, 20*(1), 81–98.

Seebacher, G., & Winkler, H. (2014). Evaluating flexibility in discrete manufacturing based on performance and efficiency. *International Journal of Production Economics, 153,* 340–351.

Seebacher, G., & Winkler, H. (2015). A capability approach to evaluate supply chain flexibility. *International Journal of Production Economics, 167,* 177–186.

Shalender, K., & Yadav, R. K. (2019). Strategic flexibility, manager personality, and firm performance: The case of Indian Automobile Industry. *Global Journal of Flexible Systems Management, 2*(1), 77–90.

Sharifi, H., & Zhang, Z. (1999). A methodology for achieving agility in manufacturing organisations: An introduction. *International Journal of Production Economics, 62*(1–2), 7–22.

Sharma, M. K., Sushil, & Jain, P. K. (2010). Revisiting flexibility in organizations: Exploring its impact on performance. *Global Journal of Flexible Systems Management, 11*(3), 51–68.

Sherehiy, B., Karwowski, W., & Layer, J. K. (2007). In A review of enterprise agility: Concepts, frameworks, and attributes. *International Journal of Industrial Ergonomics, 37*(5), 445–460.

Shukla, S. K. (2018). The flexibility in product family engineering process. In Sushil, Singh, T. P. & Kulkarni, A. J. (Eds.), *Flexibility in resource management. Flexible systems management* (pp. 17–28). Singapore: Springer, .

Shukla, S. K., & Sushil. (2016a). The dominant strategy to achieve Indian defence manufacturing objectives. In *Conference proceedings of 16th Global Conference on Flexible Systems Management,* December 4–6, 2016, UTS Sydney, Australia.

Shukla, S. K., & Sushil. (2016b). Effective sourcing strategy for Indian defence needs. *Global Journal of Business Excellence, 3–9*(1), 1–14.

Shukla, S. K., & Sushil. (2018). Examining flexibility in capability maturity model integration framework using comparative study with flexibility maturity model. In *Conference proceedings of 18th global conference on flexible systems management,* IIM Lucknow.

Shukla, S. K., & Sushil. (2020a). Generic flexibility evaluation model. In P. K. Suri & R. Yadav (Eds.), *Transforming organizations through flexible systems management. Flexible Systems Management*. Singapore: Springer.

Shukla, S. K., & Sushil. (2020b). Evaluating the practices of flexibility maturity for software product and service organizations. *International Journal of Information Management, 50,* 71–89.

Singh, R. K., & Sharma, P. B. (2014). Development of framework for analyzing flexibility in supply chain. In M. K. Nandakumar, S. Jharkharia, & A. S. Nair (Eds.), *Organisational flexibility and competitiveness* (pp. 273–283). Delhi: Springer.

Slack, N., & Correa, H. (1992). The flexibilities of push and pull. *International Journal of Operations & Production Management, 12*(4), 82–92.

Smith, C., & Hutchinson, J. (1994). Addressing gender issues in management education: An Australian initiative. *Women in Management Review, 9*(7), 29–33.

Song, H., & Yu, K. (2009). How do interpartner legitimacy and trust affect distribution flexibilities and distribution competitiveness. In *IEEE international conference on management and service science*, ISBN 978-1-4244-4639-1, pp. 1–7.

Sopelana, A., Kunc, M., & Hernaez, O. R. (2014). Towards a dynamic model of organisational flexibility. *Systemic Practice and Action Research, 27*(2), 165–183.

Steinbring, M., Motschnig, R., & Pitner, T. (2013). A multidimensional model of web-induced flexibility in organisations. *Global Journal of Flexible Systems Management, 14*(1), 57–65.

Stephenson, Z., & McDermid, J. (2005). Deriving architectural flexibility requirements in safety-critical systems. *Software IEE Proceedings, 152*(4), 143–152.

Stone, A. (2001). An ontological approach to quantifying the functional flexibility of embedded systems. *IEEE Systems Journal, 5*(1), 111–120.

Struweg, N. (2014). An integrated communication implementation model for the post-2000 business environment. *Public Relations Review, 40*(2), 184–192.

Subramanyam, R., Ramasubbu, N., & Krishnan, M. S. (2012). In search of efficient flexibility: Effects of software component granularity on development effort, defects, and customization effort. *Information Systems Research, 23*((3 PART 1)), 787–803.

Sullivan, A. S. (1994). Flexibility as a management tool. *Employment Relations, 21*(4), 393–405.

Sushil. (2011). Flexibility, vitality and sustainability. *Global Journal of Flexible Systems Management*, *12*(1–2), iii.

Sushil. (2012a). Making flowing stream strategy work. *Global Journal of Flexible Systems Management, 13*(1), 25–40.

Sushil. (2012b). Flexibility maturity model: Possibilities and direction. *Global Journal of Flexible systems Management, 13*(2), 75–76.

Sushil. (2013). *Flowing Stream Strategy: Leveraging Strategic Change with Continuity*. Springer: New Delhi.

Sushil. (2015a). Managing flexibility: Developing a framework of flexibility maturity model. In K. T. Sushil Bhal, & S. P. Singh (Eds.), *Managing flexibility people, process, technology and business*, pp. 3–19.

Sushil. (2015b). Valuation of flexibility. *Global Journal of Flexible Systems Management, 16,* 219–220.

Sushil. (2016). Strategic flexibility in ecosystem. *Global Journal of Flexible Systems Management., 17,* 247–248.

Sushil. (2017). Does flexibility mitigate or enhance risk? *Global Journal of Flexible Systems Management, 18*(3), 169–171.

Svensson, G., & Høgevold, N. M. (2017). Framing business sustainability efforts through time. In M. Stieler (Ed.), *Creating marketing magic and innovative future marketing trends Developments in marketing science: Proceedings of the academy of marketing science*. Cham: Springer.

Swafford, P. M., Ghosh, S., & Murthy, N. (2008). Supply chain agility through IT integration and flexibility. *International Journal of Production Economics, 116*(2), 288–297.

Sweet, S., & James, J. B. (2013). *What do managers really think about flexibility?* Retrieved from https://www.bc.edu/content/dam/files/research_sites/agingandwork/pdf/publications/IB26_TPM.pdf.

Tamayo-Torres, J., Gutierrez-Gutierrez, L., & Ruiz-Moreno, A. (2014). The relationship between exploration and exploitation strategies, manufacturing flexibility and organizational learning: An empirical comparison between Non-ISO and ISO certified firms. *European Journal of Operational Research, 232*(1), 72–86.

Tang, C. S. (1991). Capacity and flexibility planning of an economical manufacturing system. *Journal of Intelligent Manufacturing, 2*(2), 75–82.

Tarafdar, M., & Qrunfleh, S. (2017). Agile supply chain strategy and supply chain performance: Complementary roles of supply chain practices and information systems capability for agility. *International Journal of Production Research, 55*(4), 925–938.

TenDam, H. (1987). Managerial flexibility: A strategic asset. *Leadership & Organization Development Journal, 8*(2), 11–16.

Treville, S. D., Bendahan, S., & Vanderhaeghe, A. (2007). Manufacturing flexibility and performance: bridging the gap between theory and practice. *Flexible Services and Manufacturing Journal, 19,* 334–357.

Tseng, Y. H., & Lin, C. T. (2011). Enhancing enterprise agility by deploying agile drivers, capabilities and providers. *Information Sciences, 181*(17), 3693–3708.

Ullah, A., & Lai, R. (2013). A systematic review of business and information technology alignment. *ACM Transactions on Management Information Systems, 4*(1), 1–30.

Usui, T., Kotabe, M., & Murray, J. Y. (2017). A dynamic process of building global supply chain competence by new ventures: The case of uniqlo. *Journal of International Marketing, 25*(3), 1–20.

Van Biesebroeck, J. (2007). Robustness of productivity estimates. *The Journal of Industrial Economics, 55*(3), 529–569.

Verdu, A. J., Marıa, J., & Gras, G. (2009). Measuring the organizational responsiveness through managerial flexibility. *Journal of Organizational Change Management, 22*(6), 668–690.

Vokurka, R. J., & Leary Kelly, S. W. O. (2000). A review of empirical research on manufacturing flexibility. *Journal of Operations Management, 18,* 485–501.

Volberda, H. W. (1992). *Organizational flexibility: change and preservation: A flexibility audit and redesign method*. Wolters-Noordhoff. ISBN-13: 978-9001915407.

Volberda, H. W. (1996). Towards the flexible form: How to remain vital in hypercompetitive environment. *Organization Science, 7*(4), 359–374.

Volberda, H. W. (1997). Building flexible organizations for fast-moving markets. *Long Range Planning, 30*(2), 169–183.

Wadhwa, S., & Rao, K. S. (2003). Flexibility and agility for enterprise synchronozation: Knowledge and innovation management towards flexagility. *Studies in Informatics and Control, 12*(2), 111–127.

Wang, H., Mastragostino, R., & Swartz, C. L. E. (2016). Flexibility analysis of process supply chain networks. *Computers & Chemical Engineering, 84,* 409–421.

Wang, T., Libaers, D., & Jiao, H. (2015). Opening the black box of upper echelons in China: TMT attributes and strategic flexibility. *Journal of Product Innovation Management, 32*(5), 685–703.

Wang, Y., & Li-Hua, R. (2007). *Marketing competences and strategic flexibility in China*. ISBN: 978-0-230-28860-7 (E-Book). https://doi.org/10.1057/9780230288607.

Ward, S., & Chapman, C. (2003). Transforming project risk management into project uncertainty management. *International Journal of Project Management, 21*(2), 97–105.

White, A., Daniel, E. M., & Mohdzain, M. (2005). The role of emergent information technologies and systems in enabling supply chain agility. *International Journal of Information Management, 25*(5), 396–410.

Whyman, P. B., & Petrescu, A. I. (2015). Workplace flexibility practices in SMEs: Relationship with performance via redundancies, absenteeism, and financial turnover. *Journal of Small Business Management, 53*(4), 1097–1126.

Wu, K.-J., Tseng, M.-L., Chiu, A. S. F., & Lim, M. K. (2017). Achieving competitive advantage through supply chain agility under uncertainty: A novel multi-criteria decision-making structure. *International Journal of Production Economics, 190,* 96–107.

Xie, F. (2009). Managerial flexibility, uncertainty, and corporate investment: The real options effect. *International Review of Economics & Finance, 18*(4), 643–655.

Yazici, H. J. (2005). Influence of flexibilities on manufacturing cells for faster delivery using simulation. *Journal of Manufacturing Technology Management*, *16*(8), 825–841.

Yu, K. (2013). The effects of objective and perceived environmental uncertainty on supply chain flexibility. In *10th International Conference on Service Systems and Service Management*. https://doi.org/10.1109/ICSSSM.2013.6602647

Yu, K., Cadeaux, J., & Luo, B. N. (2015). Operational flexibility: Review and meta-analysis. *International Journal of Production Economics, 169,* 190–202.

Yusuf, Y., Sarhadi, M., & Gunasekaran, A. (1999). Agile manufacturing: The drivers, concepts and attributes. *International Journal of Production Economics, 62*(1–2), 33–43.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Key Questions

(1) What are the benefits of flexibility initiatives undertaken by management in the sphere of organizational control?
(2) What are concerns, hurdles, and mechanisms in implementing flexibility initiatives?

Sanjai Kumar Shukla is currently working with Bharat Electronics Ltd., a leading defense public sector in India. He had also worked in security product's design and development in HCL. He has rich experience in the design, development, and project management of command and control systems of Indian armed forces and network security products. His work area and research interest includes the flexible systems, agile management, command and control systems, defense projects execution and technology planning. He has received M.S. degree (Birla Institute of Technology and Science Pilani). He also holds MBA and Ph.D. degrees of Indian Institute of Technology Delhi.

Sushil is Abdulaziz Alsagar Chair Professor (Professor of Strategic, Flexible Systems, and Technology Management) and Chair, Strategic Management Group, Department of Management Studies, Indian Institute of Technology Delhi. He has served as Deputy Director (Operations) and Dean (Faculty) at IIT Delhi. He has served as Visiting Professor and delivered seminars in many leading universities. He is an active researcher and has supervised more than 60 doctoral dissertations. He has twenty books to his credit and over 300 papers in various refereed journals and conferences. He is the editor-in-chief of journal "Global Journal of Flexible Systems Management."

Manoj Kumar Sharma has 32 years of experience in industry, academics, and government. He has rich experience as faculty, administrator, manager, and policy analyst. Currently, he is the Chief Executive Officer (CEO), DTU-IIF and OSD, DTU East Campus, Delhi Technological University. He held senior positions in academics as Dean and executive council member of leading universities.

Modelling Of Viscoelastic Fluid In Finishing Processes

Krovvidi Srinivas, Q. Murtaza, A.K Agarwal

Abstract: The Traditional Finishing process mostly lapping, honing, buffing, superfinishing etc contains various disadvantages like labour intensive work, high Production cost and time. Moreover, the quality of surface finish achieved is not up to the mark but in present era there is requirement of better quality of products with desired surface finish, in order to overcome this finishing issue Abrasive Magnetic Finishing and Abrasive Flow Machining process comes into the consideration. Both the Process uses a Viscoelastic media which contains an adequate amount of Polymer, gel, Oil and abrasives. Magnetic Abrasive Finishing includes polymer media with carbonyl particle which align along the magnetic field and results in effective machining. This paper includes solid modelling of the process and simulation on Ansys Fluent module and analyse the Flow parameters of the viscoelastic polymer media and its effect on material removal.

Index Terms: viscoelastic Magnetic Abrasive Finishing (VMAF), Magnetic Abrasive Finishing (MAF), Material Removal (MR), Surface Finish, Ansys Maxwell, Polymer Media, Abrasive Flow Finishing (AFF).

1 INTRODUCTION

Magnetic abrasive finishing is an effective machining process in which magnetic field is used to machine the target surfaces of the workpiece by influencing the carbonyl particles and directing them towards the wall of the workpiece for machining. In conventional abrasive machining process both the workpiece and the target surface are polarised and due to the interplay of magnetic field, a Brush of abrasive is formed which results in the effective machining of the workpiece. This effective concept of machining got the eyeball of many scholar and initiated a large number of researches in this field like Kang and Yamaguchi [1] developed a multipole dip- system for the machining of internal surface of the capillary tube. The researchers highlighted the difficulties faced in machining of the capillary tube mainly in controlling the magnetics abrasive in MAF process, the researchers suggested that the multipole system with alternate magnetics and nonmagnetic region simultaneously applied on the workpiece which results in the effective machining of the long capillary tube. This process has allowed the application of MAF on long Capillary tube as well which was earlier confined to the short length tube only. The final conclusion of the researcher was that the double pole system finishes the workpiece in half a time as compare to the single pole system. V.K Jain [2] explained various non-traditional machining process which are used in modern times which involves the use of abrasive and the media to support it. The various process which he discussed in its paper was Abrasive Flow Finishing (AFF), Magnetic Abrasive Finishing (MAF), Magnetorheological Finishing, Magnetorheological Abrasive Flow Finishing, Elastic Emission Machining (EEM) and Magnetic Float Polishing. he laid attention on the magnetic abrasive finishing process and gave some important conclusion like involvement of slot in the magnet over non machining zone and the use of pulse Dc supply rather than smooth DC supply in the system. Wani et al. [4] gave stress on Magnetic Abrasive flow finishing which is a process made by combining the advantage of both abrasive flow machining and the magnetic abrasive finishing process, and used finite element method to find distribution of magnetic potential in the magnetic brush. Through this distribution the researcher evaluated machining pressure, surface finishing and material removal. After the successful evaluation of the results the researcher claims that these results are similar to those present in the literature. Kala et al. [4] developed a setup which aims to catch the force signature for double disc magnetic abrasive finishing process and the conventional abrasive magnetic abrasive finishing process and by these properties the researchers aims to understand the working of the flexible magnetic abrasive brush. Thus, to do this the researchers performed the magnetic abrasive finishing operation in single as well as double disc and finally understood the property of flexible magnetic abrasive brush behaviour from it. Wu et al [5] developed a model of magnetic abrasive finishing process and analyzed its properties based on the rotational speed, cutting fluid and the magnetic pole and current frequency and its effect on change in material removal and surface finish are invested. The researchers finally claimed that the neat cutting oil is required for processing which results in higher material removal and finer finish. They also concluded that finishing is better with the increasing in rotational speed and as the current frequency increases the angle variation of the particle increases and if frequency is low the surface requirement is up to few nanometers. Rathod et al. [6] correlated the difference shape factor with wide range of particle size. The researcher made use of slurry abrasion tester with silica sand slurry and commented on the effect of sliding distance slurry abrasion volume, where they argued that volume loss is proportional to the sliding distance. They also studied the effect of micro ploughing and cutting. Srivastav et al. [7] simulated a model of electrochemical magnetic abrasive finishing process in which the researchers used stainless steel workpiece whose radius was ten mm and thickness was 2.5 mm. the researchers used chromium oxide as the abrasive particle whose average diameter was 387 micrometers, the intensity of magnetic field was around 1 tesla and the rotational speed of tool was taken as 5835 rpm and its diameter was 6 mm. the electrolyte used for the experiment was brine solution with twenty percent concentration. The researchers study the thermal model in both steady and transient model. Srivastav and Amit [8] did the thermal analysis of work-brush interface and along the workpiece depth in magnetic abrasive finishing process with the help of ANSYs R15 software in both steady and thermal state. The value of magnetic field intensity given to 6 mm diameter tool was 1 tesla. The researcher used chromium oxide abrasives of about 387 micrometer Diameter. The workpiece has a radius of 10 mm and 2.5 mm thickness and the material of the workpiece was silicon nitride. Singh et al. [9] used Taguchi L orthogonal array to design an experiment which could measure surface temperature of the finishing surface of mild steel using MAF process. The researchers used current, working gap, Rotational speed and abrasive weight as

its important process parameters. After the Taguachi optimization the Fuzzy logic is used by researcher through which prediction of surface temperature is and also it is validated. Jaiswal et al. [10] made FEM model of MAF and investigated the distribution of magnetic force on the workpiece. The researcher also made a theoretical model for computing the material removal and surface roughness. For this the workpiece was considered of uniform surface without statically distribution. Numerical experiments were also computed which were validated later on through literature. Yan et all [11] performed the experiment on MA genetic abrasive finishing based on a simulated model which is based on FEM model of electromagnetic inductor. He then compared the simulated model with the actual electromagnetic inductor and predicted a marginal error of only 7 percent. Li et all [12] developed a finishing setup having new media for finishing the rotatory surface ANSYS Maxwell was used to determine the optimum angle between the poles and the mathematical model was made to compute the MRR in which pressure and velocity was computed based on arced wear model. The MRR computation model was made which predicted MRR as a function of pole rotational speed, Magnetic flux density, cam rotational speed and the ferromagnetic and abrasive particles' diameter. Li et all [13] proposed a new viscoelastic magnetic Abrasive based on analysis of field characteristic and existing finishing process. They also discussed the motion locus of the abrasives. The present scenario, abrasive flowing finishing, grinding, abrasive brush deburring, magnetorheological finishing, manual deburring etc. are broad name in term of finishing but still they possess some limitations in finishing quality and efficiency. Shimamura et al. [14], [15] modelled and tested that magnetic abrasive particles are subjected to the pressure for machining and that pressure is a function of magnetic flux density, number of abrasive particles and the permeability of abrasive medium. Fox et al [16] observed that the magnetic force (function of the volume and magnetic susceptibility of the ferrous particles in the magnetic field, magnetic field intensity and the gradients at the finishing area). Srinivas et al. [17] studied the distribution of the magnetic field intensity along with the computation of torque and force and related flow parameters of Viscoelastic Magnetic Abrasive finishing process with magnetic field intensity. The Researchers used ANSYS Maxwell and ANSYS Fluent in order to find co relation between the flow parameters and magnetic field intensity and commented on the successful application of the proposed viscoelastic media. The researchers also plotted the graph between the magnetic field intensity and the current which is applied on the single pole electromagnet to produce the desired magnetic field for machining in the magnetic abrasive finishing process. Srinivas and Anant [18] studied the pressure variation in abrasive flow machining process with the help of a simulated model made on Ansys Fluent and commented on the effect of pressure variation in the machining operation. They used viscous media similar to that used in viscoelastic magnetic abrasive finishing with the difference that in later process there is an addition of carbonyl particles which gets activated by magnetic field. The researcher computed pressure difference between 40 MPa and 20 MPa and the value of pressure so obtained was validated with the help of literature review. Ali et al. [19] reviewed various researches on AFM which uses a viscoelastic media and semisolid media which can support various abrasive based finishing process including both magnetic and non-magnetic based finishing process. Bhardwaj et al [20] compiled the researches which contained the use of viscoelastic media in various hybrids of abrasive flow machining process, the researchers also commented on the tooling and fixturing in various hybrids of AFM and the compatibility of the viscoelastic media with the hybrid process. So it can be seen that Viscoelastic Magnetic Abrasive Finishing has a vital role in Finishing department and must be consider for effective machining operation.

2 SIMULATION AND MODELING

A simulation model was proposed to overcome the research gap in the process of magnetic abrasive finishing process. For this purpose, ANSYS Fluent was used so a basic geometry was made in CREO 3.0 software and was then transferred to ANSYS fluent as an IGES File. The basic model contains a brass workpiece with alumina abrasive mixed in the viscous fluid whose property are listed in the tables. When a geometry is port into the ANSYS fluent software, it undergoes 5 steps. The first step (Figure 1) is the geometry where we do name selection and perform the Boolean operations if required. So, we did name selection in which the selections were Inlet, Outlet, workpiece and fluid Domain, next was the Boolean operation where we created path for the fluid flow within the internal surface of the workpiece. Then the file was transferred into the meshing section in which the mesh was created which is depicted by figure 2 and 3. The total elements in mesh was about $3X\ 10^3$ and the shape of the element was majorly tetragonal. After successful meshing the next step was to load the fluent setup for which is shown in figure 4. This is the major step in which we select the type of simulation, energy equation, material assignment, boundary condition and select the number of iterations for the process of simulation. For the process the laminar simulation with the boundary condition on was chosen and the adequate material is assigned to the parts. The boundary condition used for the simulation was Pressure inlet of 40 MPa and the outlet at 20 MPa, the workpiece and walls were stationary. The detail of boundary condition is given in table 2. Initially the fluid is not given any velocity but its movement would be there through the pressure difference which would do machining. The effective machining may be done when we will be placed magnet inside the system which cannot be done on ANSYS fluent, although for that ANSYS Maxwell must be taken into consideration. After all this assignment the simulation is allowed to run for iteration and the plot of energy along with the X, Y and Z velocity was plotted against the number of iterations, which is shown in figure 6.

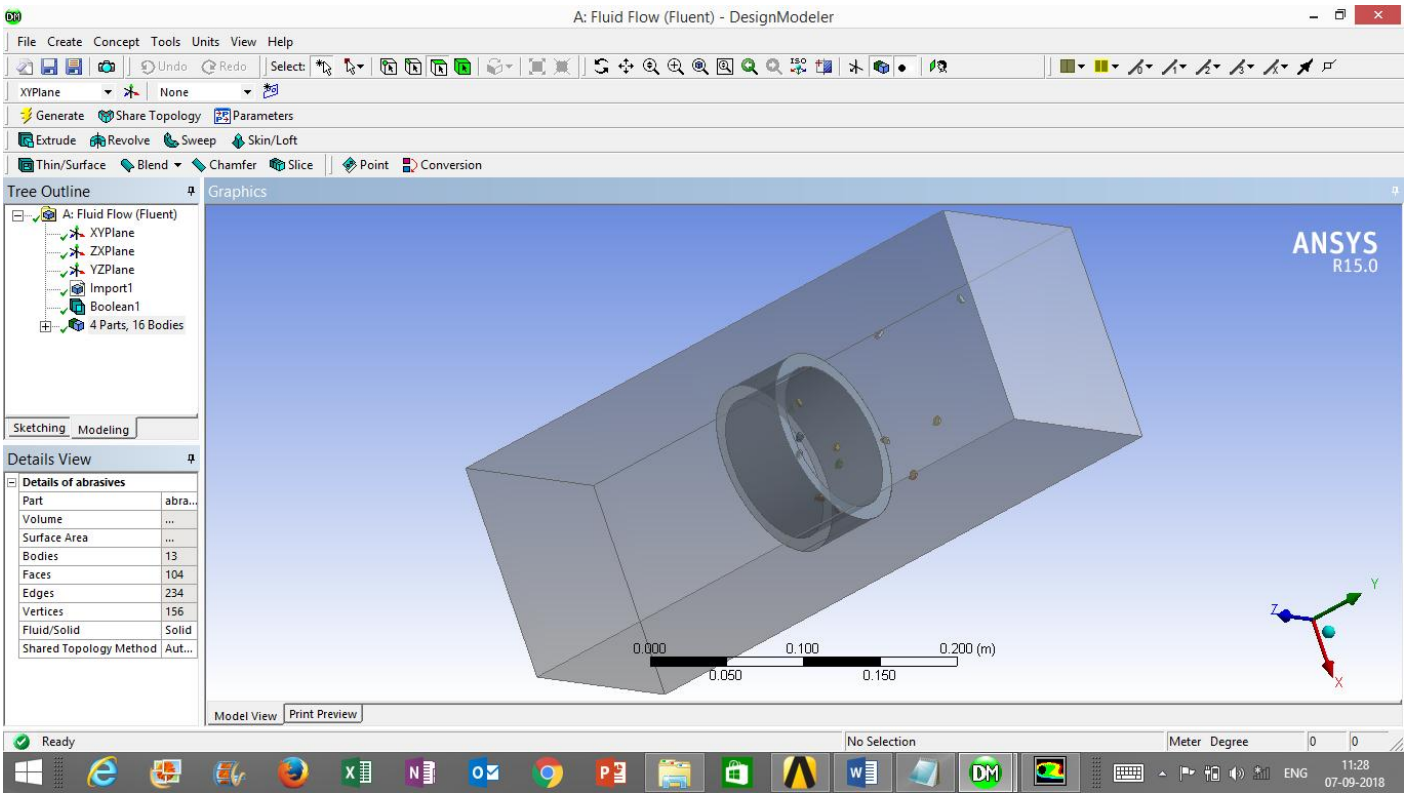


Fig.1. *Initial geometry in ANSYS fluent*

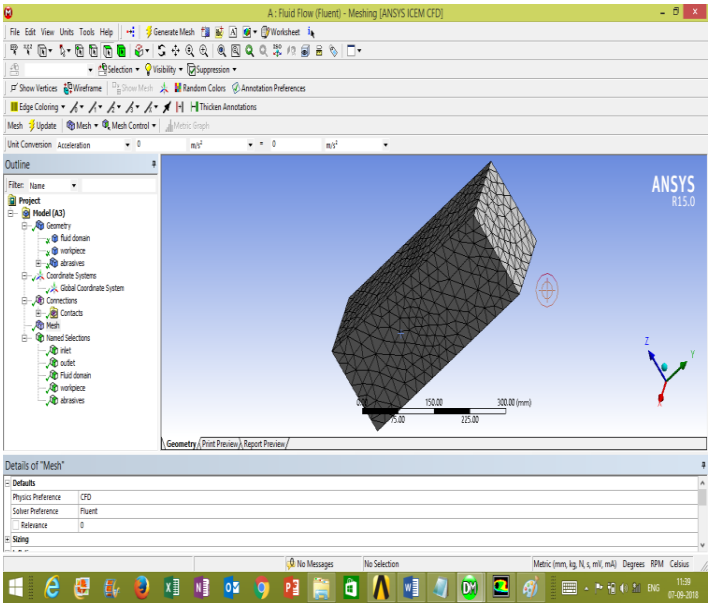


Fig. 2. *Meshing of Fluid Domain*

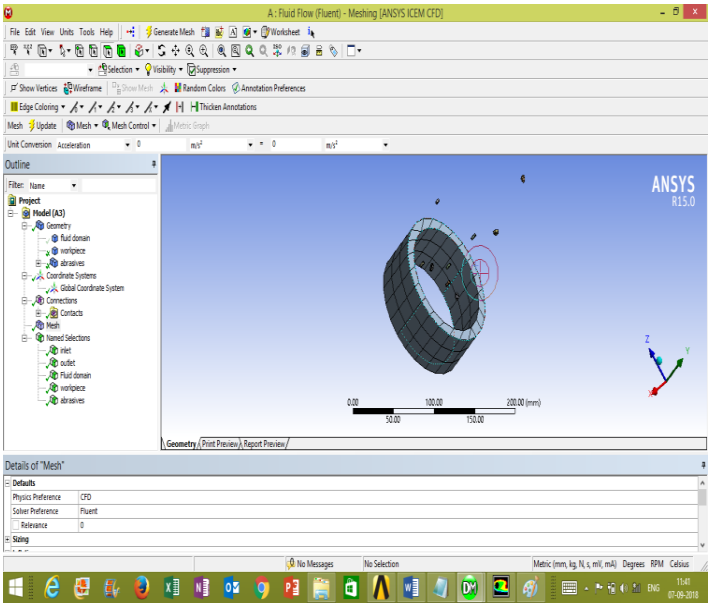


Fig. 3. *Meshing of work piece*

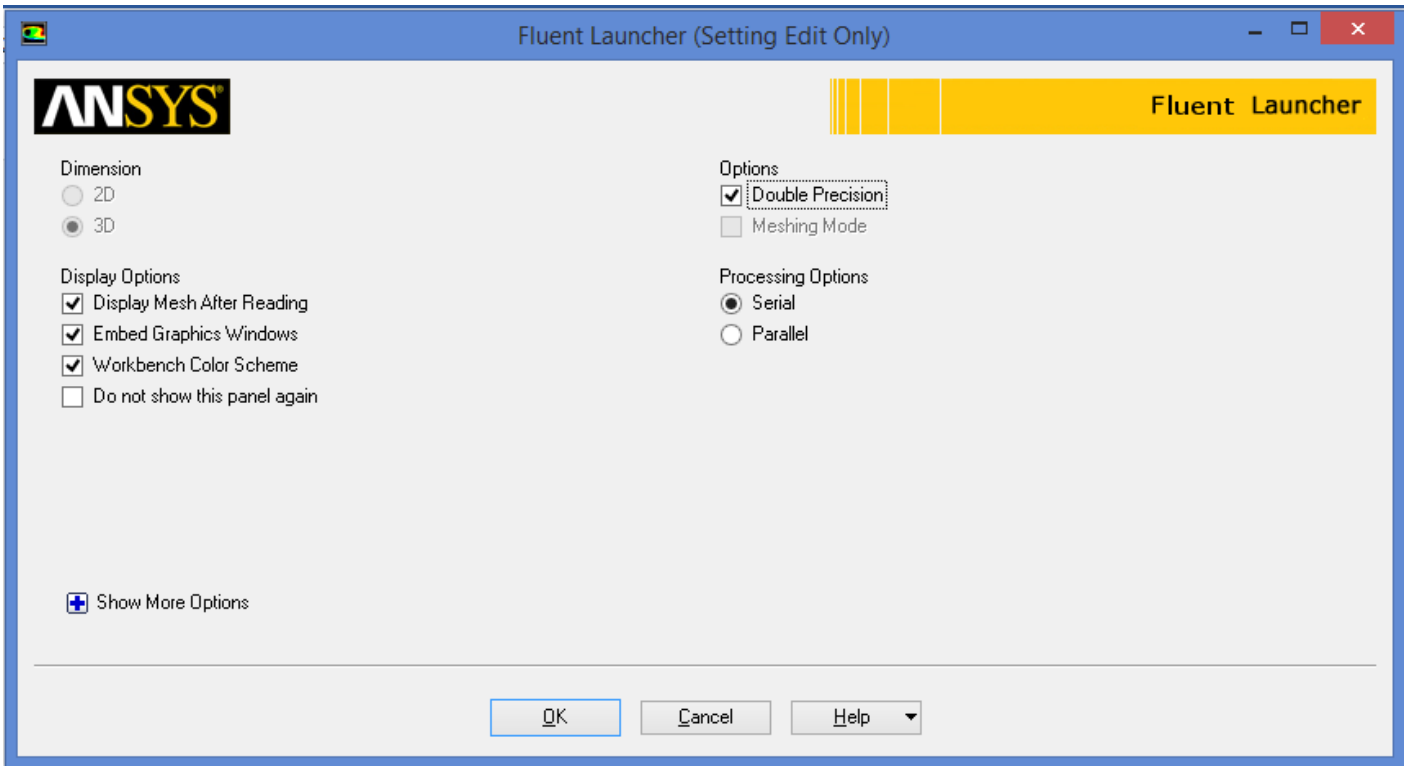


***Fig.4**. Initial setup*

TABLES

***Table 1** Showing the Domain Nodes and Elements of Simulations*

MATERIAL	TYPE	DENSITY (Kg/m³)	VISCOSITY (Kg/m-s)	THERMAL CONDUCTIVITY (W/ m.K)	SPECIFIC HEAT (J/Kg-K)
Viscous Fluid	Fluid	1212	0.782	0.22	2025
Brass	solid	8500		109	400
Aluminum abrasives	solid	2950		12	451

***Table 2** showing The Boundary condition of simulations*

Sr no	Name Selection	Boundary condition
1.	Inlet	Inlet pressure: 40 MPa
2.	Fluid	Stationary fluid at room temperature: 300 K
3.	Workpiece	Stationary workpiece at Room Temperature: 300 K
4.	Outlet	Outlet Pressure: 20 MPa

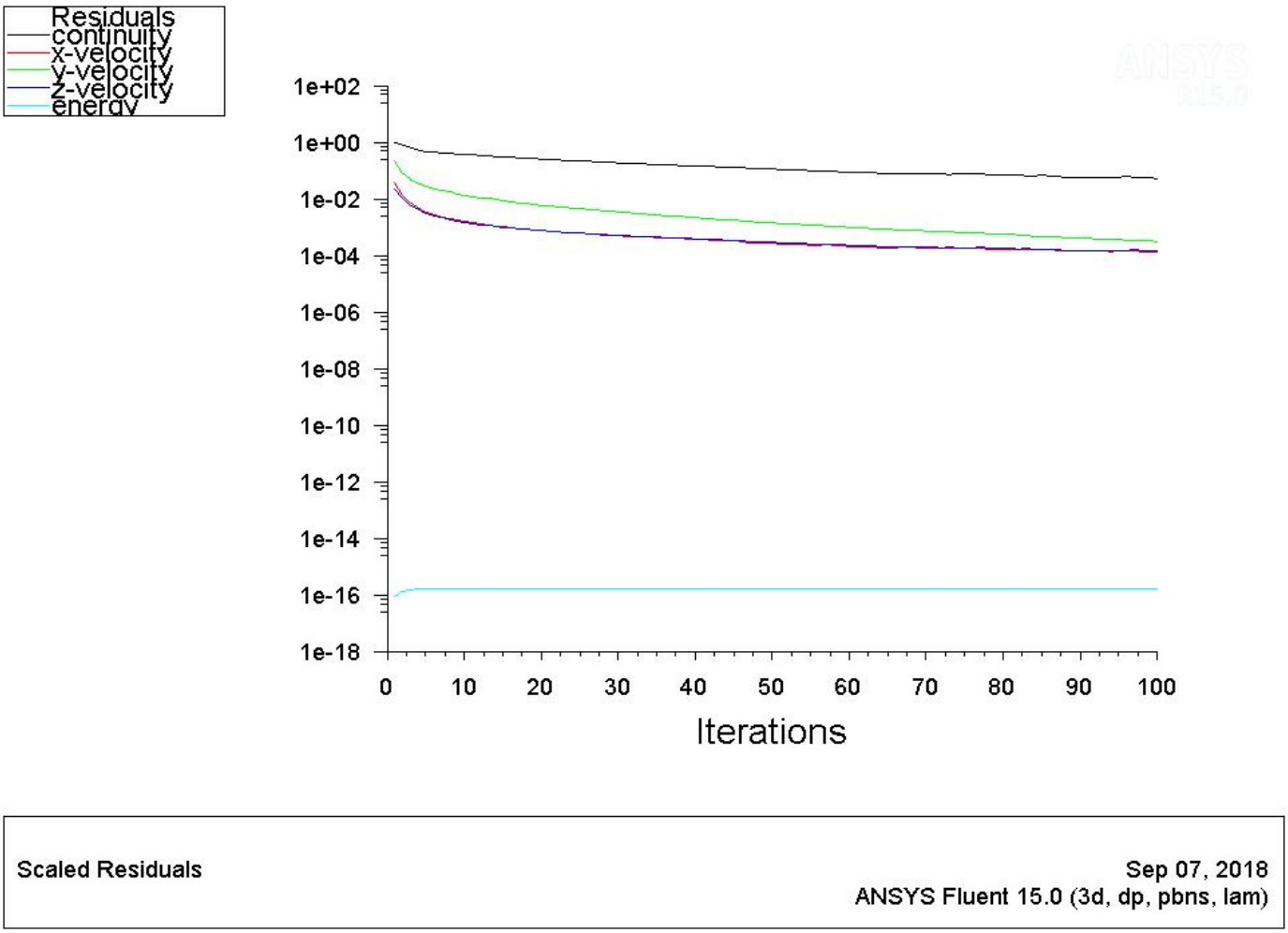


***Fig**. 5. Energy Vs Iteration curve*

3 Results and Discussions

As The results show the maximum pressure is 26.2455 MPa and the minimum Pressure obtained after the analysis is 10.4519 MPa which is an optimum value for machining the brass workpiece at the pressure range of 40 MPa to 20 MPa (Figure 7). The pressure drops in the reason of the workpiece due to the bell effect, When the surface area decreases at the neck region the velocity of the media decreases resulting in the increase in the pressure value at the neck region which is at the starting of workpiece in the direction of the flow, But the value of the pressure decreases throughout the flow of the polymer media which can be seen through the subsequent figure in which the pressure variation along the plane is given, in which the drop in the pressure is shown. Figure 8 shows the streamline variation of the Polymer media. As observed from the figure the stream lines are uniform and smooth and representing the flow due to the pressure difference only. The effective machining may also be achieved by putting a magnet around the workpiece which will attract the carbonyl particle present in the media which would further hit the magnetic particles. The figure 9 shows the force applied by the polymer media on the workpiece, though the force applied would be uniform as the abrasive would be uniformly mixed with the polymer media which results in the application of force on the workpiece. Though the maximum value of the force is 5440 N. The force would be maximum at the point where the abrasives strikes the surface of the workpiece. The figure 10 shows the distribution of velocity of the polymer media on the workpiece, the polymer media is specially made by combining various polymer, gel, oil and abrasives blended together to form a semisolid putty which is extruded up and down with the pressure difference to achieve the effective machining of the process. As Described by the process the maximum velocity is 183.2 m/s and the minimum velocity is 6.94 m/s, again this point must be taken care of that the figure is showing the velocity distribution at the region of the workpiece and not along the flow. The distribution of velocity along the flow would be different. The figure 11 shows the pressure variation along the entire length of the cylinder in the abrasive flow machine rather than the workpiece regions. It shows the pressure at the upper side is 29 MPa and at the lower side is 10.45 MPa the region for the pressure drop and is already discussed. Hence simulation was successful in showing machining characteristic of Viscoelastic media

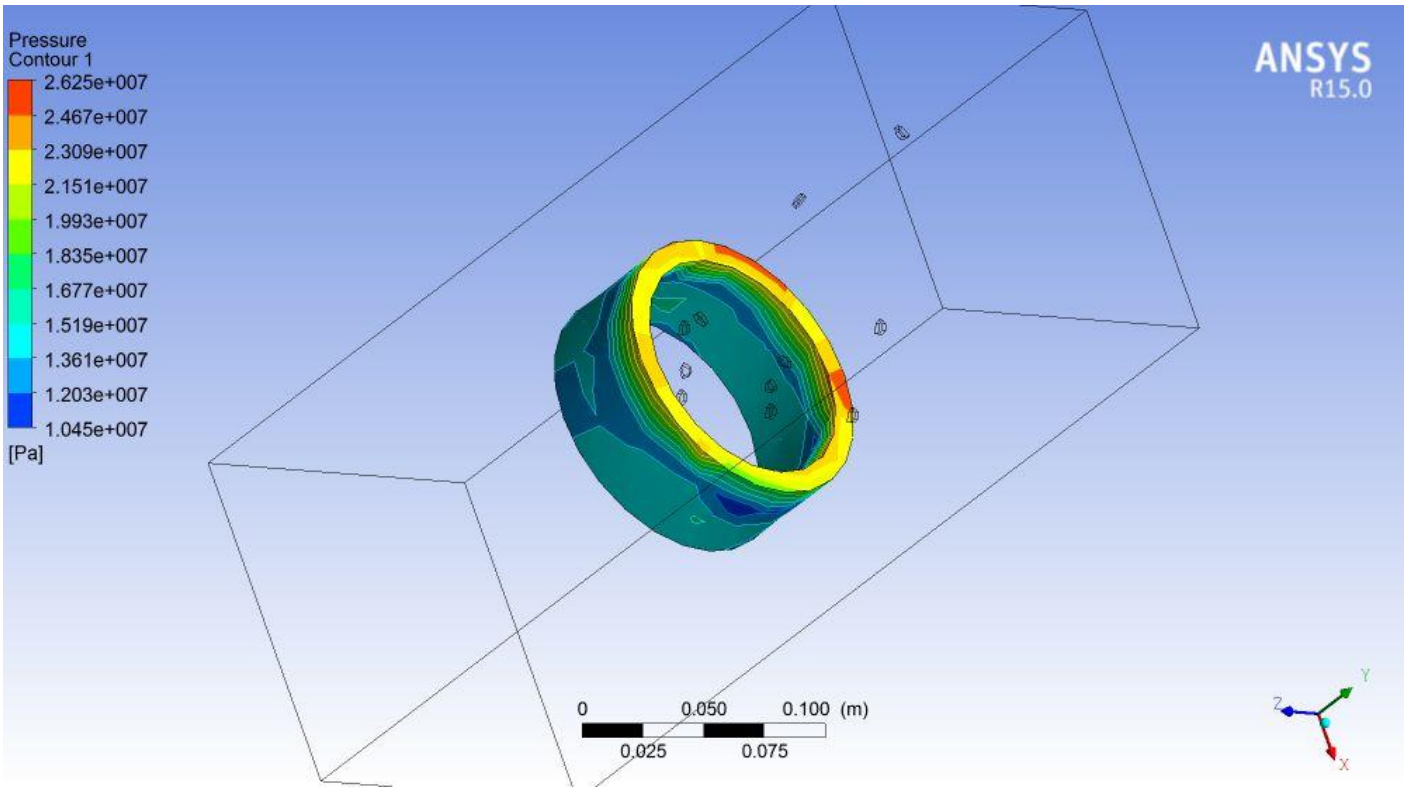


FIG.6 *EFFECT OF PRESSURE VARIATION ON WORKPIECE*

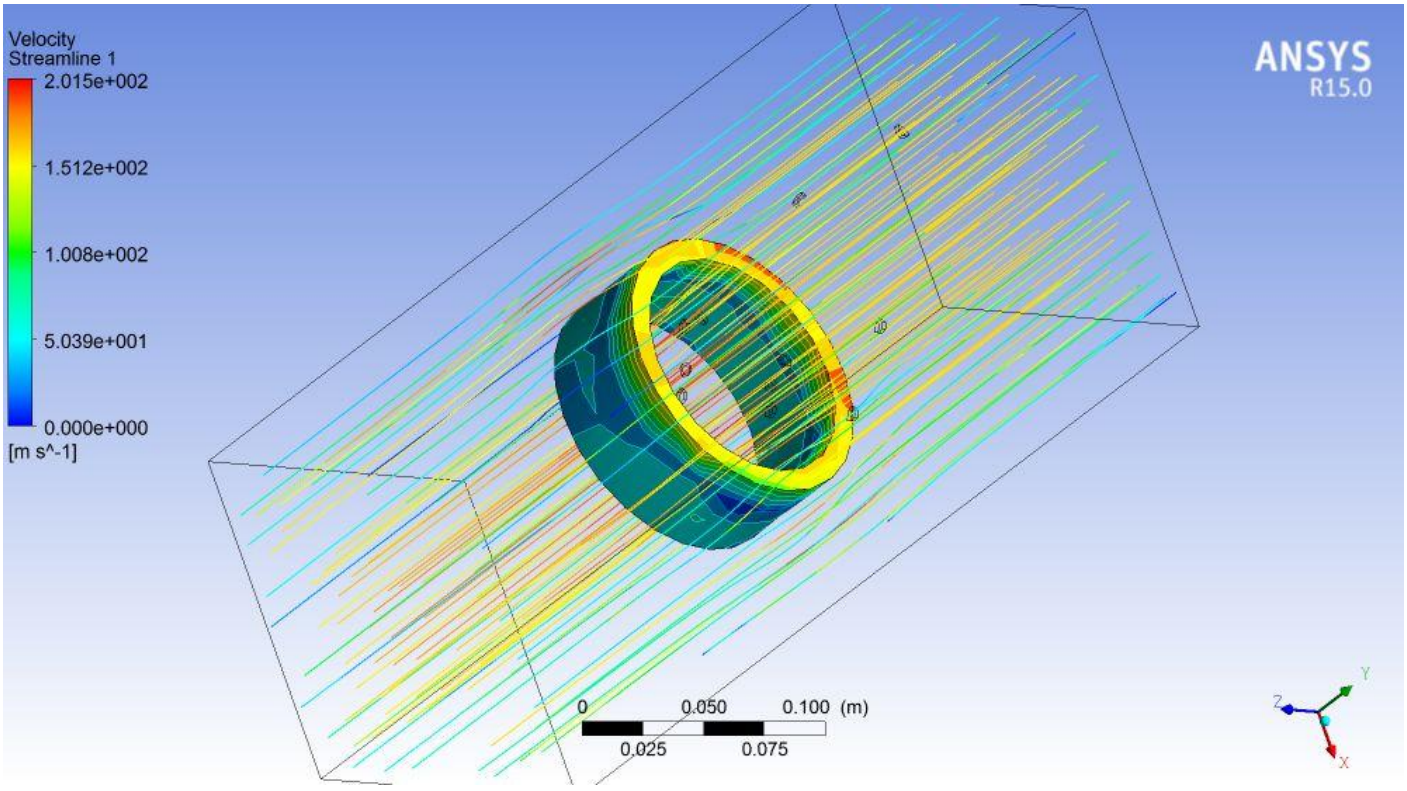


FIG.7 *STREAMLINE VARIATION OF POLYMER MEDIA*

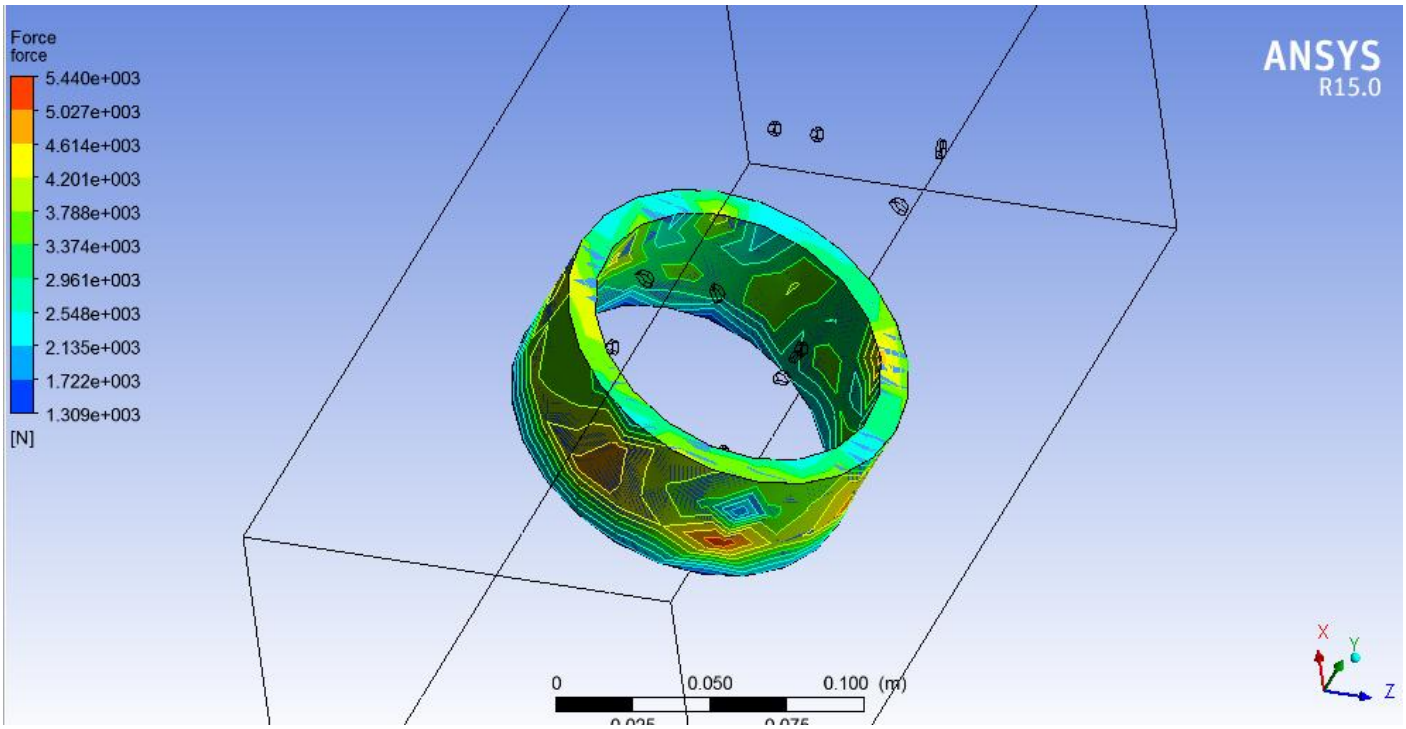


FIG.8 *FORCE APPLIED ON POLYMER MEDIA ON WORKPIECE*

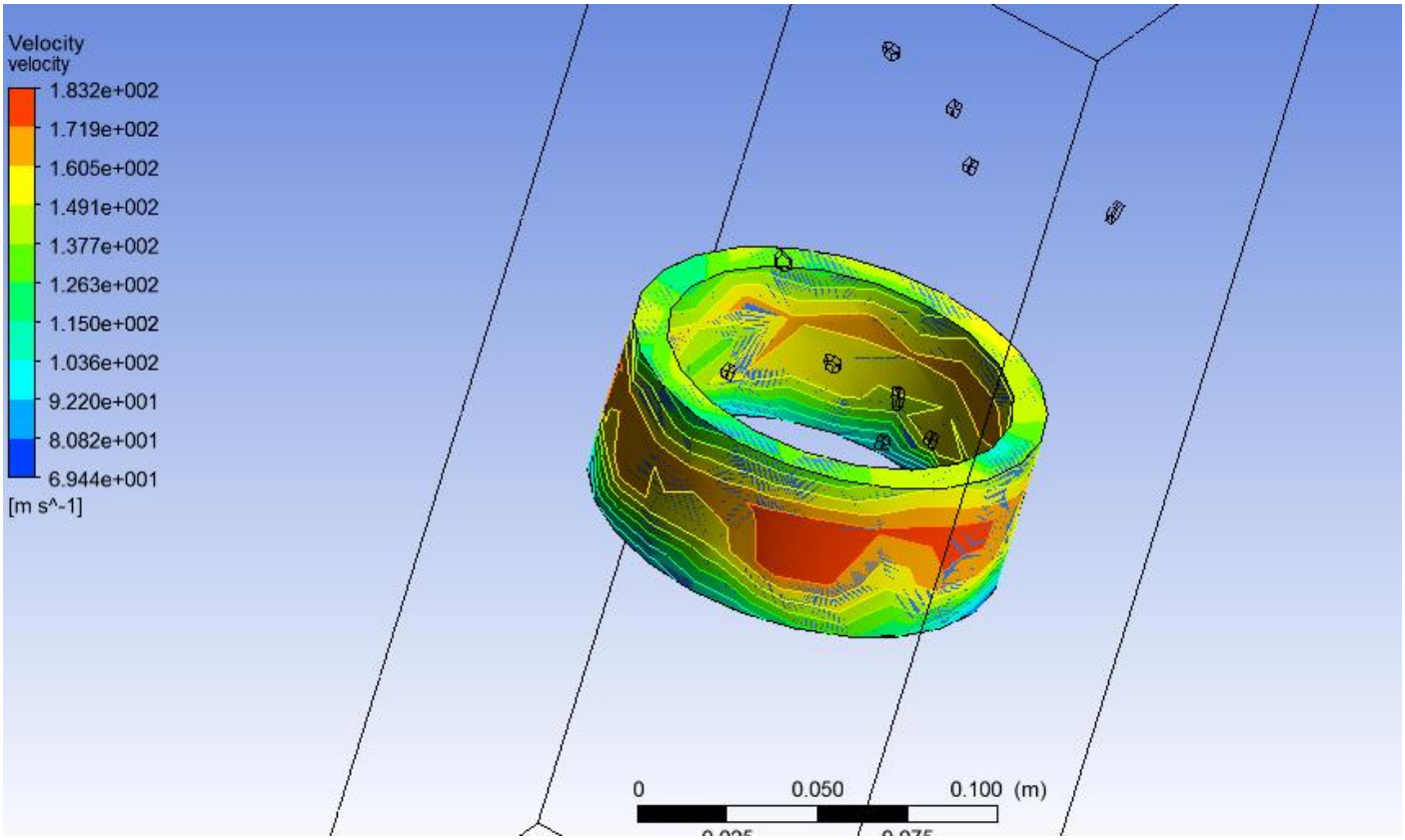


Fig 9 *VELOCITY OF POLYMER MEDIA*

IJSTR©2019
www.ijstr.org

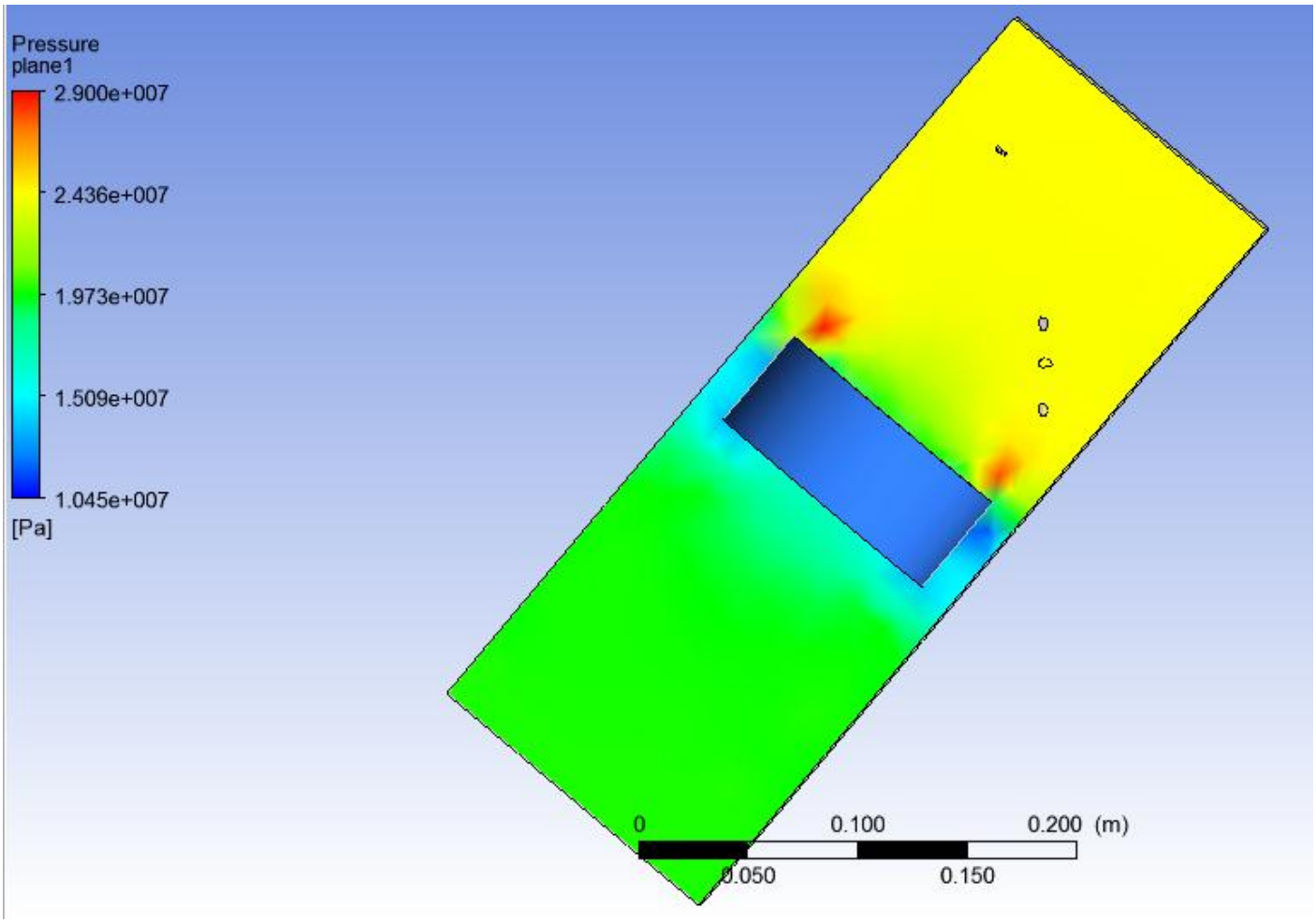


Fig 10 *PRESSURE VARIATION ALONG POLYMER MEDIA*

REFERENCES

[1] Kang, Junmo., Yamaguchi, Hitomi: Internal finishing of capillary tubes by magnetic abrasive finishing using a multiple pole-tip system, Precision Engineering 36 (2012) 510– 516, 2012.

[2] Jain, V. K: Magnetic field assisted abrasive based micro-/nano- finishing, Journal of Materials Processing Technology 209 (2009) 6022–6038, 2009.

[3] Wani, Amit. M., Yadava, Vinod., Khatri, Atul: Simulation for the prediction of surface roughness in magnetic abrasive flow finishing (MAFF), Journal of Materials Processing Technology 190 (2007) 282–290.

[4] Kala, Prateek., Pandey, Pulak. M., Verma, Girish. C., Sharma, Varun: Understanding flexible abrasive brush behaviour for double disk magnetic abrasive finishing based on force signature, Journal of Manufacturing Processes 28 (2017) 442–448, 2017.

[5] Wu Jinzhong., Zou, Yanhua., Sugiyama, Hitoshi: Study on finishing characteristics of magnetic abrasive finishing process using low-frequency alternating magnetic field, Int J Adv Manufacturing Technology 85:585– 594 DOI 10.1007/s00170-015-7962-9, (2016).

[6] Rathod, Avishkar., Sapate, S. G., Khatirkar, R. K: Shape factor analysis of abrasive particles used in slurry abrasion testing, International Journal of Mechanical and Industrial Engineering (IJMIE) ISSN No. 2231 –6477, Vol-2, Iss-4, 2012.

[7] Srivastava, Ankit. Kumar., komma, V.R: Finite Element Modelling of Electrochemical Magnetic Abrasive Finishing, Imperial Journal of Interdisciplinary Research (IJIR), Vol- 2, Issue-12, 2016, ISSN: 2454-1362.

[8] Srivastava, Ankit. Kumar., Katiyar. Amit.: Modelling and Simulation of Magnetic Abrasive Finishing for Thermal Analysis, Imperial Journal of Interdisciplinary Research (IJIR) Vol-2, Issue-12, 2016 ISSN: 2454-1362.

[9] Mr. Krovvidi Srinivas is presently working as Assistant Professor in Delhi Technological University. He is having a total teaching experience of 24 years. He guided around 75 UG projects and 7 PG projects. He published around 10 research papers in different conferences and Journals.He is a Member of Institution of Engineers (I), Life member of Indian Society for Technical Education, Life member of Powder Metallurgy Association of India.

[10] Presently, Dr. Qasim Murtaza is a Professor in the Department of Mechanical Engineering, Delhi Technological University, Delhi, India. He got his Ph.D (2006) in the area of manufacturing process, funded

by European Union as Research Assistant, from Metal Thermal Spray Lab, Dublin City University, Ireland. He has also published several papers in reputed SCI journals and conference proceedings.

[11] Dr. A K Agrawal is presently working as Professor in the Mechanical Engineering Department, Delhi Technological University. He is Having more than 30 years of Teaching and research experience. He published several papers in International journals and Conferences.

[12] Singh, Rajneesh. Kumar., Singh, D.K., Gangwar. Swati.: Fuzzy logicoptimization of process parameters affecting surface temperature in MAF, IOP Conf. Series: Materials Science and Engineering 404 (2018) 012017 doi:10.1088/1757-899X/404/1/012017.

[13] Jaiswal, S.C., Jain, V. K., Dixit, P.M.: Modelling and simulation ofmagnetic abrasive finishing process, Int J Adv Manufacturing Technology (2005) 26: 477–490, DOI 10.1007/s00170-004-2180.

[14] Yan, Q. S., Qiu, T.X., Gao, W.Q.,Meng, L: Simulation and Experimental Analysis of Electromagnetic Inductor for Magnetic Abrasive Finishing, Key Engineering Materials Vols. 389-390 (2009) pp 193-198 Online: 2008-09-26 © (2009) Trans Tech Publications, Switzerland, doi: 10.4028/www.scientific.net/KEM.389- 390.193.

[15] Li, Wenhui., Li Xiuhong., Yang, Shengqiang., Li, Weidong: A newly developed media for magenetic abrasive finishing process: Material removal behaviour and finishing performance, Journal of Materials Processing Tech 260(2018) 20-29.

[16] 13.Li, Xiuhong., Li, Wenhui., Yang, Shengqiang.: Effect of coupling Agent on Interfecial Bonding Properties of Visecoelastic Magnetic Abrasive Tools and Finishing Performance, ISSN 1473-804 online, 1473-8031 print, DOI 10.5013/IJSSST.a.17.28.17.

[17] 14.Shinmura T, Takazawa K, Hatano E, Matsunaga M. (1990) Study on magnetic abrasive finishing. Ann CIRP, 39(l): 325-328.

[18] 15.Shinmura T., Takazawa K., Hatano E., Aizawa T. (1985) Study on magnetic abrasive process-process principles and finishing possibility, Bull. Japan Soc. Prec. Eng. 19(1) 54–55.

[19] 16.Kim JD., Choi MS. (1995) Simulation for the prediction of surface accuracy in magnetic abrasive machining. J Mater Process Tech 53:630–642.

[20] 17.Srinivas. K., Murtaza. Q., Aggarwal. A.K.: Variation of Flow parameters and Magnetic Flux in Viscoelastic Magnetic Abrasive Finishing process, ISSN 0973-4562 Volume 14, Number 8, pp. 1940-1946, (2019).

[21] 18.Srinivas. K., Bhardwaj, Anant.: Pressure variation in Abrasive Flow machining: Modelling and simulation, 3rd international conference on Advanced Production and Industrial Engineering (ICAPIE'18), Paper ID ICAPIE-2018-PE MT 218.

[22] Ali, Parvesh., Ranganath, M.S., Walia, R.S., Murtaza, Q.: Various developments in Abrasive flow machining process: A Review, 3rd international conference on Advanced Production and Industrial Engineering (ICAPIE'18), Paper ID ICAPIE-2018- PE MT 215.

[23] Bhardwaj, Anant., Ali, Parvesh., Walia, R. S., Murtaza, Qasim., Pandey, S. M.: Development of Hybrid Forms of Abrasive Flow Machining Process: A Review, Advances in Industrial and Production Engineering pp 41-67, Part of the Lecture Notes in Mechanical Engineering book series (LNME), conference paper (2019).

New Failure Rate Model for Iterative Software Development Life Cycle Process

Sangeeta, *Member, IEEE,* Kapil Sharma, *Member, IEEE,* and Manju Bala

***Abstract*—Software reliability models are one of the most generally used mathematical tool for estimation of reliability, failure rate and number of remaining faults in the software. Existing software reliability models are designed to follow waterfall software development life cycle process. These existing models do not take advantage of iterative software development process. In this paper, a new failure rate model centered on iterative software development life cycle process has been developed. It aims to integrate a new modulation factor for incorporating varying needs in each phase of iterative software development process. It comprises imperfect debugging with the possibility of fault introduction and removal of multiple faults in an interval as iterative development of the software proceeds. The proposed model has been validated on twelve iterations of Eclipse software failure dataset and nine iterations of Java Development toolkit (JDT) software failure dataset. Parameter estimation for the proposed model has been done by hybrid Particle Swarm Optimization and Gravitational Search Algorithm. Experimental results in-terms of goodness-of-fit shows that proposed model has outperformed Jelinski Moranda, Shick Wolverton, Goel Okummotto Imperfect debugging, GS Mahapatra, Modified Shick Wolverton in 83.33 % of iterations for eclipse dataset and 77.77% of iterations for JDT dataset.**

***Index Terms*—Software development life cycle, Iterative software development life cycle, Optimization, Nature-inspired algorithms, Software reliability models.**

ACRONYMS

SDLC	**Software Development Life Cycle**
LLF	**Log Likelihood Function**
MLE	**Maximum Likelihood Estimation**
SSE	**Sum of Squared Error**
MSE	**Mean Square Error**
JM	**Jelinski Moranda**
PSO-GSA	**Particle Swarm Optimization and Gravitational Search Algorithm**

Sangeeta is with Delhi Technological University, New Delhi, India (e-mail: sangeeta@dtu.ac.in)

K. Sharma is with Delhi Technological University, New Delhi, India (kapil@ieee.org)

Manju Bala is with Indraprastha College for Women, Delhi University, New Delhi, India (manjugpm@gmail.com)

NOTATIONS

$\lambda(t_i)$	**Failure Intensity**
$f(t_i)$	**Probability Density Function**
$F(t_i)$	**Cumulative Distribution Function**
$R(t_i)$	**Reliability Function**
L(N)	**Likelihood Function**
γ	**Modulation factor for representing changing needs in an iteration of software development**
μ	**Modulation parameter that represents newly added functionality and user acceptance in an iteration**
N	**Number of initial faults in iteration**
p	**Probability of fault removal in iteration**
r	**Probability of fault introduction in iteration**
n_{i-1}	**Cumulative number of failures at $(i-1)^{th}$ failure interval**

I. Introduction

With the growing advances in the digital world, software development demand from industries is growing at an exponential rate. Due to enormous demand and lack of time and budget, software companies are not able to develop fault-free software. Latest tools and techniques have been applied for the development of defect-free software, but still, it is not possible for software developers to develop defect-free software practically. Software must go through exhaustive testing and debugging, which requires time and money to enhance the reliability [1] [*2*]. The occurrence of fault is inevitable in the current demand of software. There should have some means to avoid software failures so that devastating losses whether related to life or any other field could be evaded.

According to IEEE standard 729 [*3*], reliability is the most significant quality aspect of the software. If we could measure the reliability of software under development, better we can predict whether the software would be operational in the future or not. Reliability estimation process must be precise to provide information to the manager like what should be the release time of the software and amount of man-hour consumption etc. while developing any software [4]. Software reliability models are one of the ways to simulate software reliability estimation curve to predict the reliability of the system under study. Numerous reliability estimation models for software have been developed, and all are working on specific applications, specific environments, datasets and

assumptions made by them. Then, what is the need for developing new software reliability models?

First need lies in the fact that, among the available research in software reliability model development [5] [6] [7] [8] [9] [10], all developed models are basically made for reliability estimation of software's developed under traditional waterfall SDLC process. Second, research publications are screening no more than 31% experimental researches and among this only 13 % are purely experimental [11]. This low number is due to the reason that public experimental data sets in software reliability estimation are very narrow and producing reliability data through experimentation usually require long time cluster. Third, most generally developed models are NHPP based and very few failure rate based models [12] [13] [14] [15] [16] [17] [18] [19] [20] [21] [22] [*23*].

Failure rate-based models are among the earliest software reliability estimation models used in industrial and academia. These models are grounded on JM software reliability model [9]. They need further enhancement, so that more realistic assumptions like imperfect debugging and factors for exact reliability growth estimation can be incorporated into the model. Further, all software reliability models are developed under waterfall SDLC process. Waterfall SDLC process for the development of software assumes that requirements from the end users are stable and it delivers whole software at one shot in the end [24] [25]. This may generate risks for the users as they do not have information till end what they will get. These limitations construct a need for the use of another methodology while modeling software reliability. Yet now, no research in software reliability assessment has been found which are based on latest SDLC process.

Software reliability is primarily dependent on the process of software development. The software developed using latest SDLC approach can capture and implement all the user requirements within time and budget. Earlier waterfall SDLC process based reliability estimation models are very much stable and found to be unsuitable for the software industries because there are high risk and uncertainty of change in requirements. An iterative SDLC process [26] is the latest methodology of software development and is a practical method of step-wise top-down refinement approach to the software development that replaces the waterfall SDLC process. It gathers user requirements in each iteration of software development and implements software with less uncertainty in risk and user satisfaction. Nowadays the latest software development processes are based on iterative software development process and these are like Rational Unified Process (RUP) [27], Adaptive method [28] [*29*] (Agile methodology), XP (Extreme Programming) [*30*], spiral [31] model and AZ [*32*] development processes. These SDLC processes are considered as an aid of producing reliable software based on iterative model in development. Moreover, these are found to be used practically in industries and academia for software development and promote iterative development process.

Further accurate software reliability estimation is dependent on the selection of optimum parameter values of the models. The techniques for optimizing parameter values of the software reliability models are available, through the various classical methods of parameter estimation [5] [*33*]. These methods are based on number of constraints and may fall in local maxima and do not converge to global maxima in the multimodal cases. An alternative to these classical mathematical optimization methods are nature-inspired optimization algorithms for the solution of non-differential, non-linear and multimodal problems [34] [35] [36].

The proposed failure rate model incorporates iterative SDLC process by replacing earlier waterfall based SDLC process. Further, it assumes imperfect debugging during each of iteration. There is always a possibility of fault introduction with feature addition in each of the iterations of software development. All latest iterative SDLC processes can be used to predict the reliability by applying the proposed failure rate model. Existing failure rate model cannot be applied to the current Software development methodology. The proposed model takes care of complexity and paradigm shift of iterative based software development process by introducing modulation factor.

Rest of the paper is schematized as follows: Section 2 exposes the existing literature of the software reliability models and parameter estimation algorithms used for software reliability modeling. In section 3 a new classification scheme for software reliability models based on iterative SDLC process is discussed. In Section 4 a new software reliability model is proposed. Section 5 discusses the experimental setup and results of the proposed model and the last section concludes the paper.

II. Literature Review

For reliability estimation of the software more than 300 SRGMs have been developed under various categories of reliability estimation models like [6] [*5*] [37] [*38*] [39] [40] [41] [42] [43]. Experimental research work for software reliability estimation is found to be very few, reason being limited failure data and requisite of long time period. From the developed models, no model is well applicable in all type of applications and fits only with a specific application. Some of the model groups are having assumptions that are considered better and have higher predictive quality than other models. Most of the model development comprises NHPP based models and very few are centered on the failure rate behavior. Literature in this paper discusses well-known failure rate behavior-based models.

Jelinski-Moranda model [9] is the first software reliability estimation model. This model is perfect debugging based and assumes that there is fixed, constant and unknown number of initial faults in the software. The time between failures is assumed as independent and exponentially distributed. Shick-Wolverton (SW) model [10] assumed that failure rate function is proportional to the current total number of faults and the time since last failure. This idea has been evolved after modification from the basic JM Model. L.I.A. Turk and E.G. Alsolami developed a model using JM model by applying Weibull distribution function for the amount of debugging time between fault occurrences [44]. Modified Shick-Wolverton model [6] [5] [12] modified the SW model by incorporating a cumulative amount of faults at a time. These

models incorporate perfect debugging process while making assumptions for the model development. Goel and Okumotto first time proposed imperfect debugging based model that incorporates imperfect debugging phenomenon by introducing fault removal probability [5]. JM Geometric Model [13] modified JM model by assuming a geometric decrease in program failure rate at failure times. Littlewood and Sofer [14] proposed a model with Bayesian modification in the JM model and provided improvement for estimating the correct reliability of the software. Further modification to the JM model based on the cloud theory has been applied by Luo, Cao, Tang, and Wu [15]. Y.C. Chang and C. Liu proposed a generalized model by extending the JM model using a truncated distribution that matches up a self-exciting point process [45]. Mahapatra and Roy proposed a model and assumed the imperfect debugging behavior of the software with fault removal probability, probability of fault not removed and probability of fault introduction [16]. This model introduced fault introduction probability while removing any fault during debugging.

In failure rate behavior-based models, JM model assumed perfect debugging with single failure between time intervals after that multiple failure were proposed by SW but considers perfect debugging. Imperfect debugging based GOI model consider single failure in a time interval. Littlewood and Luo modified JM model according to the Bayesian and cloud theory respectively. Mahapatra *et al*. introduced probability for fault introduction in the software. In these enhancements to the earliest failure rate model, multiple faults with imperfect debugging have not been covered by any model. These entire model enhancements are made for software's developed under most basic waterfall development life cycle process. There is a need to incorporate latest software development processes for precise estimation of software reliability.

Parameter estimation process plays a crucial role in estimating accurate reliability of software product. There are number of models which are based on the classical methods for parameter estimation [33] [46]. These are numerical technique based methods (like Newtonian method, Gauss-Siedal method, etc.) [47], and these methods are doing well in a number of mathematical optimization problems. Nowadays meta-heuristic based optimization methods are found to be performing well in different domains. The algorithms like Particle Swarm Optimization [48] inspired by the bird flocks and Genetic Algorithm [49] inspired by the natural evolution phenomenon have been well applied in various fields including the field of software reliability estimation. A. Sheta and Jordan Al-Salt [50] applied PSO for SRGM parameter estimation. Malhotra and Negi applied PSO for reliability modeling [51]. Jin and Jin [52] applied improved swarm intelligent approach for parameter optimization of SRGM based on s-shaped testing effort function. Hybrid nature-inspired algorithms are found to be performing well in various domains. Mirjalili and Hashim proposed a hybrid PSO-GSA algorithm by hybridizing PSO and Gravitational Search Algorithm for mathematical function optimization [53]. Abraham *et al*. proposed a hybrid differential and ABC algorithm [54]. Mirjalili and Wang proposed a binary optimization for hybrid PSO-GSA algorithm [55]. As hybrid algorithms are performing well in different fields, these algorithms could be applied in software reliability model parameter estimation so that accurate software reliability could be estimated.

III. Classification of Software Reliability Models

This paper classifies software reliability models based on iterative SDLC process. Iterative model is like a cyclic process that mainly focuses on initial very simple implementation that progressively gains complexity and wider set of features till the final system is completed. Most recently built iteration and its feedback from evaluation are used in the next iteration, and accordingly, refinements are made in future iterations. Each of iteration provides enhancements and at-least found to be better than the last. Iterative model adapts rapidly to ever-changing needs of projects and clients within lowest time and budget.

Sharma *et al*. [56] has given classification of software reliability models according to waterfall SDLC process. In this paper, keeping iterative software development practice in mind software reliability models are categorized in Fig. 1. Every phase in iterative software development process is associated with specific requirements and future plans. Fig. 1 shows that each of the iteration is associated with a modulation parameter showing the added functionality and user acceptance level in each iteration. Initially, system is least reliable, and as the number of iterations proceeds, system moves towards refinement and gains reliability. Models are grouped into five categories. Each model category is assigned to a specific phase in iterative software development process for reliability estimation.

IV. Proposed Model

Keeping new SDLC processes and technologies, a new failure rate model centered on iterative SDLC process has been proposed. The proposed model has derived from the general class of failure rate behavior-based models and exploits iterative behavior of software development process.
In iterative process of software development, entire software is built and delivered to the end user in iterations. The process starts with simple implementation of key samples of a problem and iteratively enhances existing software releases until full system is implemented [57]. At each of iteration release feedback from the iteration is available for next iteration [58] [59] [60]. Feedback is mainly about the functionality and user aspect of the software. At each released stages of software, not only extensions but design modifications may be made. Each of the iteration makes step wise refinements in an effective way to converge to the full implementation of a problem. Main focus of feedback analysis is to find the amount of defects in an iteration that gets injected in the upcoming iteration [61]. Using feedback in each iteration, analysis of changing needs in an up-coming iteration can be made. These changing needs in each iteration are essential to know because these tells information about how much more defects may occur and how much more

effort and functionality is needed in making extensions and design modifications to the released iteration.

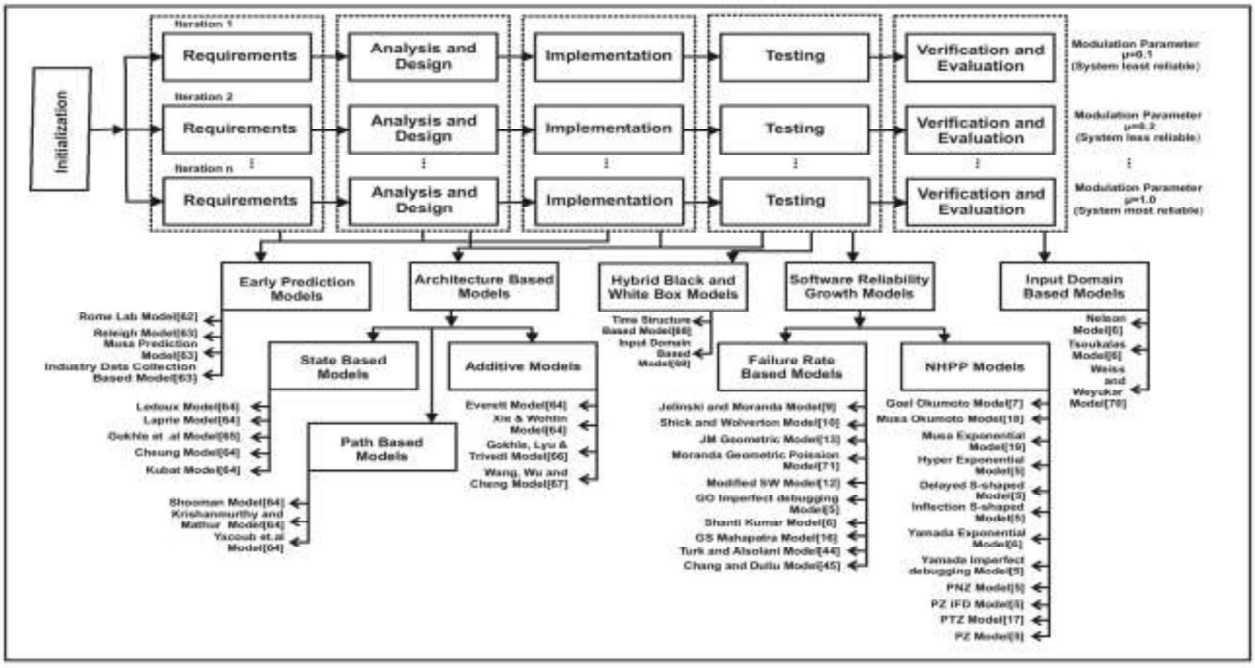


Fig. 1. Classification of Software Reliability Models based on Iterative approach.

In modeling software reliability, there is a need to introduce a factor that will engulf all the changing needs depending on the defect analysis in each of iteration. The varying needs in each of iteration of software development include bug reports, additional functionalities, and testing effort required to find the amount of defects that gets injected / removed in each upcoming iteration. These changing needs are incorporated in proposed failure rate model using modulation factor γ given in (1), so that precise reliability growth of the system could be estimated. The proposed model assumes that fault removal process is imperfect. Due to imperfect debugging regenerated faults are induced in successive iterations.

A. Proposed Model Assumptions

i. Initial software fault is unknown and constant in an iteration.
ii. Each fault in an iteration is independent and it may be equally likely to cause a failure while testing.
iii. The interval of time between fault occurrences in each iteration is independent and follows an exponential distribution.
iv. Software failure rate remains constant over the intervals between fault occurrences.
v. The software failure rate is proportional to the number of faults that remain in the software and modulation factor γ.
vi. In each iteration the detected fault is removed with a probability p, not removed perfectly with a probability q and new fault may be introduced with a probability r. Here $p+q+r=1$ and the probability $p > r$.

Faults are injected from previous iterations along with newly introduced faults in the current iteration. Newly induced faults are caused by added and modified functionalities in a respective iteration of software development. Depending on the number of initial iterative faults, there is a need to modify the amount of resources allocated for debugging in each iteration. The modulation factor γ as defined in (1) reflects the modified needs that integrate iterative development processes in software failure rate models. Modulation factor changes its value according to the modulation parameter μ as shown in (1). Moreover, changing needs in each iteration is different and vary according to (1).

$$\gamma = \mu + (1-\mu)/\mu, \quad (0 < \mu \le 1.0) \tag{1}$$

Here, μ is the modulation parameter that represents newly added functionality and user acceptance in the current iteration. Its value is almost 0 at the beginning and becomes 1.0 in the final iteration of iterative software development process. Modulation parameter μ takes its value from 0 to 1.0 by assuming that with growing number of iterations, level of user acceptance increases from lower to higher. It shows abrupt changes in its value in initial iterations due to preliminary design changes, testing effort, and user acceptance. When it's values reaches near to one then software under development is assumed to be reliable enough and has achieved all the required functionalities to fulfill the end user needs.

B. Model Formulation

Failure rate function ($\lambda(t_i)$) with imperfect debugging is modeled in (2).

$$\lambda(t_i) = \phi[\mathrm{N} - \frac{(n_{i-1})(\gamma)}{(i-1)}(p\text{-}r)], \quad i\text{-}1, 2....N \tag{2}$$

Where

γ	Modulation factor for representing changing needs in each of the iteration of software development
$n_{i\text{-}1}$	Cumulative number of failures at $(i\text{-}1)^{th}$ failure interval
N	Number of initial faults in software
ϕ	Proportionality constant

Cumulative Density Function $\mathrm{F}(t_i)$ and Reliability Function $\mathrm{R}(t_i)$ is calculated in (3) and (4)

$$F(t_i) = 1 - e^{[-\phi[\mathrm{N} - \frac{(n_{i-1})\gamma}{i-1}(p\text{-}r)]t_i]} \tag{3}$$

$$R(t_i) = e^{[-\phi[\mathrm{N} - \frac{(n_{i-1})\gamma}{i-1}(p\text{-}r)]t_i]} \tag{4}$$

When $p = 1, r = 0$ and γ varies as in (5), proposed model behaves as JM model.

$$1, \frac{2i}{i+1}, \frac{3i}{i+4}, \frac{4i}{i+8}, \frac{5i}{i+13}.... \tag{5}$$

Following a variation of γ in (5) and considering p being the probability of fault removal and r as the fault introduction probability then model behaves as the GS Mahapatra *et al.* model [*16*].

C. Parameter Estimation

In the proposed model there are three unknown parameters N, n , and ϕ these parameters are estimated at different values of the γ . MLE has been used to estimate the values of the parameters. Parameter estimation by MLE method requires solutions of complex equations by maximizing the likelihood of model parameters. Probability density function $\mathrm{f}(t_i)$ for the proposed model is given in (6).

$$\mathrm{f}(t_i) = \phi[\mathrm{N} - \frac{(n_{i-1})\gamma}{i-1}(p\text{-}r)] \cdot e^{[-\phi[\mathrm{N} - \frac{(n_{i-1})\gamma}{i-1}(p\text{-}r)]t_i]} \tag{6}$$

The likelihood function $\mathrm{L(N)}$ is calculated in (7) using (6).

$$\mathrm{L(N)} = \prod \mathrm{f}(t_i) = \prod_{i=1}^{n} [\phi[\mathrm{N} - \frac{(n_{i-1})\gamma}{i-1}(p\text{-}r)] e^{[-\phi \sum_{i=1}^{n} [\mathrm{N} - \frac{(n_{i-1})\gamma}{i-1}(p\text{-}r)]t_i]} \tag{7}$$

Taking the log of $\mathrm{L(N)}$, LLF is calculated in (8).

$$LLF = \ln L(N) = n \ln \phi + \sum_{i=1}^{n} \ln[N - \frac{(n_{i-1})\gamma}{i-1}(p-r)] - \sum_{i=1}^{n} \phi[N - \frac{(n_{i-1})\gamma}{i-1}(p-r)]\, t_i \tag{8}$$

Solution using log likelihood function for parameter estimation involves calculation of its partial derivatives with respect to N, n and ϕ respectively and then equating them to value zero. MLE of the parameters are calculated from (9), (10) and (11).

$$\frac{n}{\phi} = \sum_{i=1}^{n} [N - \frac{(n_{i-1})\gamma}{i-1}(p-r)]\, t_i \tag{9}$$

$$n = \phi \sum_{i=1}^{n} [N - \frac{(n_{i-1})\gamma}{i-1}(p-r)]\, t_i \tag{10}$$

$$\frac{1}{\sum_{i=1}^{n} [N - \frac{(n_{i-1})\gamma}{i-1}(p-r)]} = \phi \sum_{i=1}^{n} t_i \tag{11}$$

V. Experimental Setup and Results

A. Application Datasets used for Experimentation

Suitability of the proposed model has been tested using Tera Promise repository bug report files of Eclipse (DS1) and JDT (DS2) open source software at https://zenodo.org/record/268486#.W-QsPpMzY2w. These data sets have been given by An Ngoc Lam. The bug reports contains table having contents as bug_id, summary, description, report time, report time status, commit, and commit time files. Datasets have been extracted from these bug reports and reformatted in time domain format. DS1 includes eight minor releases and four major releases starting from the year 2001 to the year 2013. DS2 includes 3 major releases and 6 minor releases starting from the year 2002 to the year 2014.

Model parameters include N, n, ϕ, γ, p and r . Fault removal probability p and fault introduction probability r cannot be estimated from DS1 and DS2 and depends on the project type and skill set of persons involved in development and testing. Based on these two factors fault removed during testing is assumed 95 % and fault introduced is assumed 3%. Parameters N, n, ϕ, γ are estimated using hybrid PSO-GSA algorithm and MLE technique in (6), (7) and (8) to maximize the log-likelihood function value using (9), (10) and (11). The goodness of fit for the proposed model is measured using SSE and MSE for each application datasets [56]. These statistics are used for comparison of the proposed model with existing failure rate models.

The values of μ with respect to iteration are different for different software projects and depend on the type of project and user acceptance levels. Fig. 2 depicts variation in μ values with respect to number of iterations for DS1. The value of μ decreases in successive iterations 1.0 to 2.0, 2.1 to 3.0 and 3.3 to 3.4 by difference of 0.064, 0.081 and 0.045 respectively. These little deeps in μ values represent

Table 1 SUMMARY OF FAILURE RATE BASED SOFTWARE RELIABILITY MODELS

Sr. No.	Model Name	Failure Intensity
1	Jelinski-Moranda Model (JM) [9]	$\lambda(t_i)=\phi[\mathrm{N}-(i-1)]$
2	Schick &Wolverton Model (SW) [10]	$\lambda(t_i)=\phi[\mathrm{N}-(i-1)]t_i$
3	Goel Okumotto imperfect debugging Model(GOI) [5]	$\lambda(t_i)=\phi[\mathrm{N}-p(i-1)]$
4	G.S.Mahapatra *et al.* Model [16]	$\lambda(t_i)=\phi[N-p(i-1)+r(i-1)]$
5	Modified S-W Model(MSW) [5]	$\lambda(t_i)=\phi[\mathrm{N}-(n_{i-1})]t_i$
6	Proposed Model	$\lambda(t_i)=\phi[\mathrm{N}-\frac{(n_{i-1})(\gamma)}{(i-1)}(p-r)]$

variation in compliance of functionality requirements and lower values of user acceptance levels for eclipse software dataset (DS1). Overall values of μ shows increasing trends with successive iteration as shown in Fig. 2 and finally it reaches near to 1.0 in final iteration release.

Fig. 3 is revealing the change in functionalities and its corresponding increase in user acceptance for DS2 of JDT software product. At each of the iteration release there is increase in user acceptance level and it is illustrated with the values of μ in all successive iterations. For major iteration releases the value of μ are 0.0869, 0.1108 and 0.2144, for minor iteration these values are 0.1381, 0.2684, 0.3677, 0.6787, 0.7241 and 0.9539. The user acceptance increases in all iterations but a large variation of 0.31102 in user acceptance level is found in 3.4 to 3.5 iteration. It represents enhanced functionality and accomplishment of all requirements at end-users. Overall values of μ in successive iterations are reflecting variation in functionalities and user acceptance level in Fig. 3 that increases by small amount in each iteration and finally reaches near to 1.0 in last iteration release.

B. Result Analysis

1) CASE 1: Eclipse Software Dataset (DS1)

In this section, the goodness-of-fit of the proposed model is calculated and compared it with existing models given in Table 1 for DS1. The value of γ for major iterative release 1.0, 2.0, 3.0 and 4.1 are 2.2666, 2.8763, 3.7891 and 1.1258 respectively. In minor iteration release, the value of γ for iterative release 2.1, 3.1, 3.2, 3.3, 3.4, 3.5, 3.6 and 4.2 are 2.6343, 3.2217, 2.5788, 2.3121, 2.7295, 2.6434, 2.0435 and 1.00009 respectively.

The value of γ shows large changes in successive major iterations as compared to successive minor iterations due to varying needs in major and minor iterations. The goodness of fit of the models is shown in Table 2. From the analysis of results, it is found that proposed model fits well in term of SSE in all iterations of DS1 except at iteration number 3.1 where GOI model, GS Mahapatra model, and SW model outperforms proposed model.

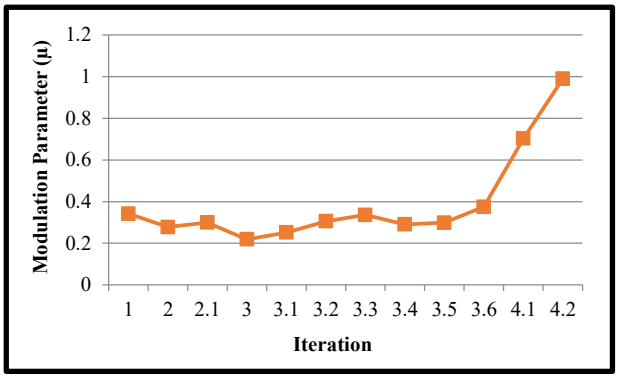


Fig. 2 Plot of μ versus Iteration for DS1

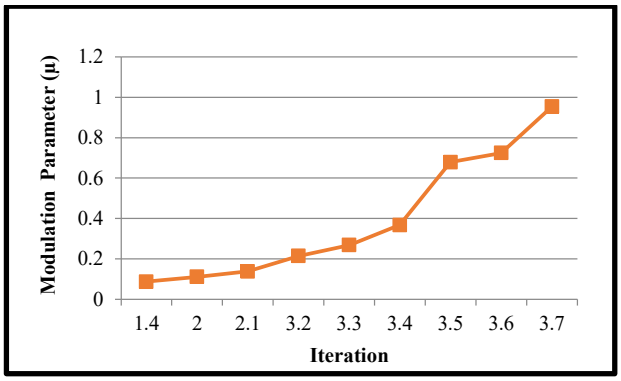


Fig. 3 Plot of μ versus Iteration for DS2

The proposed model has clear-cut outperform all models under comparison in 11 iterations for DS1. It shows the lowest values of SSE in 91.6 % iterations. In term of MSE proposed model is winner in nine iterations. In iteration 2.1 MSW model is performing well than the proposed model. In iteration 3.3, GOI model, SW model, and GS Mahapatra model are performing well than the proposed model. There is a tie among proposed model and GS Mahapatra model in iteration 2.0, where both of them outperform all other models. The proposed model has clear-cut outperform other models in 75 % iterations by achieving lowest value of MSE. Result shows that proposed model has given a significantly better fit to iterative data by adapting according to varying needs of different iterations.

The evaluation and comparison of goodness-of-fit of all models in Table 1 in terms of SSE and MSE for all iterations shows that the proposed model has promising technical merit in the sense that it provides development terms with both iterative SDLC requirements and traditional reliability measures.

2) CASE 2: JDT Dataset

To test the applicability of the proposed model in this case, dataset DS2 is used. Table 3 is showing the estimated values of model parameters and goodness of fit criteria. The value of γ for major iterative release 1.4, 2.0 and 3.2 are 10.5954, 8.1358 and 3.8778 respectively. In minor iteration releases the value of γ for iterative release 2.1, 3.3, 3.4, 3.5, 3.6 and 3.7 are 6.3815, 2.9946, 2.0872, 1.1521, 1.1052 and

1.0022, respectively. These values of γ are showing large deviations in successive major releases as compared to successive minor releases due to the varying needs in each of the major and minor iterations. Table 3 is showing the goodness-of-fit measures of all models depicted in Table 1 for DS2. Out of total nine iterations in DS2, proposed model outperform other models under comparison in eight iterations, in terms of SSE values. JM model performs well than proposed model in iteration 3.4. Proposed model attains lower-most value of SSE in 88.8 % of iterations. In terms of MSE values for nine iterations proposed model out-performs other models in six iterations. In iteration 2.0, GS Mahapatra model and GOI model are performing better than the proposed model. In iteration 3.7, proposed model outperforms JM, GOI, SW and MSW models except GS Mahapatra model. The proposed model outperforms GOI , SW, GS Mahapatra and MSW model except JM model in iteration number 3.6. The proposed model has performed better than all other models under comparison in 66.6% iterations by attaining the minimum value of MSE. Result analysis in Table 3 shows that the proposed model has significantly better fit to iterative data by fine-tuning to varying needs of different iterations.

Table 2 GOODNESS-OF-FIT ES/TIMATED USING DS1 (ECLIPSE SOFTWARE FAILURE DATASET)

		Iteration1.0			Iteration2.0			Iteration2.1			Iteration 3.0		
Sr. No.	Model	Estimated Parameter values	SSE	MSE	Estimated Parameter values	SSE	MSE	Estimated Parameter values	SSE	MSE	Estimated Parameter values	SSE	MSE
1	JM Model	Φ=5.74E-06, N=4	5.23	5	Φ=2.68E-05, N=8	375.51	17.05	Φ=5.039E-06, N=27	452.2	16.74	Φ=2.40E-06, N=124	56948	605.82
2	GOI Model	Φ=1.64E-05, N=3	1.51	1	Φ=2.89E-06, N=28	92.01	4.18	Φ=1.95E-06, N=26	133.34	4.93	Φ=1.37 E-06, N=100	45510.1	559.22
3	SW Model	Φ=2.124E-06, N=3	17.31	17.00	Φ=4.70E-06, N=28	218.81	9.91	Φ=3.83E-06, N=29	126.6	4.66	Φ=2.92 E-05, N=121	39981	494.69
4	GS Mahapatra Model	Φ=7.77E-07, N=3	1.26	1	Φ=2.53 E-05, N=23	55.46	2.5	Φ=3.20E-06, N=28	131.31	4.85	Φ=1.25 E-06, N=125	55122.2	586.40
5	MSW Model	Φ= 2.23E-05, N=2, n=58	5.0	4.9	Φ=1.52E-06, N=9, n=209	528.7	15.89	Φ=1.36E-05, N=18, n=316	118.67	3.89	Φ=1.29E-06, N=123, n=298	66890.4	679.32
6	Proposed Model	Φ= 2.86E-05, N=5, n=11, ϒ=2.2666,	1.24	0.66	Φ=4.70E-06, N=27, n=362, ϒ=2.8763,	45.12	2.5	Φ=1.67E-06, N=30,n=415, ϒ=2.6343,	101.02	4.39	Φ=2.98E-05, N=102,n=5690, ϒ=3.7891	39705.2	484.14

		Iteration3.1			Iteration3.2			Iteration3.3			Iteration 3.4		
Sr. No.	Model	Estimated Parameter values	SSE	MSE	Estimated Parameter values	SSE	MSE	Estimated Parameter values	SSE	MSE	Estimated Parameter values	SSE	MSE
1	JM Model	Φ=2.99 E-05, N=100	6091.1	45.45	Φ=2.21E-06, N=122	13940.73	119.14	Φ=2.95E-05, N=117	6810.14	58.24	Φ=8.28E-06, N=53	1153.32	23.53
2	GOI Model	Φ=3.07E-07, N=126	1018.8	7.59	Φ=8.61E-06, N=106	7246.05	61.93	Φ=1.25E-06, N=115	9371.1	80.09	Φ=3.06E-06, N=50	1545.41	1320.9
3	SW Model	Φ=2.99E-05, N=125	1123	8.38	Φ=5.40E-06, N=90	12585	107.56	Φ=2.95 E-05, N=123	7318.81	62.55	Φ=2.93 E-05, N=54	1153.73	23.53
4	GS Mahapatra Model	Φ=6.62E-06, N=130	1015	7.57	Φ=1.25E-06, N=115	7794.28	68.97	Φ=3.06E-06, N=114	2217.4	18.99	Φ=1.63E-05, N=55	905.4	18.46
5	MSW Model	Φ=1.36E-05, N=135, n=4979	2000.7	15.9	Φ=2.08E-06, N=132, n=4108	9685.12	89.67	Φ=5.31E-06, N=121,n=3129	2903.9	27.87	Φ=1.72E-05, N=52, n=2094	3589.96	29.79
6	Proposed Model	Φ=2.96E-05, N=136,n=6132, ϒ=3.2217	1589.82	12.22	Φ=2.98 E-05 , N=122,n=6910, ϒ=2.5788	952.2	8.14	Φ=2.96E-05, N=123,n=7879, ϒ=2.3121	2203.05	19.49	Φ=2.92 E-05, N=53, n=1411, ϒ=2.7295	520.23	11.55

Table 2 continued…

Table 2 continued…

		Iteration3.5			Iteration3.6			Iteration4.1			Iteration 4.2		
Sr. No.	Model	Parameter estimated values	SSE	MSE	Parameter estimated values	SSE	MSE	Parameter estimated values	SSE	MSE	Parameter estimated values	SSE	MSE
1	JM Model	Φ=1.63E-06, N=30	373.3	15.54	Φ=6.24E-07, N=28	199.91	7.65	Φ= 2.20 E-05, N=19	286.65	22	Φ=1.92E-05, N=31	441.12	15.21
2	GOI Model	Φ=3.81E-06, N=29	137.74	5.71	Φ=3.208E-06, N=28	160.01	6.15	Φ=1.58 E-05, N=17	111.21	8.54	Φ=3.20E-06, N=30	121.21	4.17
3	SW Model	Φ=1.70E-07, N=29	551.1	22.96	Φ=3.74E-06, N=27	154.54	5.92	Φ=1.93E-05, N=19	284.86	22.1	Φ=6.50E-06, N=31	193.34	6.66
4	GS Mahapatra Model	Φ=3.81E-06, N=29	137.8	5.71	Φ=6.65E-07, N=28	152.23	5.86	Φ=5.53E-06, N=17	129	9.92	Φ=7.29E-07, N=31	166.65	5.72
5	MSW Model	Φ=1.26E-05, N=17, n=2663	489.74	18.24	Φ=2.97E-05, N=17, n=2624	315.00	14.92	Φ=2.68E-05, N=8, n=1529	178.22	7.34	Φ=2.97E-05, N=30, n=3729	704.86	26.87
6	Proposed Model	Φ=2.82 E-05, N=25, n=416, ϒ=2.6434	86.62	4.3	Φ=2.94E-05, N=30, n=401, ϒ=2.0435	91.15	4.14	Φ=2.87E-05, N=16,n=136, ϒ=1.1258	49.95	5.44	Φ=2.82E-05, N=32, n=513, ϒ=1.00009	72.24	2.88

Table 3 GOODNESS-OF-FIT ESTIMATED USING DS2 (JDT SOFTWARE FAILURE DATASET)

		Iteration 1.4			Iteration 2.0			Iteration 2.1			Iteration 3.2		
Sr. No.	Model	Estimated Parameter values	SSE	MSE	Estimated Parameter values	SSE	MSE	Estimated Parameter values	SSE	MSE	Estimated Parameter values	SSE	MSE
1	JM Model	Φ= 1.46E-05, N=22	1510	62.91	Φ= 5.63E-08, N=13	1792.56	66.37	Φ= 2.81E-05, N=37	2638.12	203.02	Φ= 2.35E-06, N=50	179.79	11.19
2	GOI Model	Φ= 4.11E-06, N=15	1517	75.85	Φ= 4.80E-08, N=18	1486.81	55.04	Φ= 5.53E-06, N=35	1650.04	127.90	Φ= 8.16E-07, N=56	317.67	5.56
3	SW Model	Φ= 8.18E-07 N=9	2165	90.21	Φ= 6.99E-06, N=17	2427.45	187.9	Φ= 5.36E-06, N=38	3278.45	252.87	Φ= 2.17E-06, N=58	1056	15.68
4	GS Mahapatra Model	Φ= 9.00E-06 N=12	1871	77.96	Φ= 1.08E-06, N=16	1434.79	53.11	Φ= 1.21E-06, N=57	3808.85	293.48	Φ= 1.21E-06, N=57	56500.09	991.23
5	MSW Model	Φ=7.43E-06 N=19,n=242.98	2909.88	121.09	Φ=1.62E-05, N=15,n=282.93	1870.67	69.26	Φ=8.73E-06, N=29,n=2569.9	1741.90	164.48	Φ=2.91E-05, N=52,n=1723	2989.06	29.69
6	Proposed Model	Φ=2.99E-05, N=13, n=160, ϒ=10.5954	1379	57.54	Φ=2.97E-05, N=17, n=199, ϒ=8.1358	1292	56.34	Φ=2.82E-05, N=37, n=234, ϒ=6.3815	683.70	75.88	Φ=2.91E-05, N=60, n=1788, ϒ=3.8778	171.89	3.226

Table 3 continued…

		Iteration 3.3			Iteration 3.4			Iteration 3.5			Iteration 3.6		
Sr. No.	Model	Estimated Parameter values	SSE	MSE	Estimated Parameter values	SSE	MSE	Estimated Parameter values	SSE	MSE	Estimated Parameter values	SSE	MSE
1	JM Model	Φ= 2.78E-05, N=9	404	33.666	Φ=8.86E-06, N=18	37	12.33	Φ=2.49E-05, N=8	57	4.384	Φ=2.93 E-05, N=4	76	7.6
2	GOI Model	Φ=1.48E-05, N=16	989.9	61.825	Φ=2.35E-06, N=18	206.56	17.16	Φ=1.06E-06, N=13	76.53	25.33	Φ=2.01E-05, N=11	46721	359.39
3	SW Model	Φ=4.70E-06, N=8	589.1	40.78	Φ=3.83E-06, N=16	750.9	68.00	Φ=2.92 E-05, N=10	3019.9	119.05	Φ=2.99E-05, N=29	12089	367.83
4	GS Mahapatra Model	Φ=8.92E-06, N=26	604.01	37.75	Φ=1.03E-05, N=28	1241.01	103.41	Φ=4.78 E-07, N=5	38225	318.54	Φ=1.26E-05, N=85	11534	384.48
5	MSW Model	Φ=2.99E-05, N=9, n=1028	486.32	33.960	Φ=2.91E-05, N=13, n=1124	1839.56	146.92	Φ=2.31E-05, N=3, n=137	20687	156.94	Φ=1.34E-05, N=15, n=134	34834.09	329.21
6	Proposed Model	Φ=2.97E-05, N= 19,n=179, ϒ=2.9946	31	2.583	Φ=2.92E-05, N=18, n=119, ϒ= 2.0871	91	11.37	Φ=2.96E-05, N=7, n=25, ϒ=1.1521	18	-24.002	Φ=2.67E-05, N=18, n=171, ϒ=1.1051	192	21.33

Table 3 continued…

Iteration 3.7				
Sr. No.	Model	Estimated Parameter values	SSE	MSE
1	JM Model	Φ=1.50E-06, N=16	280	45.02
2	GOI Model	Φ=1.58E-06, N=13	90.011	9.01
3	SW Model	Φ=5.40E-06, N=26	200.9	11.08
4	GS Mahapatra Model	Φ=5.30E-06, N=12	952	8.14
5	MSW Model	Φ=2.69E-05, N=14, n=1599	1109.23	57.34
6	Proposed Model	Φ=2.74E-05,N=15, n=78,ϒ=1.0022	54	9.00

VI. CONCLUSION

Failure rate models available in literature are centered on most traditional waterfall SDLC process. However, new software development processes have been developed and found to be more beneficial than waterfall SDLC process, like iterative life cycle processes. Keeping in view, new software development environments and technologies, a new failure rate model by exploiting iterative behavior of software development process is proposed. The changing needs in each of the iteration are reflected in the proposed model using a modulation factor. Calculated values of γ parameter are significantly reflecting all changing requirements for each upcoming iterations numerically. These values are meaningfully representing how much impact is of adding and removing new features with the level of user acceptance in each of upcoming iteration. In order to compare the performance of the proposed model, five well-known software failure rate models JM, GOI, SW, GS Mahapatra and MSW have been applied for dataset DS1 and DS2. The proposed model is a clear-cut winner in 11 iterations in SSE and 9 iterations in MSE out of 12 iterations and 8 iteration in SSE and 6 iterations in MSE out of 9 iterations for DS1 and DS2 respectively. Overall, in 83.33 % of iterations for DS1 and 77.77 % of iterations for DS2, the proposed model has shown better results in terms of goodness of fit by successfully incorporating varying needs in each of iteration.

The data collected from real applications and comparison of goodness of fit shows that the proposed model successfully incorporated varying needs in each of iteration and performed better than JM, GOI, SW, GS Mahapatra and MSW models.

VII. REFERENCES

[1] Prabhat Ranjan Sushant Kumar, "A phase wise approach for fault identification," *Journal of Information and Optimization Sciences*, vol. 39, no. 1, 2018.

[2] Xuelin and Wong, W Eric and Gao, Ruizhi and Hu, Linghuan and Hosono, Shigeru Li, "Genetic Algorithm-based Test Generation for Software Product Line with the Integration of Fault Localization Techniques," *Empirical Software Engineering*, vol. 23, no. 1, pp. 1-51, 2018.

[3] "Glossary of software Engineering Terminology," IEEE standard 729, 1983.

[4] K. Sharma, C. Nagpal, R. Garg, R. Garg, R. Kumar and Sandhaya R. Garg, "Ranking of software engineering matrics by fuzzy based matrix methodology," *Software testing ,Verification and Reliability* , vol. 23, no. 2, pp. 149-168, March 2013.

[5] H. Pham, *Software Reliability Modeling, London: Springer Series in Reliability Engineering, 2006.*: Springer Series in Reliability Engineering, 2006.

[6] M.Xie, *Software Reliability Modeling*.: World Scientific Publishing Co. Ltd., 1991.

[7] A. L. Goel and K. Okumoto, "Time-Dependent Error-Detection Rate Model for Software Reliability and Other Performance Measure," *IEEE Transaction on Reliability, Vols. R-28, no. 3*, pp. pp. 206-211, 1979.

[8] R. K. Srivastava Kavita Sahu, "Revisiting Software Reliability," *Data Management, Analytics and Innovation* , pp. 221-235, 2018.

[9] z Jelinski, "software reliability reserach," in *statistical computer performance evaluation*, 1972, pp. 465-484.

[10] G.H.Schick R.W.Wolverton, "An Analysis of competing Software reliability Models," *IEEE Transaction on Software Engineering*, pp. 104-120, March 1978.

[11] A.Macedo,R.Matias,L.Borges J.Xavier, "A survey on Research in Software Reliability Engineering in the last Decade," in *SAC-14*, Gyeongju,Korea, March 14-28,2014.

[12] A.N.Sukert, "An investigation of software reliability models," in *in proceeding of annual reliability and maintainability symposium,IEEE reliability society*, Piscataway, 1977.

[13] P.B. Moranda, "An error detection model for application during software development," *IEEE Transaction on software reliability*, vol. R-28, no. 5, 1979.

[14] P. B. littlewood and P. A. Sofer, "Bayesian Modification to the Jelinski Moranda software reliability growth model," *Software engineering* , pp. 30-45, 1987.

[15] P.Cao,G.Tangand L.Wu Z.Luo, "Jm model based on cloud theory," 2011.

[16] P.Roy G.S. Mahapatra, "Modified Jelinski-Moranda Software reliability model with imperfect debugging Phenomenon," *International journal of computer application*, vol. 48, no. 18, pp. 38-46, june 2012.

[17] H.Pham, X.Zhang, and X.Teng, ""Considering Fault Removal Efficiency in Software Reliability assessment.," *"IEEE Transaction on Systems Man and Cybernetics part A :System and humans, vol. 33, no. 1*, pp. pp. 114-120, 2003.

[18] J.D. Musa K.Okumoto, ""A Logarithmic Poisson Execution Time Model for Software Reliability Measurment"," *IEEE, Whippany*, 1984.

[19] M.R.Lyu and S.Y.Kuo C.Y.Huang, "A unified Scheme of some Non-homogenous Poisson Process models for software reliability estimation," *IEEE Transaction on Software Engineering*, vol. 29, no. 3, pp. 261-269, March 2003.

[20] Georgi Spasov,Asen Rahnev,Nikolay Kyurkchiev Nikolay Pavlov, "A NEW CLASS OF GOMPERTZ–TYPE SOFTWARE," *International Electronic Journal of Pure and Applied Mathematics*, vol. 12, no. 1, pp. 43-57, 2018.

[21] G. Mahapatra and K. N. Dey P. Roy, ""An NHPP software reliabilty growh model with imperfect debugging and error generation,"" *International Journal of Reliability, Quality and Safety Engineering, vol. 21, no. 2*, pp. pp. 1450008-1450040, 2014.

[22] Meera Sharma, H. Pham V.B. Singh, "Entropy Based Software Reliability Analysis of Multi-Version Open Source Software," *IEEE TRANSACTIONS ON SOFTWARE ENGINEERING*, 2017.

[23] Yunwei Dong, Xuelin Li, W. Eric Wong Xiaomin Wei, "Architecture-level hazard analysis using AADL," *Journal of Systems and Software*, vol. 37, pp. 580-604, 2018.

[24] Barry W and others Boehm, *Software engineering economics*.: 197, Prentice-hall Englewood Cliffs (NJ), 1981.

[25] Dr. Winston and W.Royce, "Managing the development of large software systems," in *Reprinted from Proceedings, IEEE WESCON*, August 1970, pp. 1-9.

[26] Albert J.Turner Victor R.Basil, "Iterative Enhancement: A Practical Technique for Software Developemnt," *IEEE Transaction On Software Engineering*, vol. SE-1, no. 4, DEC 1975.

[27] P. Kruchten, "The Rational Unified Process—An Introduction.," in *Addison-Wesley*, 2000.

[28] Craig Larman, *Agile and iterative development: a manager's guide*. 2004: Addison-Wesley Professional.

[29] Letha H. Etzkorn, Sampson Gholston, Stephen Quattlebaum Hector M. Olague, "Empirical Validation of Three Software Metrics Suites to Predict Fault-Proneness of Object-Oriented Classes Developed Using Highly Iterative or Agile Software Development Processes," *IEEE Transaction on Software Engineering*, vol. 33, no. 6, JUNE 2007.

[30] P. Abrahamsson, "Extreme programming: First results from a controlled case study," 2003.

[31] Barry Boehm, "A spiral model of software development and enhancement," *ACM SIGSOFT Software engineering notes*, vol. 11, no. 4, pp. 14-24, 1986.

[32] Jun Sang,Arif Ali Khan,Fazal-E-Amin,Nasrullah,Muhammad Shafiq,Shahid Hussain, Haibo Hu, Manzoor Elahi, Hong Xiang Muhammad Azeem Akbar, "Improving the quality of software development process by introducing a new methodology—AZ-Model," *DOI:10.1109/ACCESS.2017.2787981, IEEE Access*, pp. 1-9, 2017.

[33] I. J. Myung, ""Tutorial on maximum likelihood estimation" ," *Journal of Mathematical Psychology*, pp. vol. 47, pp. 90-100, 2003.

[34] Samia Nefti Meziani ,Adham Atyabi Mohd nadhir Ab Wahab, "A comprehensive review of Swarm optimization algorithms," *PLOS ONE*, pp. 1-36, february 2015.

[35] Xin-She Yang, *Nature-inspired metaheuristic algorithms*.: Luniver press, 2010.

[36] Xin-she yang, "Nature inspired metaheuristic algorithms:Success and new challenges," *J Comput Eng Inf.Technol*, pp. vol1,issue 1 pg 1-3, 2012.

[37] M.ohba, "Inflection S-Shaped Software reliability Growth Model," stochastic models in reliability theory, 1984.

[38] H. Ohtera and H. Narihisa S. Yamada, "Software Reliability Growth Models with Testing -Effort," *IEEE Transaction on Reliability*, vol. R-35, no. 1, pp. 19-23, 1986.

[39] H. Othera and M. Ohba S. Yamada, "Testing-Domain Dependent Software Reliability Models," *Computers Mathematics Applications*, vol. 24, pp. 79-86, 1992.

[40] M. Ohba and S. Osaki S. Yamada, "S-Shaped Reliability Growth Modeling for Software Error Detection," *IEEE Transaction on Reliability, Vols. R-32, no. 5*, pp. pp. 475-478, 1983.

[41] M.Fumio,T.Kumiko and M.Yoshiharu X. Jianwen, ""An Imperfect Fault Coverage Model with Coverage of Irrelevant Components," *IEEE Transaction on Reliability,vol.64,No.1*, March 2015.

[42] P. Kapoor and S. Younues, "Modelling an Imperfect Debugging Phenomenon in Software Reliability," in *Microelectronic Reliability Elsevier Science Ltd*, vol. 36, 1996, pp. 645-655.

[43] H.Pham,A.G and Gurjeet P.K.Kapoor, "Two Dimensional Multi Release Software Reliability Modeling and Optimal Release Planning," *IEEE Transaction on Reliability*, vol. 61, no. 3, September 2012.

[44] E.G.Alsolami L.I. AI Turk, "Jelinski Moranda Software rEliability Growth Model :A Brief Literature and Modification," *International Journal of Software Engineering & Applications*, vol. 7, no. 2, March 2016.

[45] C.Ti Liu Y.C. Chang, "A Generalized JM Model with applications to Imperfect debugging in Software REliability," *Applied Mathematical Modeling*, vol. 33, pp. 3578-3588, 2009.

[46] M.Staron ,C.Berger,J.Hansson,M.Nilson,F.Torner R.Rana, "Comparing between maximum likelihood

estimator and Non Linear regression estimation procedures for NHPP software reliability growth models," in *Joint conference of 23rd IWSM and eighth international conference on software process and product measurement* , Mensura, 2013, pp. 213-218.

[47] SK Gupta, *numerical methods for Engineers*. kanpur: New age international pvt , 1995.

[48] J Kennedy and R Eberhart, "Particle swarm optimization," in *ieee international conference on neural networks [c] piscataway*, NJ USA, 1995.

[49] Holland JH, "Genetic Algorithm," *SCI Am1992*, vol. 267, pp. 66-72.

[50] A.sheta, "Parameter estimation of software reliability growth models by particle swarm optimization," *AIML.j*, pp. 7(1) 55-61, 2007.

[51] Ruchika Malhotra and Arun Negi, "Reliability modeling using Particle Swarm Optimization," *International Journal of System Assurance Engineering and Management*, vol. 4, no. 3, pp. 275-283, September 2013.

[52] Cong Jin and Shu-Wei Jin, "Parameter optimization of software reliability growth model with S-shaped testing-effort function using improved swarm intelligent optimization," *Applied Soft Computing*, vol. 40, pp. 283-291, 2016.

[53] Seyedali Mirjalili, Mohd Hashim, and Siti Zaiton, "A new hybrid PSOGSA algorithm for function optimization," in *2010 International Conference on Computer and Information Application (ICCIA 2010)*, 2010, pp. 374-377.

[54] Ajith,jatoth,ravi kumar,Rajasekhar,A. Abraham, "Hybrid Differential Artificial Bee Colony Algorithm ," *Journal of computational and theoritical nanoscience*, vol. 9, no. 2, pp. 249-257, february 2012.

[55] Seyedali Mirjalili and Gai-Ge Wang, "Binary optimization using hybrid particle swarm optimization and gravitational search algorithm," *Neural Computing & Applications*, vol. 25, no. 6, pp. 1423-1435, May 2014.

[56] Kapil Sharma, Rakesh Garg, CK Nagpal, and RK Garg, "Selection of Optimal Software Reliability Growth Models using Distance Based Approach," *IEEE Transaction on Reliability*, vol. 59, no. 2, pp. 266-277, June 2010.

[57] Craig and Basili, Victor R Larman, "Iterative and incremental developments. a brief history," *Computer*, vol. 36, no. 6, pp. 47-56, 2003.

[58] P.K.Kappor and R.B. Garg, "Optimal release policies for software systems with testing effort," *international journal of System and science*, vol. 22, no. 9, pp. 1563-1571, 1991.

[59] M.R. Lyu C.Y.Huang, "Optimal Release time for software systems considering cost,testing effort and testing efficiency," *IEEE Transaction on Reliability*, vol. 54, no. 4, 2005.

[60] D.Grees and G.Ruhe, ""Software Release Planning:An Evolutionary and Iterative approach," *Information and software Technology*, vol. 46, no. 4, pp. 243-253, 2004.

[61] W Eric and Gao, Ruizhi and Li, Yihao and Abreu, Rui and Wotawa, Franz Wong, "A survey on software fault localization," *IEEE Transactions on Software Engineering*, vol. 42, no. 8, pp. 707-740, 2016.

[62] "Rome Laboratory, Methodology for software reliability prediction and assessment Technical Report RADC-TR-87-171," 1987.

[63] M. Stutzke and R. W.stoddard C. Smidts, "Software Reliability Modeling :An approach to early prediction," *IEEE Transaction on Reliability*, vol. 47, no. 3, pp. 268-278, 1998.

[64] K. Popstajanova K.Trivedi, "Architecture based approach to reliability assessment of software systems," *Performance Evalvation*, vol. 45, no. 2, pp. 179-204, 2001.

[65] Swapna S and Wong, W Eric and Horgan, Joseph Robert and Trivedi, Kishor S Gokhale, "An analytical approach to architecture-based software performance and reliability prediction," in *Computer Performance and Dependability Symposium, 1998. IPDS'98. Proceedings. IEEE International*, 1998, pp. 13-22.

[66] Swapna S and Lyu, Michael R and Trivedi Kishor S Gokhale, "Analysis of software fault removal policies using a non-homogeneous continuous time Markov chain," *Software Quality Journal*, vol. 12, no. 3, pp. 211-230, 2004.

[67] M.H.Chen and W.Wang Y.Wu, "An Architecture based Software Reliability Model," in *In Proceeding of Pacific Rim Dependability Symposium*, Piscataway,IEEE Reliability Society, 1977.

[68] S.S.Gokhale and K.S. Trivedi, "A Time Structure Based Model of Software Reliability," *Annals of Software Engineering*, vol. 8, pp. 85-121, 1999.

[69] Swapna S Gokhale Peter N and Trivedi, Kishor S Marinos, "Important milestones in software reliability modeling," *Proceedings of Software Engineering and Knowledge Engineering (SEKE'96), Lake Tahoe, NV*, pp. 345--352, 1996.

[70] S.N.Weiss and E.J.Weyuker, "An extended domain based model of software reliability ," *IEEE transaction on software engineering*, vol. SE-14, pp. 1512-1524, 1988.

[71] B.Littlewood, "Software Reliability model for modular program structure ," *IEEE transaction on Reliability*, vol. R-28, no. 3, 1979.

Journal Pre-proofs

Regular paper

On the Physical Layer Security over Beaulieu-Xie Fading Channel

Puspraj Singh Chauhan, Sandeep Kumar, Sanjay Kumar Soni

PII: S1434-8411(19)31626-7
DOI: https://doi.org/10.1016/j.aeue.2019.152940
Reference: AEUE 152940

To appear in: *International Journal of Electronics and Communications*

Received Date: 28 June 2019
Accepted Date: 1 October 2019

Please cite this article as: P.S. Chauhan, S. Kumar, S.K. Soni, On the Physical Layer Security over Beaulieu-Xie Fading Channel, *International Journal of Electronics and Communications* (2019), doi: https://doi.org/10.1016/j.aeue.2019.152940

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2019 Published by Elsevier GmbH.

Journal Pre-proofs

On the Physical Layer Security over Beaulieu-Xie Fading Channel

Puspraj Singh Chauhan[a], Sandeep Kumar [1a,*], Sanjay Kumar Soni[a]

[a]*Department of Electronics and Communication Engg., M. M. M. University of Technology, Gorakhpur, Uttar Pradesh, India*
[1a] *CRL, BEL Ghaziabad, Uttar Pradesh, India*

Abstract

With the expansion of IoT and device-to-device communication applications, the security and privacy of the information have been kept at the forefront. The existence of a secondary eavesdropper, which attempts to access the confidential information of the legitimate partners, makes the data communication insecure. In this regard, we study the effect of randomness caused by the Beaulieu-Xie fading channel on the performance of physical layer security. To this end, the analysis for the secrecy outage probability (SOP) and average secrecy capacity (ASC) is carried out. Further, we derived the closed-form expressions for lower-bound of secrecy outage probability and strictly positive secrecy capacity (SPSC) for the said channel. To get more insight into the system, the high-power and low-power analysis of the ASC are also provided. Besides, the impact of different fading parameters on the system matrices is also investigated. Finally, the theoretical results are verified via extensive Monte Carlo simulations.

Keywords: Secrecy capacity; physical layer security; fading channel; secure communications

1. Introduction

With the advancement of 5G communications, where the increasing number of smart devices in IoT and device-to-device networking applications, the issue of privacy and security has been considered as the forefront requirement for the design engineer [1]. These applications work in the broadcast mode so data communications over these networks are more vulnerable to intruders i.e., a threat to the privacy and security of the host. There are some traditional cryptographic techniques, which are used for the secure transmission of information [2] between the two legitimate partners. These traditional techniques assume that the link between the trans-

*Corresponding author
Email addresses: puspraj.chauhan@gmail.com (Puspraj Singh Chauhan), sann.kaushik@gmail.com (Sandeep Kumar [1]), sksoni98fece@gbpec.ac.in (Sanjay Kumar Soni)

mitter (Tx) and receiver (Rx) was error-free, which was not always the case. There have been attempts in the literature where the random nature of signal propagation was captured using different multipath and shadowing models. In some cases, the channel estimation and modeling over different frequencies and propagation scenarios have been studied [3, 4, 5, 6]. More recently, attempts have been made to use the random nature of the wireless channel to guarantee the secure transmission of messages [7, 8]. The model of the wire-tap channel was first proposed by Wyner in which two familiar nodes interact with each other over the main channel, and a third party tries to access the information of the legitimate partners [9]. In this model, the trade-off between the information rate and the achievable secrecy level for a wire-tap channel has been proposed. The secrecy performance of wire-tap channel was provided in terms of the secrecy capacity (SC), which was defined as the difference between the capacity of the main channel and the wire-tap channel. In other words, it can be defined as the maximum rate at which the secure data transmission can be achieved. Recently, there has been an increasing interest in studying the SC of the wireless systems over different fading scenarios. Secrecy analysis over Nakagami-*m* fading channel in the presence of imperfect channel state information (CSI) was studied in [10]. The physical layer security performance of the mobile vehicular networks over *N*-Nakagami fading channels was proposed in [11]. The effect of various fading parameters on the secrecy performance of generalized fading channel has been investigated in [1, 12, 13, 14, 15]. The closed-form expressions for strictly positive secrecy capacity (SPSC) and lower-bound of secrecy outage portability (SOP) for $\kappa-\mu$ fading channel has been presented in [1] while the probability of positive secrecy capacity and upper-bound of SOP has been studied in [12]. Recently in [13], average SC (ASC) and SOP over the more generalized $\alpha-\eta-\kappa-\mu$ fading channel has been analysed. Although the results produced in [13] are novel but not in closed-form. Closed-form expressions for the SOP and ASC over the Fox-H and mixture Gamma distribution are derived and discussed in [14, 15]. In [16] and [17], secrecy performance analysis over generalized-*K* (GK) fading channel was conducted. The authors have derived the expressions of SOP, SPSC and average SC in closed-form over generalized-*K* (GK) fading channels using a mixture Gamma distribution [17]. This work was extended for single-input-multiple-output (SIMO) system in [18]. In [19] SOP performance over correlated Nakagami-*m*/Gamma composite fading channel was analyzed while $\kappa-\mu$ shadowed composite fading channel was considered in [20] and [21] for studying the performance of physical layer security. The expressions for SOP and SPSC were derived in closed-form in terms of Meijer's-

G function in [20]. A similar analysis for SPSC using moment matching method was presented in [21].

Given the benefits of the Ricean and Nakagami-*m* fading models, these models have the limitations in characterizing the broad range of signal-fades from reflections and shadowing in the case of Ricean model and both LOS and NLOS components in the case of Nakagami-*m* model. To overcome these challenges, both these models were recently merged into a new fading model by Beaulieu and Xie [22]. The Beaulieu-Xie fading model acquires the benefits of the Nakagami-*m* fading model, in that it has a flexible fading parameter and the benefits of the Ricean fading model, in that its use of a non-central chi-distribution allows it to characterize both LOS and NLOS channels. This new fading model has been shown to have a relationship with the generalized Ricean distribution and the $\kappa - \mu$ distribution [22], which makes the performance analysis of this fading model to be of utmost importance. Moreover, the new fading models can characterize the indoor propagation scenario in a femtocell or small cell areas and high-speed train channels that has applications in 5G and beyond networks [22]. The second-order characteristic of the Beaulieu-Xie fading channel has been derived in [23], while the asymptotic tight bound of the said correlated channel for different diversity combining techniques were evaluated in [24, 25]. The physical layer security performance of Beaulieu-Xie fading channel is not present in the literature and yet to be done.

Therefore we have derived the performance matrices of physical layer security over Beaulieu-Xie fading channel in this work. The main contributions of our work are listed as follows:

- We obtained and analysed the physical layer security performance parameters like ASC, SOP, and SPSC over independent Beaulieu-Xie fading channel.
- We propose new closed-form expressions for lower-bound of SOP over Beaulieu-Xie fading channels.
- The high power analysis of the ASC has been provided. The ASC result for low power 5G network applications has also been investigated.
- All the results obtained from derived expressions are verified with the help of Monte-Carlo simulations.

The benefits of our study is to provide the efficient analytical expressions over Beaulieu-Xie fading channel that do not exist in the literature for the secrecy performances as functions of system parameters. In doing this, insights on the system performances can be obtained, as shown in Section 4.

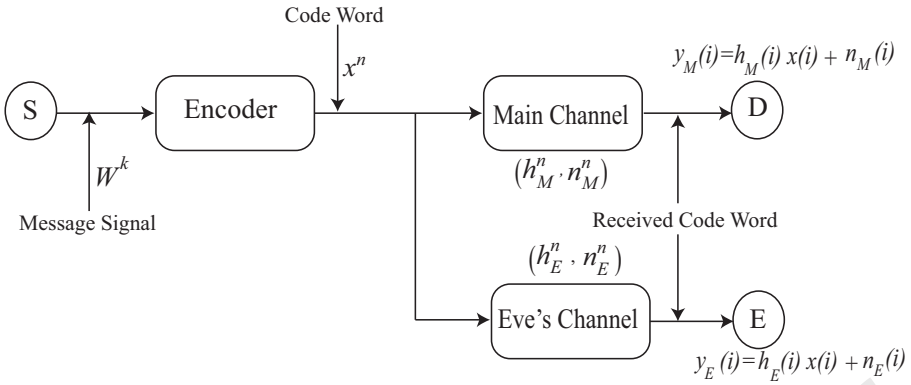


Figure 1: System model for SISO channel

The rest of this paper is organized as follows: In Section 2, the system model is introduced. Section 3 derives the expressions of the performance parameters. Numerical results and discussion are given in Section 4. Finally, Section 5 concludes this work.

2. System Model and fundamental channel statistics

The system model of the secure data transmission over single-input and single-output (SISO) channel is shown in Fig. 1. The legitimate source node (S) wishes to communicate secretly with the legitimate destination node (D), while an Eve node (E) is attempting to eavesdrop. The channel statistics of both the main and eavesdropper channel is assumed to undergo Beaulieu-Xie fading with CSI available at Tx. At the Tx, the message W^k is encoded into a code word, x^n for transmission over the channel. The signal at the receiving end, can be written as

$$y(i) = h(i)x(i) + n(i) \tag{1}$$

where $h(i)$ and $n(i)$ are the quasi-static Beaulieu-Xie fading coefficients and additive white Gaussian noise (AWGN) random variable with zero mean and unit variance at the receiving ends, respectively

The signal-to-noise-ratio (SNR) probability density function (PDF) of Beaulieu-Xie fading is given by [26, eq. (2)]

$$f_Y(\gamma) = \left(\frac{\sqrt{2}}{\lambda}\right)^{m-1} \left(\sqrt{\frac{m}{\bar{\gamma}}}\right)^{m+1} \gamma^{\frac{m}{2}-\frac{1}{2}} e^{-\frac{\lambda^2}{2}} e^{-\frac{m\gamma}{\bar{\gamma}}} I_{m-1}\left(\sqrt{\frac{2m\gamma}{\bar{\gamma}}}\lambda\right) \tag{2}$$

where $I_a(.)$ is the modified Bessel function of type 1 and order a. The fading parameters m, $\bar{\gamma}$, and λ controls the shape, spread, height, and location of the PDF.

The definition of the cumulative distribution function (CDF) is given by [27, eq. (2)]

$$F_Y(\gamma_{th}) = \int_0^{\gamma_{th}} f_Y(\gamma) d\gamma \tag{3}$$

where γ_{th} is the threshold SNR. Now substituting (2) into (3) and with the aid of [28, eq. (8.445)] and [28, eq. (3.381.1)] yields

$$F_Y(\gamma_{th}) = e^{-\frac{\lambda^2}{2}} \sum_{l=0}^{\infty} \frac{\left(\frac{\lambda}{\sqrt{2}}\right)^{2l}}{\Gamma(l+1)\Gamma(l+m)} \mathcal{G}\left(m+l, \frac{m\gamma_{th}}{\bar{\gamma}}\right) = 1 - Q_m\left(\lambda, \sqrt{\frac{2m\gamma_{th}}{\bar{\gamma}}}\right) \tag{4}$$

where $\Gamma(.)$ is the Gamma function, $\mathcal{G}(,)$ is the lower incomplete gamma function and $Q_m(,)$ is the generalised Marcum-Q function.

3. Secrecy analysis

In this section, we derive analytical expressions of SC, SOP, and strict positive secrecy capacity (SPSC) along with approximate solutions.

3.1. Secrecy capacity

The maximum data rate achieved under the eavesdropper scenario where perfect secrecy of the transmitted message is guaranteed a key assumption of the physical layer security. The instantaneous secrecy capacity is given by [1, eq. (11)]

$$S_c = \begin{cases} \log_2(1+\gamma_M) - \log_2(1+\gamma_E) & \gamma_M > \gamma_E \\ 0 & \gamma_M \leq \gamma_E \end{cases} \tag{5}$$

where $\log_2(1+\gamma_M)$ and $\log_2(1+\gamma_E)$ are the capacity of the main and the eavesdropper channel with instantaneous SNR γ_M and γ_E, respectively and $\log_2(.)$ is the logarithmic function with base 2.

Under the practical environment where fading arises the average secrecy capacity $\bar{S}_c$ (the average of the maximum transmission rate reached in a secure channel) expression is given by

[17]

$$\bar{S}_c = \int_0^{\infty} \log_2(1+\gamma_M) f_M(\gamma_M) \left(\int_0^{\gamma_M} f_E(\gamma_E) d\gamma_E \right) d\gamma_M - \int_0^{\infty} \log_2(1+\gamma_E) f_E(\gamma_E) \left(\int_{\gamma_E}^{\infty} f_M(\gamma_M) d\gamma_M \right) d\gamma_E$$
$$= \underbrace{\int_0^{\infty} \log_2(1+\gamma_M) f_M(\gamma_M) F_E(\gamma_M) d\gamma_M}_{I_1} + \underbrace{\int_0^{\infty} \log_2(1+\gamma_E) f_E(\gamma_E) F_M(\gamma_E) d\gamma_E}_{I_2} - \underbrace{\int_0^{\infty} \log_2(1+\gamma_E) f_E(\gamma_E) d\gamma_E}_{I_3} \tag{6}$$

where $f_E(\gamma_E)$, $f_M(\gamma_M)$ are the PDFs of the eve and the main channel, respectively. Putting (2) and (4) into I_1 and (2) into I_3, we get

$$I_1 = \mathcal{A}_m \int_0^{\infty} \log_2(1+\gamma_M) \gamma_M^{\frac{m_M}{2}-\frac{1}{2}} e^{-b\gamma_M} I_{m_M-1}\left(\sqrt{2b\gamma_M}\lambda_M\right) \left(1 - Q_{m_E}\left(\lambda_E, \sqrt{2\tilde{b}\gamma_M}\right)\right) d\gamma_M \tag{7}$$

$$I_3 = \mathcal{A}_E \int_0^{\infty} \log_2(1+\gamma_E) \gamma_E^{\frac{m_E}{2}-\frac{1}{2}} e^{-\tilde{b}\gamma_E} I_{m_E-1}\left(\sqrt{2\tilde{b}\gamma_E}\lambda_E\right) d\gamma_E \tag{8}$$

where $b = \frac{m_M}{\bar{\gamma}_M}$, $\tilde{b} = \frac{m_E}{\bar{\gamma}_E}$, $\mathcal{A}_M = \left(\frac{\sqrt{2}}{\lambda_M}\right)^{m_M-1} b^{\frac{m_M+1}{2}} e^{-\frac{\lambda_M^2}{2}}$ and $\mathcal{A}_E = \left(\frac{\sqrt{2}}{\lambda_E}\right)^{m_E-1} \tilde{b}^{\frac{m_E+1}{2}} e^{-\frac{\lambda_E^2}{2}}$. The previous integral with marcum-Q function in I_1 possess analytical intractability while evaluating the closed-form solution. Thus, using alternate expression of $F_E(\gamma_M)$ with lower incomplete Gamma function given in (4) and then employing [28, eq. (8.352.1)] and [28, eq. (8.445)], (7) can be written as

$$I_1 = \exp\left(-\frac{\lambda_M^2+\lambda_E^2}{2}\right) \sum_{l=0}^{\infty} \sum_{l_1=0}^{\infty} \frac{\left(\frac{\lambda_M}{\sqrt{(2)}}\right)^{2l_1} \left(\frac{\lambda_E}{\sqrt{(2)}}\right)^{2l} b^{m_M+l_1} (m_E+l-1)!}{\Gamma(l+1)\Gamma(l+m_E)\Gamma(l_1+1)\Gamma(m_M+l_1)} \left[I_1' - \sum_{l_2=0}^{m_E+l-1} \frac{\tilde{b}^{l_2}}{l_2!} I_1'' \right] \tag{9}$$

where $I_1' = \int_0^{\infty} \gamma_M^{l_1+m_M-1} \log_2(1+\gamma_M) e^{-b\gamma_M} d\gamma_M$ and $I_1'' = \int_0^{\infty} \gamma_M^{l_1+m_M+l_2-1} \log_2(1+\gamma_M) e^{-(b+\tilde{b})\gamma_M} d\gamma_M$. With the aid of [29, eq. (2.6.23.4)], the solution to I_1' and I_1'' is evaluated as

$$I_1' = \frac{\pi \log_2(e)}{(l_1+m_M) sin((m_M+l_1)\pi)} {}_1F_1\left(m_M+l_1; m_M+l_1-1; b\right) - \Gamma(m_M+l_1) b^{-m_M-l_1}$$
$$\times \left(\left[\ln(b) - \psi(m_M+l_1)\right] + \frac{b}{(1-m_M-l_1)} {}_2F_2\left(1,1; 2, 2-m_M-l_1; b\right) \right) \tag{10}$$

$$I_1'' = \frac{\pi \log_2(e)}{(\mathcal{T})sin(\mathcal{T}\pi)} {}_1F_1\left(\mathcal{T}; \mathcal{T}-1; b+\tilde{b}\right)$$
$$-\Gamma(\mathcal{T})\left(b+\tilde{b}\right)^{-\mathcal{T}} \left\{\left[\ln\left(b+\tilde{b}\right) - \psi(\mathcal{T})\right] - \frac{\left(b+\tilde{b}\right)}{(1-\mathcal{T})} {}_2F_2\left(1,1;2,2-\mathcal{T}; b+\tilde{b}\right)\right\} \tag{11}$$

where $\mathcal{T} = m_M + l_1 + l_2$, ${}_1F_1(.)$, ${}_2F_2(.)$, $\psi(.)$, and $\ln(.)$ are the confluent hypergeometric function and generalised hypergeometric function, psi function, and the natural logarithmic function. Similarly, analytical solution to I_2 can be written as

$$I_2 = \exp\left(-\frac{\lambda_M^2+\lambda_E^2}{2}\right) \sum_{l=0}^{\infty}\sum_{l_1=0}^{\infty} \frac{\left(\frac{\lambda_E}{\sqrt{(2)}}\right)^{2l_1}\left(\frac{\lambda_M}{\sqrt{(2)}}\right)^{2l} \tilde{b}^{m_E+l_1}(m_M+l-1)!}{\Gamma(l+1)\Gamma(l+m_M)\Gamma(l_1+1)\Gamma(m_E+l_1)} \left[\hat{I} - \sum_{l_2=0}^{m_M+l-1} \frac{b^{l_2}}{l_2!}\tilde{I}\right] \tag{12}$$

where the solution to $\hat{I}$ and $\tilde{I}$ is obtained from (10) and (11) by replacing the suffix $M \longleftrightarrow E$. The solution to integral (8) is deduced by employing [28, eq. (8.445)] and [29, eq. (2.6.23.4)]

$$I_3 = \exp\left(-\frac{\lambda_E^2}{2}\right) \sum_{l_1=0}^{\infty} \frac{\log_2(e)\left(\frac{\lambda_E}{\sqrt{(2)}}\right)^{2l_1}}{\Gamma(l_1+1)} \begin{cases} \dfrac{\pi\tilde{b}^{m_E+l_1} {}_1F_1\left(m_E+l_1; m_E+l_1-1; \tilde{b}\right)}{\Gamma(l_1+m_E+1)sin(\pi(m_E+l_1))} \\ -\left[\ln(\tilde{b}) - \psi(m_E+l_1)\right] \\ -\dfrac{\tilde{b}^{m_E+l_1+1} {}_2F_2\left(1,1;2,2-m_E-l_1;\tilde{b}\right)}{(1-m_E-l_1)} \end{cases} \tag{13}$$

The above solutions given by (10), (11), and (13) are valid for $2-m_{M/E}-l_1 > 0$ and $2-m_{M/E}-l_1-l_2 > 0$. Further at high SNR $\gamma \longrightarrow \infty$ $\log_2(1+\gamma) \approx \log_2(\gamma)$, thus (6) becomes

$$\bar{S}_c^{\rightarrow\infty} = \underbrace{\int_0^{\infty} \log_2(\gamma_M) f_M(\gamma_M) F_E(\gamma_M) d\gamma_M}_{I_1^{\rightarrow\infty}} + \underbrace{\int_0^{\infty} \log_2(\gamma_E) f_E(\gamma_E) F_M(\gamma_E) d\gamma_E}_{I_2^{\rightarrow\infty}}$$
$$\underbrace{-\int_0^{\infty} \log_2(\gamma_E) f_E(\gamma_E) d\gamma_E}_{I_3^{\rightarrow\infty}} \tag{14}$$

Now substituting (2) and (4) in the above equation and finally using [28, eq. (8.445)], [28, eq. (8.352.1)], and [29, eq. (2.6.21.2)], the solutions to $I_1^{\rightarrow\infty}$, $I_2^{\rightarrow\infty}$, and $I_3^{\rightarrow\infty}$ are deduced as follows

$$I_1^{\rightarrow\infty} = \exp\left(-\frac{\lambda_M^2+\lambda_E^2}{2}\right) \sum_{l=0}^{\infty}\sum_{l_1=0}^{\infty} \frac{\left(\frac{\lambda_M}{\sqrt{(2)}}\right)^{2l_1}\left(\frac{\lambda_E}{\sqrt{(2)}}\right)^{2l} b^{m_M+l_1}(m_E+l-1)!}{\Gamma(l+1)\Gamma(l+m_E)\Gamma(l_1+1)\Gamma(m_M+l_1)} \left[P_1 - \sum_{l_2=0}^{m_E+l-1} \frac{\tilde{b}^{l_2}}{l_2!}P_2\right] \tag{15}$$

Journal Pre-proofs

where $P_1 = b^{-m_M-l_1}\Gamma(m_M+l_1)\log_2(e)\left[\psi(m_M+l_1)-\ln(b)\right]$ and $P_2 = (b+\tilde{b})^{-\mathcal{T}}\Gamma(\mathcal{T})\log_2(e)\left[\psi(\mathcal{T})-\ln\left(b+\tilde{b}\right)\right]$. Similarly, solution to $I_2^{\rightarrow\infty}$ is obtained by doing the replacement of suffix as $M \longleftrightarrow E$ in (15).

$$I_3^{\rightarrow\infty} = \exp\left(-\frac{\lambda_E^2}{2}\right)\sum_{l_1=0}^{\infty}\frac{\log_2(e)\left(\frac{\lambda_E}{\sqrt{(2)}}\right)^{2l_1}}{\Gamma(l_1+1)}\left[\psi(m_E+l_1)-\ln\left(\tilde{b}\right)\right] \tag{16}$$

Further at low SNR $\gamma \to 0$, $\log_2(1+\gamma) \approx \gamma\log_2(e)$, thus utilising the previous definition in 6, it follows immediately that

$$\bar{S}_c^{\rightarrow 0} = \log_2(e)\left[\underbrace{\int_0^{\infty}\gamma_M f_M(\gamma_M)F_E(\gamma_M)d\gamma_M}_{I_1^{\rightarrow 0}} + \underbrace{\int_0^{\infty}\gamma_E f_E(\gamma_E)F_M(\gamma_E)d\gamma_E}_{I_2^{\rightarrow 0}} - \underbrace{\int_0^{\infty}\gamma_E f_E(\gamma_E)d\gamma_E}_{I_3^{\rightarrow 0}}\right] \tag{17}$$

Substituting (2) and (4) into $I_1^{\rightarrow 0}$, setting $\sqrt{\gamma} = t$, and finally utilising the identities [28, eq. (8.352.1)] and [30, eq. (2.15.5.4)], one obtains

$$I_1^{\rightarrow 0} = \mathcal{A}_M e^{-\frac{\lambda_E^2}{2}}\sum_{l=0}^{\infty}\frac{\left(\frac{\lambda_E}{\sqrt{2}}\right)^{2l}(m_E+l-1)!}{\Gamma(l+1)\Gamma(l+m_M)}\left[\bar{P}_1 - \sum_{l_2=0}^{m_E+l-1}\frac{\tilde{b}^{l_2}}{l_2!}\bar{P}_2\right] \tag{18}$$

where $c = \sqrt{2b}\lambda_M$, $\bar{P}_2 = \frac{c^{m_M-1}\Gamma(m_M+l_2+1)}{2^{m_M-1}(b+\tilde{b})^{m_M+l_2+1}\Gamma(m_M)}{}_1F_1\left(m_M+l_2+1; m_M; \frac{c^2}{4(b+\tilde{b})}\right)$, and $\bar{P}_1 = \frac{c^{m_M-1}\Gamma(m_M+1)}{2^{m_M-1}b^{m_M+1}\Gamma(m_M)}{}_1F_1\left(m_M+1; m_M; \frac{c^2}{4b}\right)$. The result for $I_2^{\rightarrow 0}$ is deduced by replacing the suffix $M \longleftrightarrow E$ in (18).

The closed-form solution to $I_3^{\rightarrow 0}$ is obtained by putting (2) into $I_3^{\rightarrow 0}$ and employing the identity [30, eq. (2.15.5.4)] as

$$I_3^{\rightarrow 0} = \mathcal{A}_E\frac{\hat{c}^{m_E-1}\Gamma(m_E+1)}{2^{m_E-1}\tilde{b}^{m_E+1}\Gamma(m_E)}{}_1F_1\left(m_E+1; m_E; \frac{\hat{c}^2}{4\tilde{b}}\right) \tag{19}$$

where $\hat{c} = \sqrt{2\tilde{b}}\lambda_E$.

3.2. Secrecy outage probability

The probability of outage occurs when instantaneous secrecy capacity falls below a target secrecy rate. Thus, it is defined as

$$P_{out}(R_s) = \mathcal{P}(C_s < R_s) \tag{20}$$

where the secrecy rate R_s is expressed in terms of threshold SNR γ_{th} as $R_s = \log_2(1+\gamma_{th})$. After performing mathematical manipulations, (20) can be defined as [1, eq. (27)]

$$\begin{aligned} P_{out}(\gamma_{th}) &= \mathcal{P}\left[\gamma_M \le (1+\gamma_{th})(1+\gamma_E) - 1\right] \\ &= \int_0^{\infty} f_E(\gamma_E) F_M(\gamma_{th} + \gamma_{th}\gamma_E + \gamma_E) d\gamma_E \end{aligned} \tag{21}$$

Plugging (2) and (4) into (21) and employing [28, eq. (8.352.1)], it immediately follows

$$P_{out}(\gamma_{th}) = \mathcal{B}\begin{cases} \int_0^{\infty} f_E(\gamma_E) d\gamma_E - e^{-b\gamma_{th}} \sum_{l_2=0}^{m_M+l-1} b^{l_2} \dfrac{\mathcal{A}_E}{l_2!} \\ \int_0^{\infty} \gamma_E^{\frac{m_E-1}{2}} (\gamma_{th} + \gamma_{th}\gamma_E + \gamma_E)^{l_2} e^{-\chi\gamma_E} I_{m_E-1}\left(\sqrt{2\tilde{b}\gamma_E}\lambda_E\right) d\gamma_E \end{cases} \tag{22}$$

where $\mathcal{B} = e^{-\frac{\lambda_M^2}{2}} \sum_{l=0}^{\infty} \dfrac{\left(\frac{\lambda_M}{\sqrt{2}}\right)^{2l} (m_M + l - 1)!}{\Gamma(l+1)\Gamma(l+m_M)}$ and $\chi = b(1+\gamma_{th}) + \tilde{b}$. Next, using the binomial expansion with the aid of [28, eq. (1.111)], setting $\gamma_E = t^2$ and [30, eq. (2.15.5.4)], one can deduce the analytical expression of SOP as

$$P_{out}(\gamma_{th}) = \mathcal{B}\begin{cases} 1 - e^{-b\gamma_{th}} \sum_{l_2=0}^{m_M+l-1} b^{l_2} \dfrac{\mathcal{A}_E}{l_2!} \sum_{k=0}^{l_2} \binom{l_2}{k} {\gamma_{th}}^{l_2-k} (1+\gamma_{th})^k \\ \dfrac{\hat{c}^{m_E-1}\Gamma(m_E+k)}{2^{m_E-1}\chi^{m_E+k}\Gamma(m_E)} {}_1F_1\left(m_E+k; m_E; \dfrac{\hat{c}^2}{4\chi}\right) \end{cases} \tag{23}$$

Further assuming that both destination and Eve nodes are super close to source such that received instantaneous SNR γ_M and γ_E are high, then usually the lower-bound of the SOP $SOP_L(\gamma_{th})$ is considered. The development for $SOP_L(\gamma_{th})$ is as follows

$$P_{out}(\gamma_{th}) \ge SOP_L(\gamma_{th}) = \int_0^{\infty} f_E(\gamma_E) F_M(\gamma_{th} + \gamma_{th}\gamma_E) d\gamma_E \tag{24}$$

Plugging (2) and (4) into (24), employing [28, eq. (8.352.1)], it immediately follows

$$SOP_L(\gamma_{th}) = \mathcal{B} - \mathcal{B}e^{-b\gamma_{th}} \sum_{l_2=0}^{m_M+l-1} b^{l_2} \frac{\mathcal{A}_E \gamma_{th}^{l_2}}{l_2!} \int_0^{\infty} \gamma_E^{\frac{m_E-1}{2}} (1+\gamma_E)^{l_2} e^{-\hat{b}\gamma_E} I_{m_E-1}\left(\sqrt{2\tilde{b}\gamma_E}\lambda_E\right) d\gamma_E \quad (25)$$

where $\hat{b} = b\gamma_{th} + \tilde{b}$. Similarly, following the identical steps as stated in (22)-(23), the lower-bound of SOP is eventually deduced as

$$SOP_L(\gamma_{th}) = \mathcal{B}\left\{1 - e^{b\gamma_{th}} \sum_{l_2=0}^{m_M+l-1} b^{l_2} \frac{\mathcal{A}_E \gamma_{th}^{l_2}}{l_2!} \sum_{k=0}^{l_2} \binom{l_2}{k} \frac{\hat{c}^{m_E-1}\Gamma(m_E+k)}{2^{m_E-1}\hat{b}^{m_E+k}\Gamma(m_E)} {}_1F_1\left(m_E+k; m_E; \frac{\hat{c}^2}{4\hat{b}}\right)\right. \quad (26)$$

The previous derived solution is valid only for integer values of m_M. Thus, more generalised solution $SOP_L(\gamma_{th})$ can be deduced by putting (2) and (4) into (21) and employing identities [28, eq. (8.445)] and [30, eq. (2.10.3.2)]

$$\begin{aligned} SOP_L(\gamma_{th}) = e^{-\left(\frac{\lambda_M^2}{2}+\frac{\lambda_E^2}{2}\right)} \sum_{l=0}^{\infty} \frac{\left(\frac{\lambda_M}{\sqrt{2}}\right)^{2l}}{\Gamma(l+1)\Gamma(m_M+l)} \sum_{l_1=0}^{\infty} \frac{\left(\frac{\lambda_E}{\sqrt{2}}\right)^{2l_1}}{\Gamma(l_1+1)\Gamma(m_E+l_1)} \\ \times \frac{\bar{\chi}^{m_M+l}\Gamma(m_M+m_E+l+l_1)}{(m_M+l)\tilde{b}^{m_M+l}} {}_2F_1\left(m_M+l, m_M+m_E+l+l_1; m_M+l+1; -\frac{\bar{\chi}}{\tilde{b}}\right) \end{aligned} \quad (27)$$

where $\bar{\chi} = b(1+\gamma_{th})$ and ${}_2F_1(.)$ is the Gauss hypergeometric function.

3.3. Strict positive secrecy capacity

Existence of the SPSC is guaranteed when $R_s = 0$, thus it can be defined using [1, eq. (19)] as

$$\begin{aligned} P_0 &= \mathcal{P}(C_s > 0) \\ &= 1 - \underbrace{\int_0^{\infty} F_M(\gamma_E) f_E(\gamma_E) d\gamma_E}_{I_7} \end{aligned} \quad (28)$$

Now substituting (2) and (4) into the integral part of (28), setting $\gamma_E = t^2$, and finally employing [28, eq. (8.352.1)] and [30, eq. (2.15.5.4)] results

$$I_7 = \mathcal{B}\left[1 - \mathcal{A}_E \sum_{k=0}^{m_M+l-1} \frac{b^k \hat{c}^{m_E-1}\Gamma(m_E+k)}{2^{m_E-1}k!(b+\tilde{b})^{m_E+k}\Gamma(m_E)} {}_1F_1\left(m_E+k; m_E; \frac{\hat{c}^2}{4(b+\tilde{b})}\right)\right] \quad (29)$$

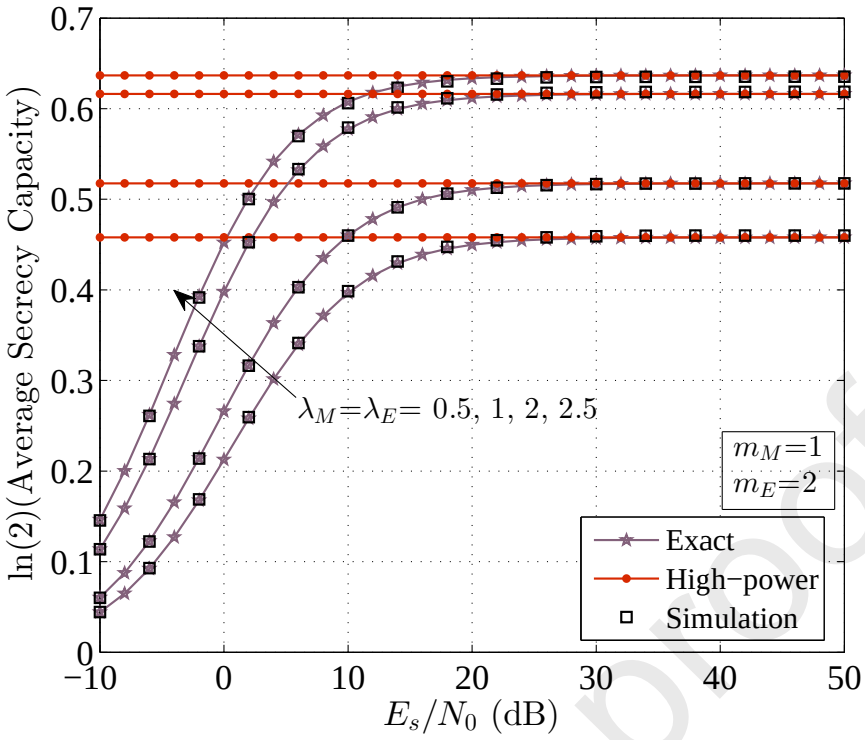


Figure 2: ASC as a function of E_s/N_0 for various values of λ_M and λ_E.

Plugging (29) into (28), one can easily deduce the analytical expression of the SPSC.

4. Numerical analysis

In the present section, the numerical results obtained in the preceding sections are plotted and discussed. Useful insights into the system performance has also been provided. Monte-Carlo simulations are used to verify the accuracy of the solutions provided here. In the Monte-Carlo simulations 10^6 random samples are generated, which are sufficient for the desired accuracy of the proposed methodology.

In Fig. 2, the ASC versus E_s/N_0 for different values of λ_M and λ_E is depicted. It can be seen that with the initial increase in E_s/N_0, the value of ASC increases and the plots get saturated with further increase in E_s/N_0. Early saturation in the value of ASC is attained for higher values of λ_M and λ_E. The high power approximation plots are also shown along with simulation plots. The analytical plots exactly match with their simulated counterpart. The variation of ASC with different values of m_M and m_E is shown in Fig. 3. The similar behaviour of ASC as discussed in Fig.1 is observed with the different fading parameters. It is noticed that as the channel severity increases the communication on the main channel becomes more secure. In 5G applications most of the devices operate in low-power mode. For this reason the low SNR ASC results are

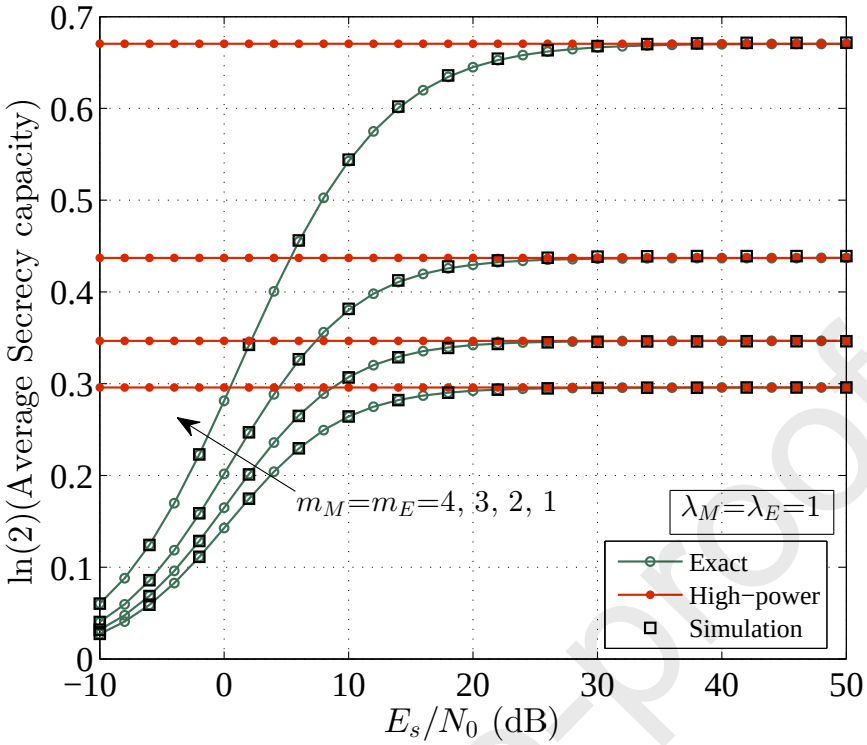


Figure 3: ASC as a function of E_s/N_0 for various values of m_M and m_E.

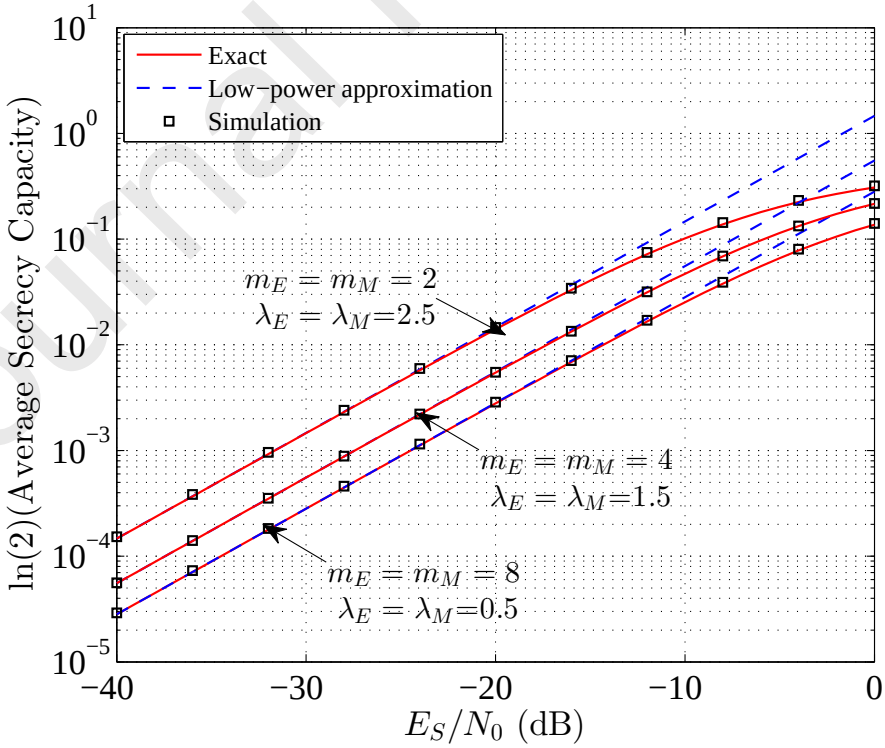


Figure 4: Low-power ASC as a function of E_s/N_0 with different fading parameters.

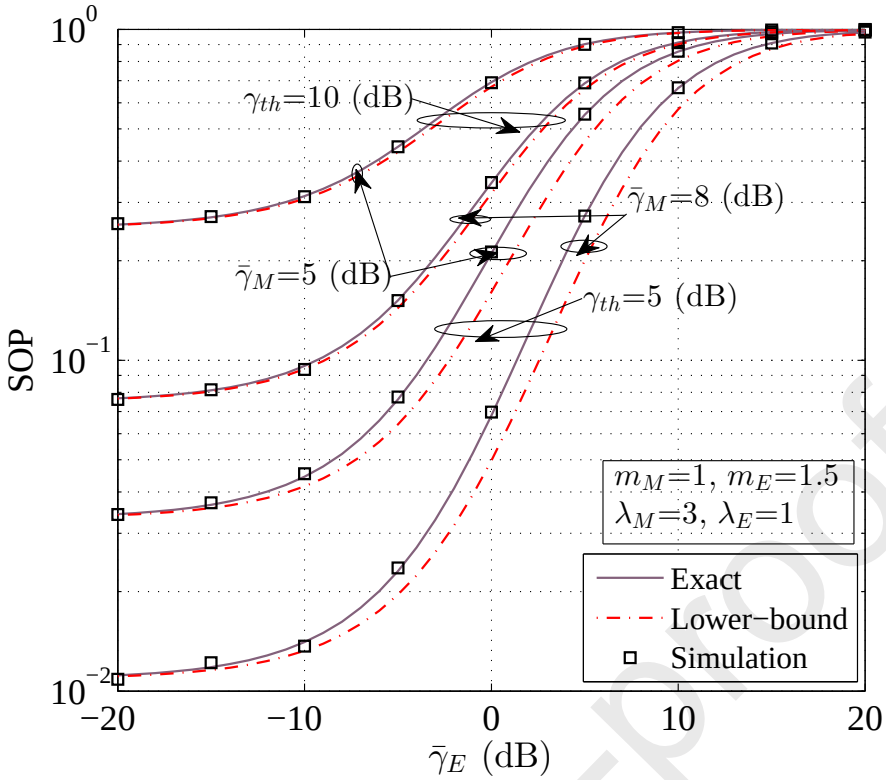


Figure 5: SOP and lower bound of SOP as a function of $\bar{\gamma}_E$

demonstrated in Fig. 4, where plots perfectly follow the simulation results at low SNR.

The SOP plots are shown in Fig. 5 and Fig. 6. As shown in Fig. 5, the plot shifts upwards with an increase in γ_{th} while it move downwards with increase in $\bar{\gamma}_M$. The lower bound of SOP plot exactly matched with the SOP plot for lower values of $\bar{\gamma}_E$ and the gap between the two increases with increase in $\bar{\gamma}_E$. It is clearly seen from the plots that the difference between the SOP and lower bound of SOP is decreasing with increase in γ_{th} while the opposite behaviour is observed for $\bar{\gamma}_M$. As shown in Fig. 6 the gaps between the SOP and its lower-bound are less for small values of $\bar{\gamma}_E$. The variation of SPSC for different values of system parameters are shown in Fig. 7. As expected, the SPSC decreases with the increase in severity fading parameter, m of the main or/and the wire-tap channels while it increases with the increase in the value of $\bar{\gamma}_M$. The simulation results almost match with the closed-form results, verify the accuracy of the proposed methodology. Finally, we remark that the results achieved in this section strongly support the concept of secure communication over the main channel in severe fading scenarios.

5. Conclusion

In this study, the impact of Beaulieu-Xie channel fading parameters on the physical layer security performance matrices has been investigated. The work derived the closed-form ex-

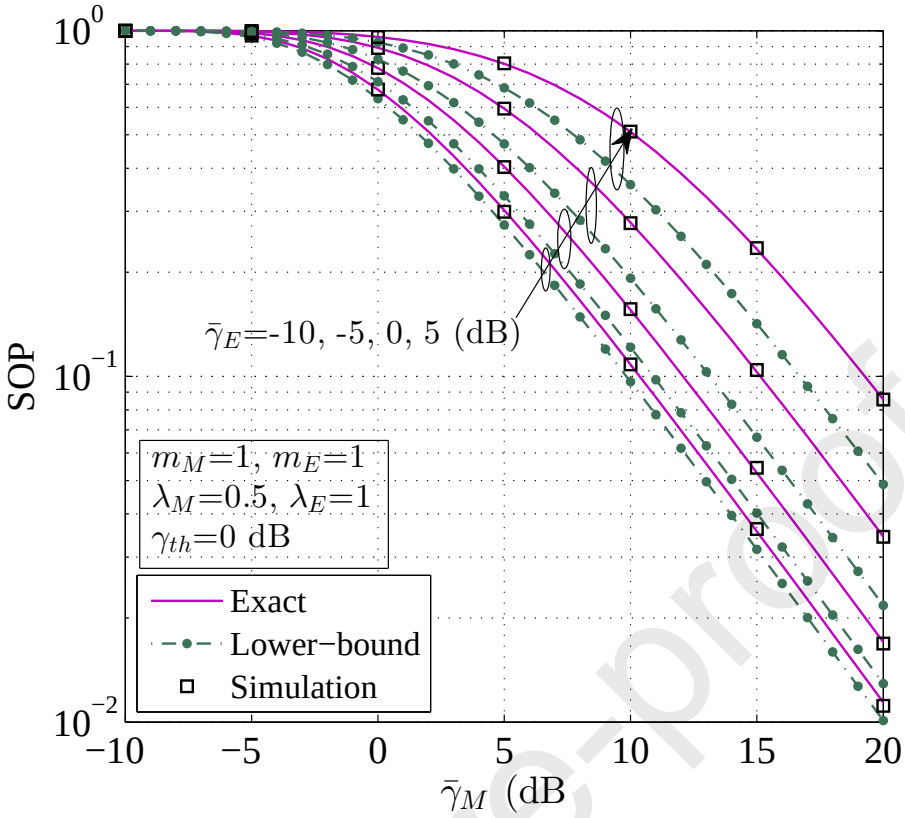


Figure 6: SOP and lower bound of SOP as a function of $\bar{\gamma}_M$

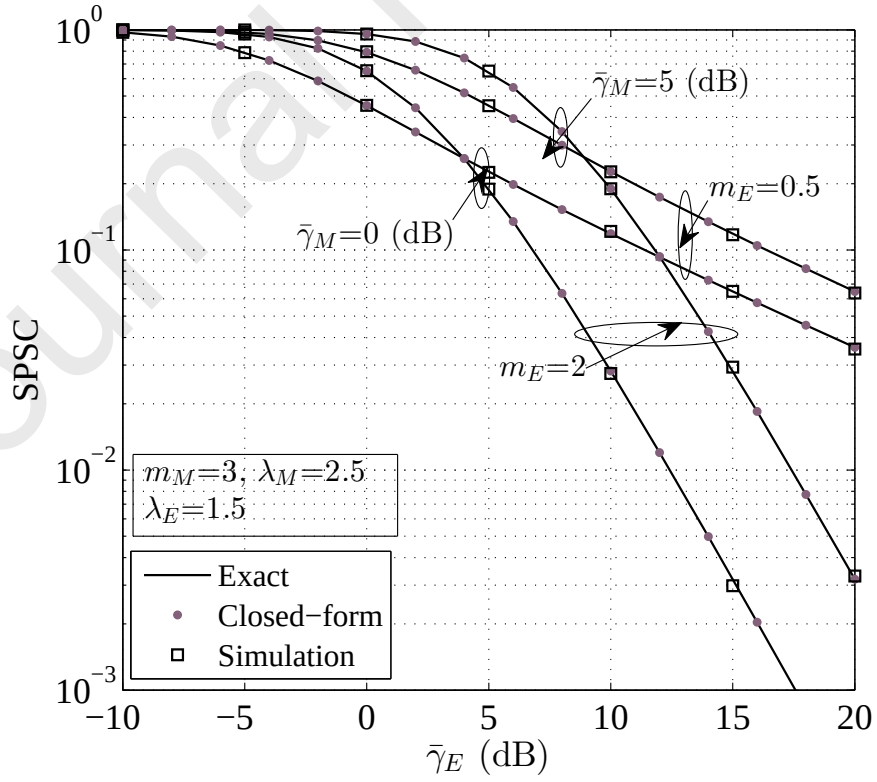


Figure 7: SPSC for various values of fading parameters

Journal Pre-proofs

pressions for ASC, SOP, lower-bound of SOP, and SPSC. To get more insight into the system, the high power analysis of the ASC has been provided. The low SNR ASC results were also produced for low power 5G network applications. Besides, the impact of different fading parameters on the system matrices a investigated. Finally, our theoretical results were verified via extensive Monte Carlo simulations. It has been observed that the expressions obtained in this work can be easily evaluated using the standard mathematical software packages such as MATLAB and MATHEMATICA..

References

[1] Bhargav N, Cotton SL, Simmons DE. Secrecy capacity analysis over $\kappa-\mu$ fading channels: Theory and Applications. IEEE Trans. on Commun., 2016; 64 (7): 3011-3024.

[2] Rivest RL., Shamir A., Adleman L. A method for obtaining digital signatures and public-key cryptosystems. Commun. of the ACM., 1978; 21 (2): 120-126.

[3] Olutayo OO, Stanley HM. Review of Channel Estimation for Wireless Communication Systems. IETE Technical Review, 2012; 29 (4): 282-298.

[4] Idrees NM, Haselmayr W, Petit M, Springer A. Improving time variant channel estimation for 3GPP LTE-downlink. in proc IEEE 23rd Int Symp on Pers, Indoor and Mobile Radio Commun-(PIMRC), 2114-2119: Sydney, NSW, 2012.

[5] Li Y, Wang Y, Jiang T. Sparse-aware set-membership NLMS algorithms and their application for sparse channel estimation and echo-cancellation. Int J of Elec and Commun, 2016; 70 (7): 895-902.

[6] Li Y, Jiyang Z, Omer OOO, Han X, Jingwei Y. Mixed Norm Constrained Sparse APA Algorithm for Satellite and Network Echo Channel Estimation. IEEE Access, 2018; 6: 65901-65908.

[7] Maurer UM. Secret key agreement by public discussion from common information. IEEE Trans. on Inf. Theory., 1993; 39 (3): 733-742.

[8] Bloch M., Barros J., Rodrigues MR, McLaughlin SW. Wireless information-theoretic security. IEEE Trans. on Inf. Theory., 2008; 54 (6): 2515-2534.

[9] Wyner AD. The wire-tap channel. Bell Syst. Tech. J., 1975; 54 (8): 1355-1387.

[10] Kumar R, Chauhan SS. Secrecy analysis of MRT/RAS system under Nakagami-*m* fading channels in the presence of imperfect channel state information. Int. J. Elec. Commun., 2018; 85: 68-73.

[11] Xu L, Yu X, Wang H *et al.* Physical layer security performance of mobile vehicular networks. Mobile Networks and App, 2019; 1-7. https://doi.org/10.1007/s11036-019-01224-8

[12] Kong L, Tran H, Kaddoum G. Performance analysis of physical layer security over $\alpha - \mu$ fading channel. Electronics Lett., 2016; 52, (1): 45-47.

[13] Mathur A, Ai Y, Bhatnagar MR, Cheffena M, Ohtsuki T. On Physical Layer Security of $\alpha - \eta - \kappa - \mu$ Fading Channels. IEEE Commun. Lett., 2018; 22 (10): 2018.

[14] Kong L, Kaddoum G, Chergui H. On physical layer security over Fox's H-function wiretap fading channels. IEEE Trans on Veh Technol, 2019; 68 (7): 6608-6621.

[15] Kong L, Kaddoum G. Secrecy characteristics with assistance of mixture Gamma distribution. IEEE Wireless Commun Lett, 2019; 8(4): 1086-1089.

[16] Lei H, Ansari IS, Gao C, Guo Y, Pan G, *et al.* Physical-layer security over generalised-K fading channels. IET Commun., 2016; 10 (16): 2233-2237.

[17] Lei H, Zhang H, Ansari IS, Gao C, Guo Y, *et al.* Performance analysis of physical layer security over Generalized-*K* fading channels using a mixture Gamma distribution. IEEE Commun. Lett., 2016: 20 (2): 408-411.

[18] Lei H, Ansari IS, Zhang H. Security performance analysis of SIMO Generalized-*K* fading channels using a mixture Gamma distribution. Veh. Technol. Conf., Montreal, QC, Canada, 2016; 1-6.

[19] Alexandropoulos GC, Peppas KP. Secrecy Outage Analysis over Correlated Composite Nakagami-*m*/Gamma Fading Channels. IEEE Commun. Lett., 2018; 22 (1): 77-80.

[20] Sun J, Li X, Huang M, Ding Y, Jin J, *et al.* Performance analysis of physical layer security over $\kappa - \mu$ shadowed fading channels. IET Commun., 2018; 12 (8): 970-975.

[21] Srinivasan M, Kalyani S. Secrecy capacity of $\kappa - \mu$-shadowed fading channels. IEEE Commun. Lett., 2018; 22 (8): 1728-1731.

[22] Beaulieu NC, and Jiandong X. A Novel Fading Model for Channels with Multiple Dominant Specular Components. IEEE Wireless Commun. Lett., 2015; 4 (1): 54-57.

[23] Olutayo A, Ma H, Cheng J. and Holzman JF. Level Crossing Rate and Average Fade Duration for the Beaulieu-Xie Fading Model. IEEE Wireless Commun. Lett., 2017, 6 (3):326-329.

[24] Olutayo A, ChengJ, and Holzman J. Asymptotically tight performance bounds for equal-gain combining over a new correlated fading channel. 15th Canadian Workshop on Information Theory (CWIT), 2017, Quebec City: 1-5.

[25] Olutayo A, Cheng J, and Holzman J. Asymptotically tight performance bounds for selection diversity over Beaulieu-Xie fading channels with arbitrary correlation. IEEE International Conference on Communications (ICC), 2017, Paris: 1-6.

[26] Singh D. and Joshi HD. Generalized MGF based analysis of line-of-sight plus scatter fading model and its applications to mimo-ofdm systems. Int. J. Elec. Commun., 2018, 91, 110-117.

[27] Chauhan PS, Rana V, Kumar S, Soni SK, Pant D. Performance analysis of wireless communication system over nonidentical cascaded generalised gamma fading channels. Int J Commun Syst., 2019; 1-15.

[28] Gradshteyn I, Ryzhik I. Table of Integrals, Series, and Products (7th edn). Academic Press; 2007.

[29] Prudnikov AP, Brychkov YA, Marichev OI, Integral Series, vol. 1, Elementary functions, The Netherlands: Gordon and Breach, 1986.

[30] Prudnikov AP, Brychkov YA, Marichev OI. Integrals and Series Volume 2: Special Functions, 1st edn., Gordon and Breach Science Publishers, 1986.

AUTHORS' BIOGRAPHIES

Puspraj Singh Chauhan was born in UP, India. He received his Bachelor of

Technology in Electronics and Communication Engineering from Uttar Pradesh Technical University, India in 2009, the M.Tech degree in Digital Signal Processing from Govind Ballabh Pant Engineering College, Pauri, Uttarakhand, India in 2011. Currently he is working towards the PhD. degree in Electronics and Communication Engineering from Govind Ballabh Pant Engineering College, Pauri, Uttarakhand, India.. His research interests include Wireless Communication and Propagation Channel Modeling.

Sandeep Kumar received his B. Tech. in electronics and communication from Kurukshetra University, India in 2004 and Master of Engineering in Electronics and Communication from Thapar University, Patiala, India in 2007. He received his Ph.D. from Delhi Technological University, Delhi, India in 2018. He is currently working as Member (Senior Research Staff) at Central Research Laboratory, Bharat Electronics Limited Ghaziabad, India. His research interests include the study of wireless channels, performance modeling of fading channels and cognitive radio networks. He is also serving as a reviewer for IEEE, Elsevier and Springer journals.

Sanjay Kumar Soni was born in UP, India on March, 1975. He received his Bachelor of Engineering in Electronics Engineering from Madan Mohan Malviya Engineering College, Gorakhpur, Uttar Pradesh, India in 1997, the M.Tech degree in Communication Engineering from IIT Kanpur in 2004, and PhD in wireless Communication Engineering from IIT Kharagpur, India in 2011. Currently he is working as Associate Professor in Electronics and Communication department, G.B.Pant Engineering College, Pauri. His research interests include Wireless Communication, Propagation Channel Modeling, Cognitive Radio, MIMO-OFDM, UWB Communication, Statistical Characterization of Wireless Channel.

Journal Pre-proofs

Performance and Emission Studies of a Variable Compression Ratio CI Engine using Bio Diesel Made from Waste Cooking Oil

Jatinder Kataria[1*], SK Mohapatra[1], K Kundu[2], R Singh[3]
[1]Department of Mechanical Engineering, Thapar University, Patiala-147004, Punjab, India
[2]Department of Bio fuel, Centre of Excellence for Farm Machinery, Ludhiana-141006, Punjab, India
[3]Department of Mechanical Engineering, Delhi Technological University, Delhi, India

Article Info
Article history:
Received 1 January 2019
Received in revised form
20 July 2019
Accepted 28 May 2019
Available online 15 September 2019

Keywords: Bio diesel, Cooking oil, Alternative fuel, CI Engine

Abstract
Compression ignition (CI) engine is the un-debated choice for power applications, stationary or mobile. There is an urgent need of alternative high potential fuel for CI engines in order to fulfill energy needs without hampering the thermal performance and stringent emission standards. Here waste cooking oil is chosen as an alternative fuel, which is upgraded into biodiesel in the laboratory using mechanical stirring and ultrasonic cavitation technique of biodiesel production.
Four blends of cooking oil are made by increasing 20% its amount in conventional diesel fuel. Performance is analyzed for two compression ratios, 15 and 17.5, on the existing VCR diesel engine set up interfaced with "Engine Soft" software. The results show that as bio-diesel concentration in a blend increases, the thermal performance (i.e., brake thermal efficiency) and emissions (i.e., CO, HC and NOx, and smoke intensity) are observed to be marginally higher; on the other hand, as compression ratio increases, the thermal performance improves; CO and smoke opacity decreases.

1.Introduction

The scarcity of fast diminishing resources of fossil fuels, increasing prices of crude oil, and environmental concerns have been the various reasons for exploring the use of vegetable oils as alternative to diesel oil [1-4]. Vegetable oils offer almost the same output with slightly lower thermal efficiency when used in diesel engines [5]. Reduction of engine emissions is a major research aspect in engine development with the increasing concern over environmental protection and the stringent exhaust gas regulation [6]. The use of neat vegetable oils poses some problems such as coking and trumpet formation on the injectors, carbon deposits, oil ring sticking and thickening and gelling of lubricating oil, when subjected to prolonged usage in CI engines. These problems are attributed to high viscosity, low volatility and polyunsaturated character of vegetable oils [7]. Some of the common problems posed by using vegetable oils in diesel engines are as a result of contamination by the vegetable oils Different methods such as preheating, blending and trans stratification are being used to reduce the viscosity and to produce bio-diesel, suitable for engine applications. In the present investigation, bio-diesel is prepared from waste cooking oil. The performance and emission characteristics were studied on a four-stroke, single cylinder, variable compression ratio direct-injection diesel engine to ensure its suitability as CI engine fuel.

2. Experimental System Development for Bio Diesel Production

2.1 Methodology

Huge quantities of waste cooking oils and animal fats are available throughout the world, especially in the developed countries. Management of such oils and fats pose a significant challenge because of disposal problems and possible contamination of the water and land resources. Even though some of this waste cooking oil is used for soap production, a major part of it is discharged into the environment. The use of waste cooking oil as a biodiesel source has a potential to reduce CO2, particulate matter and other greenhouse gases as the carbon contained in biomass-derived fuel is largely biogenic and renewable.
Waste cooking oil, which is otherwise wasted, is one of the most economical choices to produce biodiesel. It is recognized that the production of waste cooking oil will be the function of the frying temperature and length of use as well as the material used for frying. In this experiment, waste cooking oil was collected from an hotelier industry.

Corresponding Author,
E-mail address: jatinder80@gmail.com
All rights reserved: http://www.ijari.org

2.2 Magnetic Stirring Technique

In this process mixing of oil and alcohol is done by a magnetic stirrer. A magnetic capsule is dipped in the immiscible liquid (oil and alcohol are not miscible with each other), as the capsule starts rotating, a whirl is formed which disturb the phase boundary between two immiscible liquid and resulted in emulsification of mixture. Magnetic stirrer used is as shown in figure 1.

Fig. 1: Mechanical Stirrer

2.3 Ultrasonic Cavitation Technique

Methanol is immiscible with the cooking oil. The mixture of oil, methanol and catalyst is poured in ultrasonic processor transducer (model TU-50). Test is performed for the ultrasonic frequency (28.5 KHz). During the reaction the temperature of mixture is kept between 45-60^{0}C When reaction is completed the beaker is kept for the separation. Fatty acid has higher specific weight therefore it will settle at bottom. Separation of methyl ester and glycerol will take 2 to 3 hr duration. After complete separation bio-diesel (methyl Ester) is visible in the upper layer and glycerol at the bottom as shown in figure 2.. Bio-diesel is separated from beaker for purification process. The catalyst present in the methyl ester is impurity. Excess methanol present in biodiesel has been removed by vaporization process.

3. Experimental Setup

The setup consists of single cylinder, four stroke, VCR (Variable Compression Ratio) Diesel engine connected to eddy current type dynamometer for loading. The compression ratio can be changed without stopping the engine and without altering the combustion chamber geometry by specially designed tilting cylinder block arrangement. Provision is also made for interfacing airflow, fuel

flow, temperatures and load measurement. The setup has stand-alone panel box consisting of air box, two fuel tanks for duel fuel test,manometer, fuel measuring unit, transmitters for air and fuel flow measurements, process indicator and engine indicator. Rota meters are provided for cooling water and calorimeter water flow measurement. The setup enables study of VCR engine performance for brake power, indicated power, frictional power, BMEP, IMEP, brake thermal efficiency, indicated thermal efficiency, Mechanical efficiency, volumetric efficiency, specific fuel consumption, A/F ratio and heat balance. The blends B20, B-40, B-60, B-80 and B100 were prepared in the laboratory by mixing appropriate volume of diesel with waste cooking oil.

Fig. 2: Biodiesel from WCO

A set of reading was obtained first by running the engine with diesel at CR of 17.5 and varying the load from idle to rated load of 3.5 KW in steps of 1 up to 3 KW and then to 3.5 KW. The engine performance characteristics were recorded by using the software Engine Soft and instrumentation provided by the National Instruments. The emissions were recorded for each load by using Gas Analyzer (Model: AVL DI Gas 444) and the opacity was recorded by Smoke meter (Model: AVL 437).

The engine was started at compression ratio of 17.5 and then the compression ratio was changed by using the tilting head arrangement. The compression ratio was indicated by the rings made on the lever of the arrangement. The VCR engine test rig is shown in figure 3.

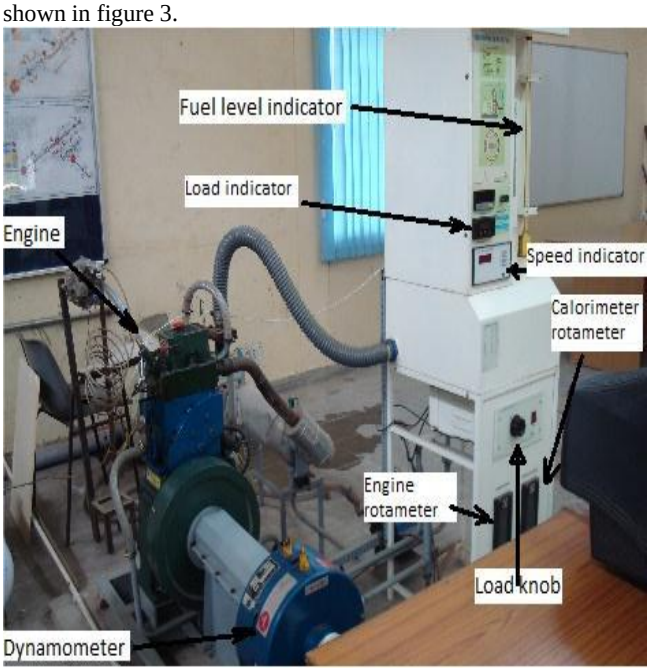


Fig. 3: VCR Engine Test Rig

4. Results and Discussion

4.1 Engine Performance

4.1.1 Brake Power

The effect of variation of brake power output at two different compression ratios has been studied for thermal performance (in terms of brake thermal efficiency, specific fuel consumption, mechanical efficiency volumetric efficiency) and emissions (in terms of carbon-monoxide, nitrogen oxides, hydrocarbons emissions and smoke opacity.

4.1.2 Brake Specific Fuel Consumption

Brake specific fuel consumption (bsfc) as a function of load obtained during engine operation on bio-diesel blends and diesel fuel at compression ratio of 17.5:1 is shown in Figure 4. For all fuels tested, bsfc decreased with increase in load at both compression ratios. At the compression ratio 15: 1, the bsfc of 0.29 kg/kWh was obtained for the Diesel fuel and at compression ratio 17.5:1 the bsfc of 0.27kg/kW/h. The data showed that at higher compression ratio there is slightly decrease in brake specific fuel consumption., the fuel blends B20 maintain about the same bsfc as that of diesel fuel, The brake specific fuel consumption at compression ratio 17.5:1 at B40 and brake specific fuel consumption at compression ratio 15:1 at B20 are same i.e 0.31kg/Kw/h. The reason for the above trend is more likely due to lower calorific values of bio-diesel blends compared to those of neat diesel [8]. So in order to produce equivalent power to that of diesel more bio-diesel blended fuel is required. The reason for increase in bsfc for all loads is the lower calorific value of bio-diesel blends than pure diesel.

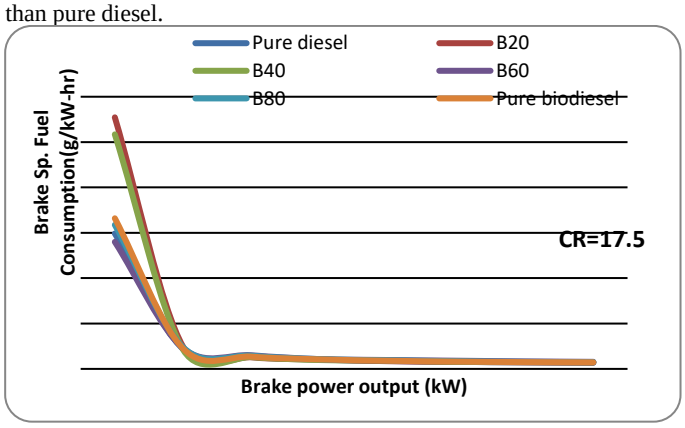


Figure 4: bsf vs. BP at CR 17.5:1

4.1.3 Brake Thermal Efficiency

The brake thermal efficiency increases with an increase in load at different compression ratios and at various biodiesel blends. This can be attributed to reduction in heat loss and increase in power with increase in load. The maximum brake thermal efficiencies obtained is about 31.18% for B40 at 17.5:1 and at 15:1 the brake thermal efficiencies is about 29.77% as shown in figure 5.

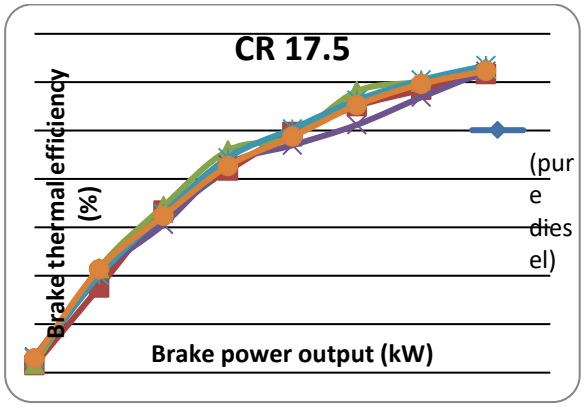


Figure 5 Break thermal efficiency at CR 17.5:1

This shows that at lower compression ratio there is decrese of 1.41% for the B40 biodiesel blends. The mixing of bio-diesel in diesel oil yields good thermal efficiency curves. Initially the thermal efficiency of the engine is improved by increasing concentration of the bio-diesel in the blend. The possible reason for this is the additional lubricity provided by bio-diesel. The molecules of bio-diesel (i.e. methyl ester of the oil) contain some amount of oxygen,

which takes part in the combustion process. It is noticed that after a certain limit with respect to diesel ester blend, the thermal efficiency trend is reversed and it starts decreasing as a function of the concentration of blend.

4.2 Emission Parameters

4.2.1 Carbon Monoxide

Figure 6 shows the plots of carbon monoxide (CO) emissions of blends and diesel fuel at different load conditions and compression ratio of 17.5:1.

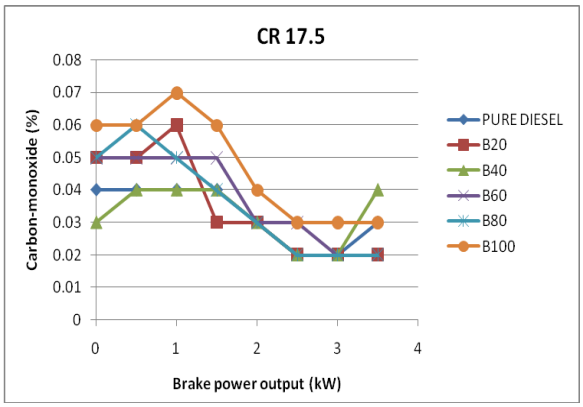


Figure 6 CO emissions at CR 17.5:1

This is typical with all internal combustion engines since air/fuel (A/F) ratio decreases with increased load. Also the data showed increase in CO emissions as the load was increased up to 50% when running on bio-diesel blends.

4.2.2 Un- burnt Hydrocarbon

The emission of un-burnt hydrocarbons (HC) for all fuels and compression ratios is small, 25-122 ppm, increasing slightly with load and proportion of fuel injected. It is quite difficult to determine any reliable dependencies; the waste cooking oil operation showed very low HC emissions throughout the load range. The plot indicates that HC emissions peaked at a particular A/F ratio and then dropped with further increase in load. The plot showed increase in HC emissions as the quantity of diesel fuel in the blend increases. A very significant difference was noted at low and high load conditions at the compression ratio 17.5:1 as shown in figure 7.

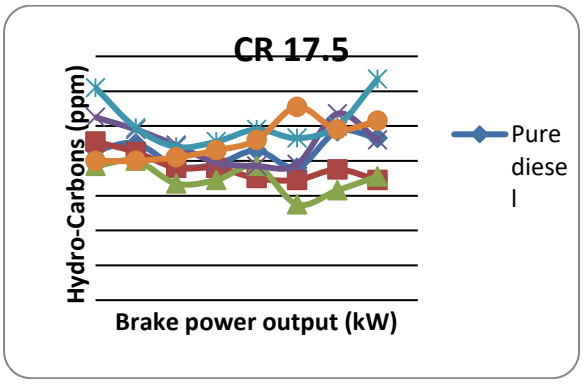


Fig. 7. HC emissions at CR 17.5:1

The increase in diesel/WCO blend showed highest HC concentration in the exhaust when compared with the results of increase in WCO/diesel blend. However at light load, the biodiesel blends did not show any marked difference in HC emissions irrespective of compression ratios. For minimum load the reduction was practically unaffected by the addition of bio-diesel in the diesel fuel at all Compression ratios. The most beneficial reduction appeared at intermediate loads. Hydro carbon emissions as a function of engine load at various compression ratios.

4.2.3 Oxides of Nitrogen

The emission of nitric oxide (NO) increases gradually with load, (figure 8) reaching the maximum values of 1578 ppm for blend80 at compression ratio 17.5:1 and its value decreases to 1478 ppm for blend 100 at the same compression ratio When operating at the compression ratio of 15:1, the minimum NO emissions at adequate loads were maintained by the B20 blend. Biofuels with higher oxygen contents, 1.45% to 4.075%, and neat bio-diesel produce higher or a bit lower NO emissions, depending on the load.

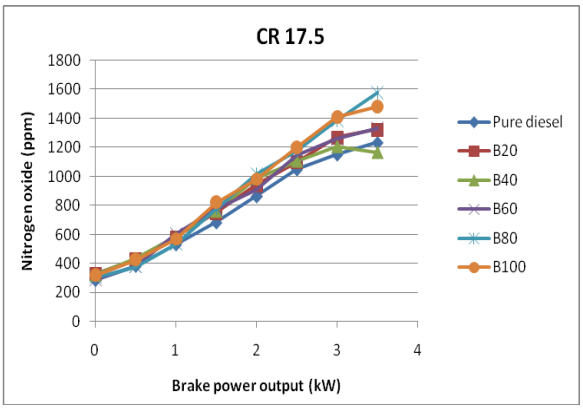


Figure 8. NO emissions at CR 17.5:1

The reason for the decrease in NO was that the cetane number of the bio-diesels was higher than that for diesel fuel and this is associated with lower NO emissions [9]. Increasing cetane number reduces the size of the premixed combustion by reducing the ignition delay. This results in lower NO formation rates since the combustion pressure rises more slowly, giving more time for cooling through heat transfer and dilution and leading to localized gas temperatures [9].Furthermore the emission of nitrogen oxide from engine exhaust are highly dependent on oxygen concentration and thus the combustion temperature.

4.2.4 Smoke Opacity

The smoke (soot) opacity with brake mean effective pressure of exhaust gas for neat diesel fuel and bio-diesel blends is shown in figure 9.

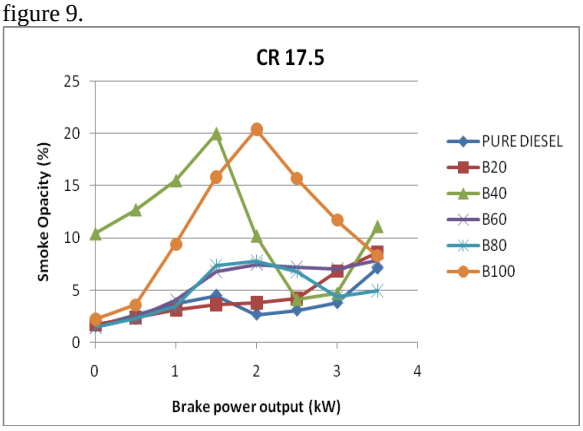


Figure 9 Smoke Opacity at CR 17.5:1

One can observe that soot emitted by all bio-diesel blends is lower than neat diesel at low loads and lower compression ratio. This is attributed to the combustion being mixed controlled for these blends, as is also the case for neat diesel, which is however assisted by the presence of the fuel bound oxygen [10]. However, there is no definite trend observed in smoke density with increase in blend percentage of bio-diesel. At all compression ratios, B100 was found to emit maximum smoke at full load. The minimum and maximum smoke opacities produced for B60 and B100.The smoke emissions generally increase or decrease in relation to the sulphur concentration. Sulphur in the fuel, results in sulphates that are absorbed on soot particles and increase the smoke emitted from diesel engines. In addition, the increase of oxygen content in the

fuel contributes to a complete fuel oxidation even in locally rich zones, leading to a significant decrease of smoke [9-10].

5. Conclusions

Four blends of cooking oil are made by increasing 20% its amount in conventional diesel fuel. Performance is analyzed for two compression ratios, 15 and 17.5, on the existing VCR diesel engine set up interfaced with "Engine Soft" software. The results show that as bio-diesel concentration in a blend increases, the thermal performance (i.e., brake thermal efficiency) and emissions (i.e., CO, HC and NOx, and smoke intensity) are observed to be marginally higher; on the other hand, as compression ratio increases, the thermal performance improves; CO and smoke opacity decreases.

References

[1]. AK Agarwal. Biofuels (alcohols and bio-diesel) applications as fuels for internal combustion engine. Journal of Progress in Energy Combustion Science. 33, 2006, 233-271.
[2]. AS Ramadhas, S Jayaraj, C Muraleedharan. Use of vegetable oils as I.C. engine fuels- A review. Journal of Renewable Energy. 29, 2004, 727-742.
[3]. M Muralidharan, MP Thariyan, S Roy, JP Subrahmanyam, PMV Subbarao. Use of pongamia bio-diesel in CI engines for rural application. SAE Paper No. 28-0030, 2004.
[4]. Vellguth G. 1983. Performance of vegetable oils and their monoesters as fuels for diesel engines. SAE Paper No. 831358.
[5]. D Agarwal, AK Agarwal. Performance and emission characteristics of a Jatropha oil (preheated and blends) in a direct injection compression ignition engine. Journal of Applied Thermal Engineering. 27, 2007, 2314-2323.
[6]. NJ Barsic, AC Humke. Performance and emission characteristics of a naturally aspirated diesel engine with vegetable oil fuels. SAE Paper No. 810262, 1981.
[7]. A Moneym, JH Gerpen. The effect of bio-diesel production on engine performance and emissions. Journal of Biomass and Bio energy. 20, 2001, 317-325.
[8]. PR Lee, CH Hobbs. Fuel quality impact on heavy duty diesel emissions- A literature review. SAE Paper No. 982649, 1998.
[9]. Y Akasaka, T Suzuki, Y Sakurai. Exhaust emissions of a DI diesel engine fueled with blends of bio-diesel and low Sulpher diesel fuel. SAE Paper No. 972998, 1997.
[10]. K Owen, T Coley. Automotive fuels reference book. Society of Automotive Engineers, USA, 1995.

OPEN ACCESS

International
Journal
Of Modern Engineering Research (IJMER)

Preparation of Low Cost Magnetorheological Fluid and its Testing

R. C. Singh [1], Ranganath M. S.[2] , Akshay Bhagat[3], Kaushik K.[4], Vaibhav Batra[5], Shiven Mittal[6], Gaurav Jain[7]

[1]Professor, [2]Associate Professor, [3,4,5,6,7]Student
Department of Mechanical and Production Engineering
Delhi Technological University, Delhi, India

ABSTRACT : *A magnetorheological fluid is a type of smart fluid which when subjected to a magnetic field, its viscosity tends to increase. The overall purpose of the research carried out in this paper is to develop low cost MR fluids without affecting its important characteristics such as sedimentary stability and a suitable increase in viscosity with increase in magnetic field. Two MR fluids were prepared and compared on the basis of their performance in the sedimentary stability test. Also, an experiment was conducted to see how the viscosity of the fluid reacted to change in the magnetic field. With the help of these two experiments, it is shown that one of the fluids can be used for practical purposes.*

Keywords: *Magnetorheological, Magnetoviscometer, Sedimentation, Surfactant, Viscosity*

I. INTRODUCTION

The advent of microprocessors and electronic technology has allowed us to incorporate electronic control into a lot of machines and materials and hence the demand for "smart materials" like Electrorheological and Magnetorheological fluids in applications like Automotive and Biomechanical components is increasing rapidly [1]. However, the magnetorheological fluids available in the market currently do not prove to be financially viable in most of their application and hence the need to develop low cost Magnetorheological fluid. Hence, development of a low cost Magnetorheological fluid is the need of the hour. A lot of work has been done on the use of mineral oils as the carrier fluid [2] in MR Fluids but the work wasn't focused towards reducing the cost of the fluid.

The aim of the organization of this paper is to develop a low cost magnetorheological fluid using carrier fluids like soybean oil and heavy paraffin oil along with Electrolytic Iron particles and suitable surfactants and then conducting tests to determine their suitability for practical purposes.

II. LITERATURE REVIEW

Horia Chiriac and George Stoian [3] have investigated the influence of two aspects which characterize magnetic powders on the magneto rheological (MR) fluids: particles size and particles size distribution. Fe powders were obtained with micrometric particles of different sizes and narrow size distribution by sieving commercially available Fe powder by means of a sieve shaker. Also, three powders with particles having normal size distributions with the same mean size and different variances were tailored using the sieved Fe powders. They prepared MR fluids based on the resulting magnetic powders and studied them by means of a rheometer with a magneto rheological module. Further, a correlation between the MR performance and particles size/particles size distribution was observed. MR fluids in 10% volumic concentrations of magnetic particles in the fluid were prepared using these powders. The MR fluid's behaviour in magnetic field was studied by means of a rheometer provided with a magneto rheological cell. For each fluid, yield stress curves were traced having the magnetic field as parameter, and based on these curves, the dependence of the yield stress as function of applied magnetic

field was plotted. The data indicated that the particles size and also the particles size distribution have an influence on the MR fluid performance.

They found out that greater stress responds to bigger particles. Therefore, there is a direct correlation between yield stress and particle size. But, larger particles lead to a more rapid sedimentation of the solid particles in the fluid, but this problem can be overcome in a large extent by using a suitable formulation of the MR fluid. Also it seems that powder comprising particles with sizes following a Gaussian distribution, it is preferable to have the particles within a narrower size distribution. It was suggested that for future work, to study some magnetorheological fluids with bi-disperse particle mixes using, for example, two powders with particle size situated at the extremities of the normal distribution T.M. Simon et al. [4] have studied the effective magnetic behaviour of an MR composite as a function of the inter-particle distance. They have presented and employed a mathematical model for the effective magnetic properties of MR fluids with periodic microstructure that is based on the theory of homogenization. They have also discussed an interpolating formula for the effective permeability of MR fluids. They have further presented an experimental setup for calculating the permeability of an MR fluid sample from the experimental B-H curve. Experiments were conducted (at the Lord Corporation) to measure the magnetic properties of three MR fluids with particle loadings of l0%, 20%, and 30% by volume. This relation showed good agreement with experimental data Ji Eun Kim et al. [5] measured the MR characteristics of the CI dispersed in PIB/PB by both static and dynamic tests using a rotational rheometer with and without an external magnetic field. Static and dynamic yield stress, shear stress, and storage modulus from the rheological measurements were obtained to be increased with applied magnetic field strengths. The system also showed improved sedimentation stability due to the elastic property of PIB/PB viscoelastic fluid. The CI-PIB/PB based MR viscoelastic fluid was found to overcome the sedimentation problem, demonstrating that the system which consists with a Boger fluid has a potential to be used as an MR fluid without sedimentation.

III. PREPARATION OF FIRST SAMPLE

Magnetorheological fluids consist of three main components namely carrier fluid, additives and ferromagnetic particles. We prepared two samples of the magnetorheological fluid by selecting two different materials for each of the three components.

Magnetic Particles

We did a thorough market research and searched for different types of metal particles, carrier fluids and additives. After completing the survey of the market, we found that sourcing Carbonyl Iron particle which is solely manufactured by BASF, will not justify our aim of manufacturing cost effective MR Fluid. So we went with the best alternative particle that is Electrolytic Iron.

Table II: ***Properties of Magnetic Material***

Carbonyl Iron		Electrolytic Iron	
Iron as (Fe)	99.5%	Iron as (Fe)	> 95%
Apparent Density “Carney”	1.1-1.3g/cc	Apparent Density “Carney”	2.5-2.7g/cc
Particle Size	3-10 µ	Particle Size	20-45 µ

Carbonyl Iron powder has a spherical particles with size between 1-8 microns and has excellent compaction and sintering properties. Moreover it has a high chemical purity (99.5%) and an excellent electromagnetic performance. Electrolytic Iron powder has roughly spherical particle with slightly larger size of 20-45 microns and has a high purity of about (>95%). It also shows decent electromagnetic properties.
Therefore based on the above mentioned properties and due to its easy availability we chose Electrolytic Iron powder. Price of electrolytic iron is $1/10^{th}$ that of carbonyl iron.

Figure 1: Electrolytic Iron Particles

Carrier Fluid

For the purpose of manufacturing a cheaper alternative to the commercially available MR Fluids we decided to use an easily and cheaply available vegetable oil. We made a list of the commonly available mineral oils and noted down their physical properties for the purpose of comparison. The listed comparisons helped us narrow down the list of more than 10 to two oils, coconut and soybean oil.
Properties of commonly available vegetable oils are as follows:

Table III: ***Properties of Carrier Fluid***

Vegetable Oil	**Kinematic viscosity (mm_2/s)**	**Flash Point (ºC)**	**Density (kg/l)**
Coconut	30	230 ºC	0.924
Rapseed	37	246 ºC	0.911
Castor	227	230 ºC	0.961
Palm	36	164 ºC	0.880
Peanut	39.6	271 ºC	0.902

Soybean	32.6	254 ºC	0.914
Sunflower	33.9	174 ºC	0.916
Damper Oil	40	220 ºC	0.860

Although Coconut Oil was the closest to the damper oil in terms of properties but due to its high freezing temperature (25 ºC) we chose the Soybean Oil due to its easy availability and low freezing temperature (-20 ºC).

Figure. 2: Carrier Fluid

Additives

Since the nanoparticles are difficult to procure we decided to use surfactants to improve the sedimentation properties of the MR Fluid.

Commonly available surfactants are:

- Oleic Acid -
  Density: 0.895g/cc, Boiling Point: 360 ºC
  Oleic acid is a fatty acid that occurs naturally in various animal and vegetable fats and oils. It is an odourless, colourless oil, although commercial samples may be yellowish.

- Tetramethyl Ammonium Hydroxide -
  Density: 1.01g/cc, Boiling Point: Decomposes
  Tetramethyl Ammonium Hydroxide (TMAH or TMAOH) is a quaternary ammonium salt and the solid and solutions are colourless, or yellowish if impure.

- Soy Lectin -
  Density: 0.92g/cc, Boiling Point: 150 ºC
  Lecithin is a generic term to designate any group of yellow-brownish fatty substances occurring in animal and plant tissues composed of phosphoric acid, choline, fatty acids, glycerol, glycolipids, triglycerides, and phospholipids
  Lecithin has emulsification and lubricant properties, and is a surfactant.
- Glycerol -
  Density: 1.26g/cc, Boiling Point: 290 ºC
  Glycerol is a simple polyol (sugar alcohol) compound. It is a colourless, odourless, viscous liquid that is widely used in pharmaceutical formulations.

Due to the easy availability and the relatively high boiling point, we decided to use Glycerol as the surfactant.

Figure. 3: Surfactant (Glycerine)

Preparation of MR Fluid

We took 50 ml of the chosen carrier fluid (soya oil) in a beaker. Then we measured 14 gms of iron powder (300 mesh) and added glycerine as a surfactant to it so as the surface tension can be reduced when the iron powder is mixed with the carrier liquid. Then the carrier fluid was mixed with the mixture of iron powder and glycerine. The resulting mixture was 70% by weight of iron. The mixture was then stirred with the help of a electric motor powered fan for 3-4 hrs. The rpm of the fan was fixed at 1400. A similar approach was taken to make a mixture of 60% and 80% by weight of iron.

IV. PREPARATION OF SECOND SAMPLE

Magnetic Particles

We continue using the same magnetic particle as the first sample, that is the electrolytic iron as it satisfies our most criteria.

Carrier Fluid

We had already gone over all the suitable mineral oils in the first sample, so we decided to use synthetic oil as the carrier fluid in the second sample. Many types of synthetic oil satisfy most of the criteria we look for in a carrier fluid, but the viscosity of most mineral oils is highly dependent on temperature which is undesirable[6] and the process of sealing of synthetic oils is tough. Thus, taking into consideration all the desirable properties, we came to the conclusion that Heavy Paraffin oil would be the best choice.

Table IV: ***Properties of Paraffin oil***

Appearance	Colorless and clearly oily liquid free from odour
Solubility	Insoluble in alcohol and water, soluble in chloroform and solvents
Weight/100 ml (in gm)	86
Kinematic Viscosity (37.8 C)	67 CST

The kinematic viscosity of paraffin oil is nearly double to that of the soybean oil (33 CST) we used in the first model.

Surfactant

Since the nanoparticles are difficult to procure we decided to use surfactants to improve the sedimentation properties of MR Fluid. [7]
Commonly available surfactants are:

1. Oleic Acid
2. Glycerol
3. AP 3 Grease.

We chose AP3 Grease because of its anti-friction properties, better stabilizing ability and similar density to that of base liquid i.e. heavy Paraffin oil.

Preparation Technique

Apparatus

For the preparation of Magneto Rheological Fluid material, we designed and fabricated an electrical stirrer as shown in the Fig. 4. The stirrer has an RPM range of 1400 -1800.

Figure 4: Mechanical Stirrer

We also prepared a mount to hold the stirrer in position. The stirrer head was machined from Nylon 6 which has four vertical blades which was fixed at the end of the 15cm long rod.

Specifications of the Stirrer:

- 50Hz at 220V
- 1400-1800 RPM
- 50 Wattage
- 4 Nylon-6 blades

Procedure

To prepare the Magneto Rheological fluid samples, the material composition is fixed. From the literary survey it was found that there are two methods by which the constituents of the Magneto Rheological fluid are selected

- Volume Percentage
- Weight Percentage

In this work, the constituents are used as per the volume percentage method. Two samples are prepared by using different volumes of Magneto Rheological fluid components. We varied the percentages of components used in the sample. The levels fixed for preparing those two samples are shown in the Table 5.

Table V: Fluid Composition

Sample Name	Fe Particle (ml)	Paraffin Oil (ml)	Grease (ml)	Vol % of materials w.r.t total fluid volume		
				Fe Particles	Paraffin Oil	Grease
Sample 1	60	120	20	60	30	10
Sample 2	50	140	10	70	25	5

V. EXPERIMENTATION

Sedimentation Rate

Sedimentation is the tendency of the particles in suspension to settle down in the fluid in which they are entrained and come to rest against a barrier. This is due to their motion through the fluid in response to the forces acting on them. These forces may be due to gravity, centrifugal acceleration or electromagnetism. The sedimentation of particles, in the M R Fluid, is determined by neglecting the magnetic effect on the M R Fluid particles. The inherent density difference of the carrier fluid and the iron particles results in the sedimentation of the particles, leaving an upper volume as supernatant fluid (the clarified fluid above the mud line). The upper mud line is the boundary between the supernatant fluid and turbid part of carrier oil. The sedimentation of particles is measured by visually observing the position change of the mud line. [2]

For determination of the sedimentation of the Magneto Rheological Fluid samples, a setup is used and is shown in Fig. 5. This set up contains a tube holding device and a ruler which is placed adjacent to the test tube to measure the height of supernatant fluid. With passage of time, the height of supernatant fluid increases and the height of turbid part decreases.

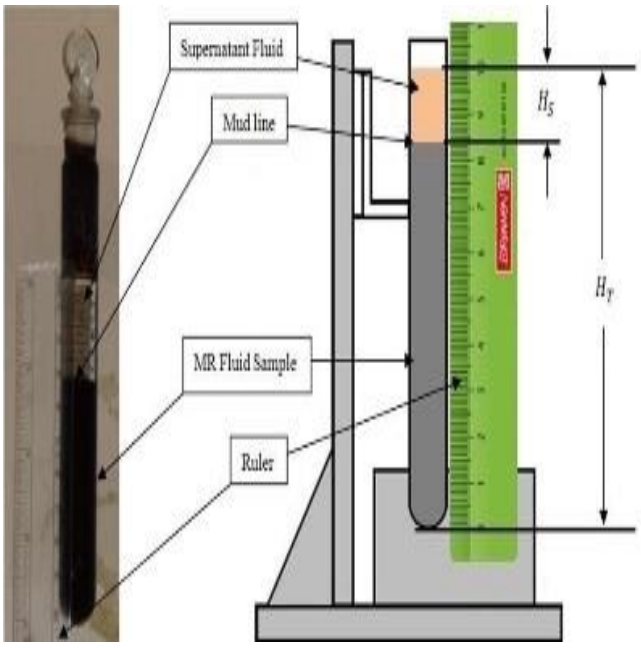


Fig 5: Sedimentation test rig

The sedimentation is observed visually under light to get a clear view of the observation being made for the height of supernatant and the turbid part of the MR Fluid sample 2. Prepared samples were filled up to 10 cm in a cylindrical glass test tube for few hours at vertical stationary position. The height of the supernatant fluid (Hs) is measured after every 5 minutes with the help of ruler and stop watch. Thus, sedimentation ratio (SR) is calculated. The sedimentation ratio is defined as the ratio of the height of the supernatant oil to the total height of the M R Fluid filled and is given as

$$SR = (HS/HT) X 100 \%$$

Where, SR is the sedimentation ratio, HS is the height of supernatant fluid and HT is the total height of fluid filled in the tube. [2]

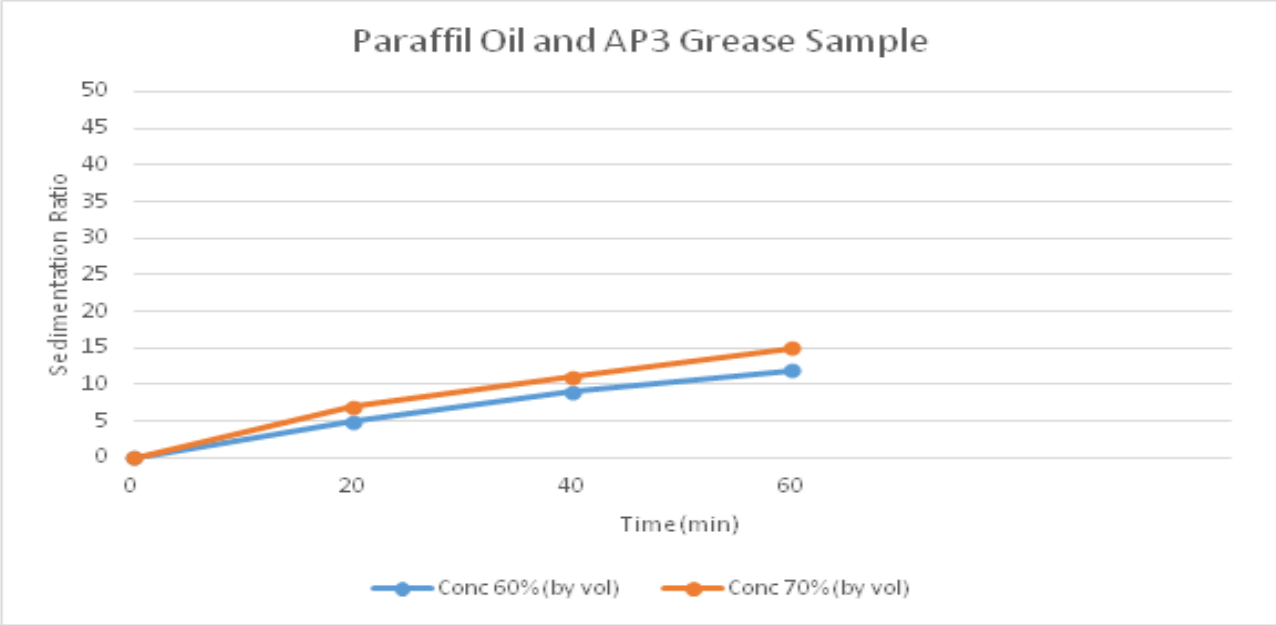


Fig. 6: Sedimentation vs. Time graph (Paraffin Oil and Ap3 Grease)

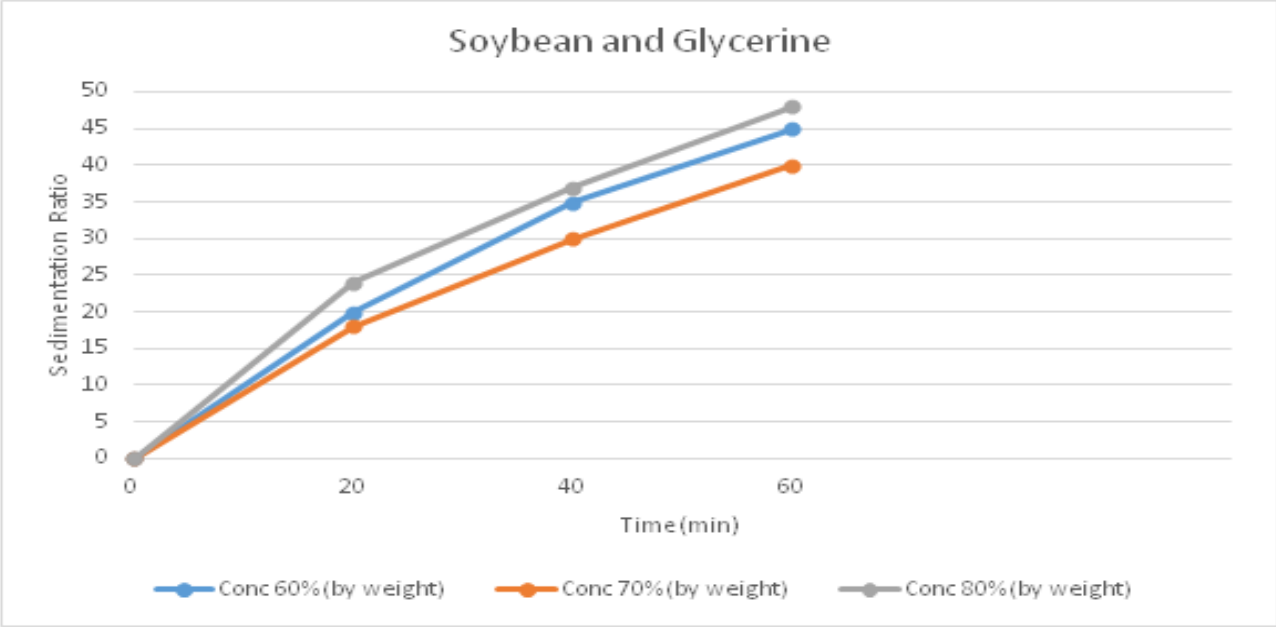


Fig. 7: Sedimentation vs. time graph (Soybean Oil and Glycerin)

Magnetoviscometer

The purpose of this experiment was to determine the coefficient of viscosity of a Magneto-Rheological fluid for different values of the magnetic field and to determine the relationship between them to develop a mechanism for controlling the damping characteristics of an automobile shocker. To determine the coefficient of viscosity, a method of indirect measurement [8] was implemented using a reference fluid with known properties.

Apparatus:

To determine the influence of the magnetic field strength on the viscosity of a fluid, a test stand was constructed, as shown in the Fig. 8. The component parts of the stand are Glass Capillary, Electrical Coil mounted coaxially with the capillary, DC power adapter making it possible to regulate voltage and current, and a manual stop watch for measurement of time. The stand is additionally equipped with magnetic field strength gauge for measurements inside in the capillary, inside the electrical coil.

Fig. 8: Test Stand

The rest of the apparatus is as shown in the figure 9.

- Stand
- Glass Capillary
- Electric Coil
- DC Current Source with means to control voltage and current
- Manual Stop watch
- Gauge of Magnetic field strength with a Hall's Probe
- Observer
- 100ml of MR Fluid sample

Fig. 9: Testing Rig

The coefficient of the viscosity of the magneto-rheological fluid was determined by an indirect method, using paraffin oil as a reference fluid of the known properties. A formula derived by Poiseuille was used for this purpose.

$H = (\rho.t)/ (\rho_g.t_g) X \eta_g$

where:

ρ – Density of the Magnet- Rheological fluid.

T – Outflow time of magneto-rheological fluid from the capillary

ρ_g – Density of reference fluid

t_g – outflow time of the reference fluid from the capillary

η_g – viscosity coefficient of the reference fluid

Procedure

1. The capillary was at first filled with the reference fluid and the capillary was allowed to drain off completely and the time taken by it to drain off was measured with the help of the stop watch.
2. The capillary was then filled with the MR fluid sample and the time taken for the fluid to drain out was measured.
3. Now a magnetic field was applied using a solenoid placed coaxially to the capillary tube. The ends of the solenoid were connected to a DC power source.
4. Step 2 was now repeated for different values of the applied magnetic field by varying the voltage applied to the solenoid.

Observations and Calculations

Table VI: Readings of Viscosity Test

Magnetic Field (mT)	Current (A)	Voltage (V)
0.4	0.1	1.0
0.5	0.4	1.5
0.8	1.1	4.0
3.1	1.8	6.5
3.4	2.5	9.0
5.2	3.5	15.0

The value of viscosity for different values of magnetic field is calculated using the given formula.

$H = (\rho.t)/ (\rho_g.t_g) X \eta_g$

Table VII: Viscosity vs. Magnetic field readings

Viscosity (Pascal-sec)	**Magnetic Field (m-Tesla)**
1.98	0
2.3	0.4
2.42	0.8

2.64	3.1
2.79	3.5
2.94	3.9
3.01	4.1
3.45	4.5

A graph is plotted between viscosity and the applied magnetic field.

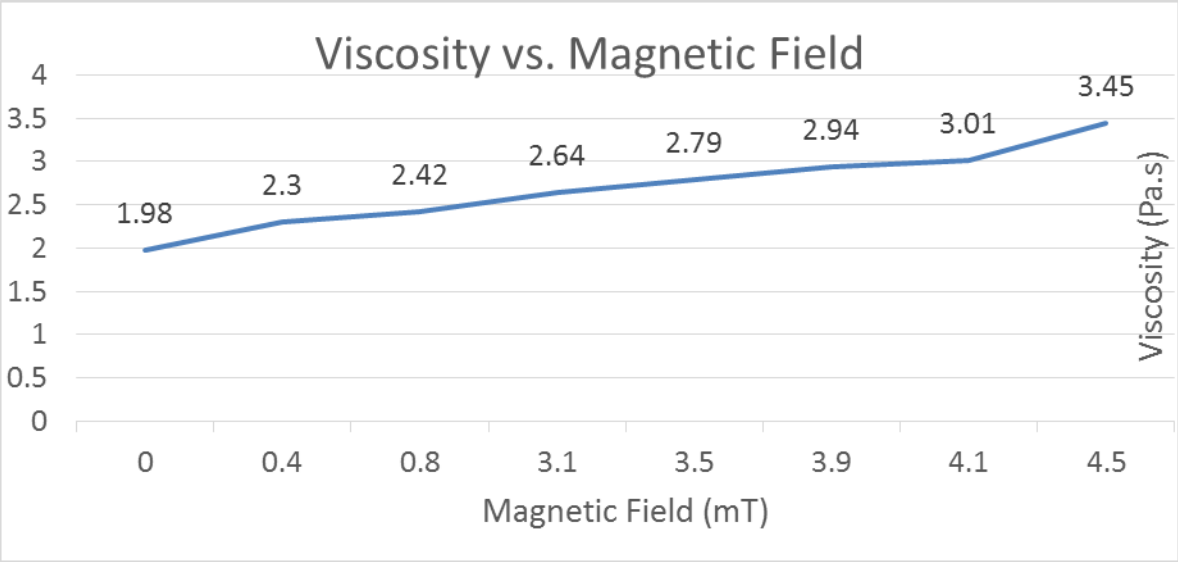


Fig 10: Viscosity vs. Magnetic Field graph

Conclusions

- After performing the sedimentation tests on the different fluid samples, we found that soybean oil and glycerin are not suitable components for making a magneto-rheological fluid. It was also observed that heavy paraffin oil and AP3 grease can be used to prepare a more stable magneto-rheological fluid.
- The aim of the second experiment was to determine the viscosity of a magneto-rheological fluid at different values of the magnetic field. From the graph in Fig. 10, it is clear that the viscosity of the fluid increases with the increasing magnetic field. The theoretical aim of this experiment was to determine this range of viscosity variations that is useful for control purposes.

REFERENCES

Journal Papers:

[1] Mark R. Jolly, Jonathan W. Bender, and J. David Carlson; "Properties and Applications of Commercial Magneto rheological Fluids"; *Journal of Intelligent Material Systems and Structures, Vol.10, No.1, PP. 5-13; January 1999.*

[2] S.K. Mangal, Mukul Kataria, Ashwani Kumar; "Synthesis of Magneto Rheological Fluid"; *IJERT, ISSN: 2249-8958,Volume-2, Issue-6; PP. 20-25; August 2013.*

[3] Horia Chiriac, George Stoian; "Influence of the Particles Size and Size Distribution on the MR Fluids Properties"; Magnetics, IEEE Transactions on (Volume:45, Issue:10) PP: 4049-4051; October 2009.

[4] T.M Simon et al. , "The Effective Magnetic Properties of Magnetorheological Fluids", Centre of Research in Scientific Computations, Mathematical and Computer Modelling 33; PP. 273-284, 2001.

[5] Ji Eun Kim, Hyoung Jin Choi; "Magnetic Carbonyl Iron Particle Dispersed in Viscoelastic Fluid and Its Magnetorheological Property"; Magnetics, IEEE Transactions on, vol.47, no.10, PP. 3173-3176, Oct. 2011.

[6] Bhau K. Kumbhar, Satyajit R. Patil, "A Study on Properties and Selection Criteria for Magneto rheological (MR) Fluid Components", International Journal of ChemTech Research; Aug/Sep2014, Vol. 6 Issue 6, p3303-3306.

[7] Kang Hyun Song, Bong Jun Park, and Hyoung Jin Choi; "Effect of Magnetic Nanoparticle Additive on Characteristics of Magneto rheological Fluid"; *Magnetics, IEEE Transactions on (Volume:45, Issue:10); PP: 4045-4048; October 2009*.

[8] A. Roszkowski, M. Bogdan, "Testing viscosity of MR Fluid in Magnetic field"; Measurement Science Review, *Volume 8, Section 3, No. 3; PP, 58-60; July 2008.*

Science Diplomacy Case Studies

2

Resumption of International Civil Nuclear Cooperation

A Diplomatic Triumph based on Strength in Science

R. B. GROVER

FISD

Forum for Indian Science Diplomacy

ABOUT THE AUTHOR

Ravi B Grover is working as an Emeritus Professor in the Homi Bhabha National Institute (HBNI) and is a member of the Atomic Energy Commission. Working as a nuclear engineer in BARC, he specialized in nuclear reactor thermal hydraulics and energy studies with special reference to nuclear energy. He worked on fuel and core thermal hydraulic analysis for a research reactor, and process safety analysis, plant dynamic analysis, process design of primary systems and equipment for a compact nuclear reactor. He studied growth of electricity requirements in India with a view to delineate the role of nuclear energy in India's electricity mix. He is a strong advocate of nuclear energy and opines that one has to factor in totality of costs while comparing different technology options for providing electricity to consumers.

He participated in negotiations with other countries and international agencies leading to resumption of international civil nuclear trade. He and his team negotiated India's entry to the ITER project in 2005. He was the sous-Sherpa of the Government of India for the Nuclear Security Summits held in 2010, 2012, 2014 and 2016.

He played a very significant role in conceptualizing and setting up HBNI and concurrent with other responsibilities, he was its Vice Chancellor for eleven years. He was Principal Adviser in the Department of Atomic Energy during 2010-13.

He was conferred with a Padma Shri in 2014 and has received several awards including Lifetime Achievement Award – 2011 by the DAE, Distinguished Alumnus Award by the Indian Institute of Science and its Alumni Association – 2011, and Lifetime Achievement Award by the Delhi College of Engineering Alumni Association – 2016. He is a fellow of the Indian National Academy of Engineering, the Maharashtra Academy of Sciences, and the World Academy of Art and Science.

© RIS, 2019

Published in September 2019 by:

RIS
Research and Information System for Developing Countries
विकासशील देशों की अनुसंधान एवं सूचना प्रणाली

Core IV-B, Fourth Floor, India Habitat Centre,
Lodhi Road, New Delhi-110 003, India
Ph.: +91-11-24682177-80, Fax: +91-11-24682173-74
E-mail: dgoffice@ris.org.in
Website: www.ris.org.in

Disclaimer: The views expressed are those of the author and need not necessarily reflect the views of RIS or DST.

CASE STUDY

Resumption of International Civil Nuclear Cooperation

A Diplomatic Triumph based on Strength in Science

R. B. GROVER

FISD
Forum for Indian Science Diplomacy

No. 2 | September 2019

Resumption of International Civil Nuclear Cooperation

A Diplomatic Triumph based on Strength in Science

1. Introduction

Energy is the driver for economic growth, and of the various options available for carrying energy, electricity is the best from the consideration of convenience of end use. As compared to the size of its population, India is poor in energy resources. Domestic production of crude and gas is very modest, therefore, coal is the mainstay of electricity generation. Considering growing electricity demand, early in this century, it was realised that coal is not going to last beyond this century and it is necessary to look for alternate sources for electricity generation. This realisation was quantified, based on a study by the Department of Atomic Energy, in a report in 2004, which was later published (Grover and Chandra 2006). The study concluded that generation by nuclear needs to be stepped up as soon as possible. However, uranium resources, as estimated at that time, were considered sufficient to support an installed capacity of only 10,000 MW of Pressurised Heavy Water Reactors (PHWRs) and it was concluded that growth in nuclear installed capacity has to come from fast breeder reactors. As we will see later in this paper, an installed capacity of 10,000 MW based on PHWRs is very modest. A two-pronged strategy was formulated by the country: (i) launch aggressive campaign to locate more uranium resources in the country, and (ii) pursue a policy initiative to open international civil nuclear trade with the objective of importing uranium, which was denied to India for not having signed the Treaty on the Non-Proliferation of Nuclear Weapons (NPT).

The NPT divides the globe into 'haves' and 'have-nots' (that is states with, and states without nuclear weapons) and it provides for peaceful nuclear explosions. In early 1970s, the International Atomic Energy Agency (IAEA) organised annual conferences on peaceful application of nuclear explosions. Once India conducted a Peaceful Nuclear Explosion (PNE) in 1974, the concept of peaceful nuclear explosion disappeared from the lexicon on the subject and IAEA stopped organising conferences. According to NPT, states that have built and exploded a nuclear device before 1 January 1967 are classified as Nuclear Weapon States (NWS). After peaceful nuclear explosion by India in 1974, votaries of the NPT cast India "into a netherworld" (Tellis 2005) and it was abruptly denied international collaboration on the peaceful uses of nuclear energy. While the NPT is based on the motto, "early bird gets the nuke" (ElBaradei, 2003), even

now there are calls[1] on India by NPT votaries to sign NPT as a Non-Nuclear Weapon State (NNWS). India's opinion is the direct opposite that India can sign NPT as an NWS, (Fidler and Ganguly 2010).[2]

A two-pronged strategy was formulated by the country: (i) launch aggressive campaign to locate more uranium resources in the country, and (ii) pursue a policy initiative to open international civil nuclear trade with the objective of importing uranium.

Returning to electricity requirements, it is desirable to understand the necessity of having nuclear energy as a part of India's electricity mix. Figure 1 is a plot of Human Development Index (HDI) versus per capita electricity consumption. HDI increases rapidly as per capita electricity consumption increases, but saturation effect is seen around per capita consumption of 5000 units.[3] One can argue that correlation between two variables cannot be called causation; but the messaging from the plot is clear that for achieving a high (close to 0.9) HDI, it is necessary to provide about 5000 units of electricity per annum per capita to the citizens of India. It is pertinent to note that some of the countries in our neighbourhood (that is with similar climate conditions) have a per capita electricity consumption, much higher than India. For example, per capita consumption in Malaysia is 4656 units, Thailand – 2621 and the world average is 3052 units (IEA 2018). In the year 2018-19, India generated about 1600 billion units (including generation from both utilities and non-utilities). Assuming India's population to be about 135 crore, in per capita terms, generation was about 1200 units. Assuming India's population will stabilise at 1.6 billons,[4] transmission and distribution losses will come down to about 7 percent, then to provide 5000 units per capita per annum, and India has to generate 8600 billion units per annum. This is about five and half times the generation in 2018-19.

After peaceful nuclear explosion by India in 1974, votaries of the NPT cast India "into a netherworld" and it was abruptly denied international collaboration on the peaceful uses of nuclear energy. […] while the NPT is based on the motto, "early bird gets the nuke", even now there are calls on India by NPT votaries to sign NPT as a Non-Nuclear Weapon State (NWS).

There are calls, at the global level as well as at the national level, by many to meet all electricity requirements by renewable sources. These calls are based on an under-estimation of total electricity requirements based on the principle of frugality and an over-estimation of the potential of renewable sources. Renewable sources, including large and small hydro, solar and wind, cannot provide more than a quarter of the total projected electricity requirements of India (Grover 2017a). The balance has to come from coal and nuclear. Climate change imperative calls for deep decarbonisation and over a period of time India has, to the extent possible, substitute nuclear for coal. Thus, nuclear has to be an essential part of India's electricity mix and the sooner its share is ramped up, better it is for climate change. That can be done only if uranium is available to fuel the reactors. Therefore, it was necessary for the Government of India to launch a policy initiative to enable India to access international market to buy uranium.

2. The Beginning of the Atomic Energy Programme

To harness the power of atom, Atomic Energy Commission was set up in 1948, that is within months of getting independence. To take care of governance issues, Atomic Energy Act was enacted in 1948 and the Department of Atomic Energy was set up in 1954. A swimming pool type research reactor APSARA was constructed in Trombay and achieved first criticality in 1956. It was followed by construction, in collaboration with Canada, of a 40 MWt research reactor CIRUS in 1960. Collaboration was established with several other countries as well including the USA, the UK, France, and Russia. As the programme was moving ahead very fast, need was felt to modernize the Atomic Energy Act and, therefore, Atomic Energy Act, 1962 replaced the 1948 Act.

Renewable sources, including large and small hydro, solar and wind, cannot provide more than a quarter of the total projected electricity requirements of India. The balance has to come from coal and nuclear.

During early years of the programme, science diplomacy was at work and India did derive benefit from international collaboration. Two Boiling Water Reactors were set up in 1960s at Tarapur, near Mumbai, on turn-key basis by a US company. Two PHWRs were set up at Rawatbhata in Rajasthan in collaboration with Canada, but the second unit had to be completed by indigenous efforts as all collaborations came to an abrupt halt after the PNE in 1974.

Nuclear has to be an essential part of India's electricity mix and the sooner its share is ramped up, better it is for climate change. That can be done only if uranium is available to fuel the reactors.

India remained steadfast in its resolve to develop nuclear technologies and achieved notable successes. Early on it was clear that India's domestic reserves are modest[5] and, therefore, it was decided to pursue a closed fuel cycle approach. It was articulated by Bhabha as a three-stage programme, which is a brilliant way to convey science-based complexities to all citizens. Over the years, technological capability has been developed for (i) exploration and mining of uranium, (ii) fabrication of a variety of fuels, (iii) heavy water production, (iv) designing and setting up research reactors, PHWRs, compact Pressurised Water Reactors (PWRs) and fast breeder reactors, (v) spent fuel reprocessing, and waste management including partitioning of actinides.

In parallel with the development of the nuclear power programme, India conducted three underground tests on 11 May 1998 and two on 13 May 1998. India became a declared nuclear weapon state.[6] Intense diplomatic activity followed, which can be gauged by the fact that there were fourteen meetings between India and the U.S. with India represented by Jaswant Singh and the U.S. by Strobe Talbott (Talbott, 2004). India's technical achievements were appreciated at the global level and became a trigger for the resumption of international civil nuclear cooperation.

3. Triggers for the Initiative and the Joint Statement of July 18, 2005

Early on it was clear that India's domestic reserves are modest and, therefore, it was decided to pursue a closed fuel cycle approach. It was articulated by Bhabha as a three-stage programme, which is a brilliant way to convey science-based complexities to all citizens.

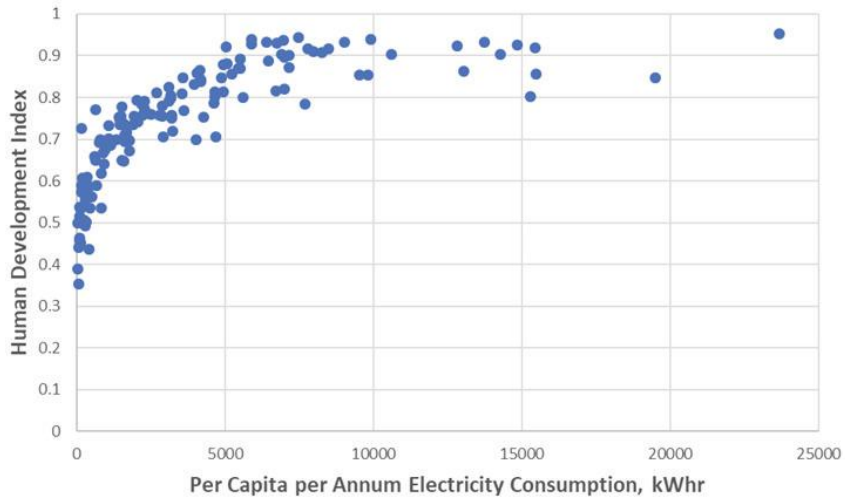
3.1 Trigger on the Indian side

Having realised the importance of nuclear as a part of electricity mix of the country, trigger for India was to have access to international uranium supplies. Uranium shortage was so acute that capacity factors of the operating reactors started coming down, reaching as low as 50 percent in 2008-09, and India could not think of constructing more PHWRs.

3.2 Triggers on the US side

On the US side, there were two different narratives for triggering the initiative: one strategic and the other scientific. Monographs by Itty Abraham and Ashley Tellis provide details about the debate in the USA (Abraham 2007; Tellis 2005) and a brief summary is given here. First public hint of the thinking about India by the US President Bush was provided during his election campaign by Condoleezza Rice (Rice 2000). She talked about the role of India in providing regional balance

Figure 1: Human Development Index versus Electricity Consumption for Different Countries



Source: Author's Compilation; Data for electricity consumption is from Key World Energy Statistics, 2018 published by International Energy Agency (IEA, 2018). Data for HDI is from "Human development Indices and Indicators" 2018 Statistical Update, published by United Nations Development Programme (UNDP, 2018).

Note: (2) HDI is a composite index calculated based on three dimensions health, education and living standard. Indicators for three dimensions are life expectancy at birth, mean years of schooling and expected years of schooling, and gross national income per capita; (2) Some important points are India (918, 0.640), Vietnam (1616, 0.694), Thailand (2868, 0.755), China (4279, 0.759), Malaysia (4656, 0.802), USA (12825, 0.924) and Iceland (not plotted) (53913, 0.938)

in Asia. After Bush took over as President, intense diplomatic engagement followed. First result was the release of a matched set of statements (Next Steps in Strategic Partnership) on 13 January 2004 by the PM of India in New Delhi and the President of the USA in Washington. Its benefit for the nuclear industry was symbolic. However, it did become a building block for a future dialogue.

Condoleezza Rice visited India on 15 March 2005, and the two sides agreed on a dialogue on energy including nuclear energy and she extended an invitation to Prime Minister to visit the USA. As a first step, the External Affairs Minister, Natwar Singh, led a delegation to the USA for a strategic dialogue in April 2005. The two visits paved the way for the visit of the Prime Minister of India to the USA in July 2005.

A few days prior to the visit of the Prime Minister to the United States, Carnegie Endowment for International Peace released a monograph "India as a global nuclear power: An action agenda by the US" (Tellis 2005). In the monograph, Tellis advocated that nothing should be done to cap India's nuclear deterrent as it would put it at a severe disadvantage vis-à-vis Beijing. He outlined several options in terms of what he called as six end states. Fifth end state proposed permitting India to have access to safeguarded nuclear fuel in return for bringing all present and future power reactors[7] under safeguards with no restriction on producing fissile material from research reactors. Sixth end state envisaged integrating India as a legitimate NWS with all privileges. Itty Abraham writes about the rise of China and a need to balance its rise, and adds India's economic growth as an important factor (Abraham 2007).

We'll return to the trigger provided by India's scientific strength in a later section.

In parallel with the development of the nuclear power programme, India conducted three underground tests on 11 May 1998 and two on 13 May 1998. India became a declared nuclear weapon state. […] India's technical achievements were appreciated at the global level and became a trigger for the resumption of international civil nuclear cooperation.

3.3 The Joint Statement

Ashley Tellis did write about growing capability of India in science and technology,[8] and an acknowledgement of scientific strength of India was embedded in the joint statement itself. Long after the Joint Statement, there was an explicit indication about the scientific strength of India having triggered the USA to act. To keep the historical perspective, let us look at the Joint Statement first. It was issued on 18 July 2005 in Washington during the visit of Prime Minister Manmohan Singh to Washington DC and involved tense negotiations. Relevant extracts are given in Annex 1. In the Joint Statement, the USA:

- appreciated India's commitment to prevent WMD proliferation and stated that as a responsible state with advanced nuclear technology, India should acquire the same benefits and advantages as other such states;

An acknowledgement of scientific strength of India was embedded in the joint statement itself. Long after the Joint Statement, there was an explicit indication about the scientific strength of India having triggered the USA to act.

India reciprocally agreed to assume the same responsibilities and practices and acquire the same benefits and advantages as other leading countries with advanced nuclear technology, such as the United States.

- committed to work to achieve full civil nuclear energy co-operation with India to enable it to realize its goals of promoting nuclear power and achieving energy security;
- agreed to seek agreement from Congress to adjust U.S. laws and policies;
- agreed to work with friends and allies to adjust international regimes to enable full civil nuclear energy co-operation and trade with India;
- agreed to consult with partners about India's interest for participation in the ITER[9] project; and
- offered to consult with the other participants in the Generation IV International Forum[10] with a view toward India's inclusion.

India reciprocally agreed to assume the same responsibilities and practices and acquire the same benefits and advantages as other leading countries with advanced nuclear technology, such as the United States. These responsibilities and practices consist of:

- identifying and separating civilian and military nuclear facilities and programmes in a phased manner;
- voluntarily placing its civilian nuclear facilities under IAEA safeguards;
- signing and adhering to an Additional Protocol with respect to civilian nuclear facilities;
- continuing India's unilateral moratorium on nuclear testing;
- working with the United States for the conclusion of a multilateral Fissile Material Cut-off Treaty;
- refraining from transfer of enrichment and reprocessing technologies to states that do not have them and supporting international efforts to limit their spread; and
- ensuring that the necessary steps have been taken to secure nuclear materials and technology through comprehensive export control legislation and through harmonization and adherence to Missile Technology Control Regime (MTCR) and Nuclear Suppliers Group (NSG) guidelines.

. Identification of facilities as civilian or military is a determination by India and is not linked to any condition. Further, placing its civilian facilities under safeguards by IAEA by India is on a voluntary basis. This again raises the status of India to an NWS

3.4 An Acknowledgement of India's Scientific Strength

Certain words in the Joint Statement are very important. These include recognition of India as "a state with advanced nuclear technology", which is a diplomatically acceptable equivalent for being a nuclear weapon state and is a recognition of India's scientific strength. Identification of facilities as civilian or military is a determination by India and is not linked to any

condition. Further, placing its civilian facilities under safeguards by IAEA by India is on a voluntary basis. This again raises the status of India to an NWS. IAEA has a standard template for the additional protocol for NNWS, but here the words are "an additional protocol" and not "the additional protocol."

A clear of recognition of India's scientific strength is also reflected in the offer to enable India's participation in the ITER project, and Generation IV international forum. India's scientific strength was explicitly acknowledged by Anish Goel,[11] when he commented that during the period of isolation, "India continued to indigenously develop its nuclear program to ever-increasing sophistication." He also wrote that the decision to negotiate with India was taken "[a]fter years of careful analysis and foundational work, supported by scientific expertise" and that India had carried extensive research in thorium technologies (Goel 2014). It appears that the US scientific community has been monitoring scientific progress in India and reporting it to the Government as can be seen by the testimony given by Hecker[12] at a Hearing of the US Senate Committee on Appropriations, Subcommittee on Energy and Water Development, on April 30, 2008: "I found that whereas sanctions slowed progress in nuclear energy, they made India self-sufficient in nuclear technologies and world leaders in fast reactor technologies. While much of the world's approach to India has been to limit its access to nuclear technology, it may well be that today we limit ourselves by not having full access to India's nuclear technology developments. Such technical views should help to advise the diplomatic efforts with India."

A clear of recognition of India's scientific strength is also reflected in the offer to enable India's participation in the ITER project, and Generation IV international forum.

One may also note the commitment made by India regarding refraining from transfer of enrichment and reprocessing technologies reflects the fact that India had developed these technologies.

The commitment made by India regarding refraining from transfer of enrichment and reprocessing technologies reflects the fact that India had developed these technologies

Growing technological capability of India was a source of concern for the USA as no treaty-based obligation to enforce export controls was applicable to India. It was realised that the pursuit of a closed fuel cycle will leave India "awash with more plutonium than it could ever use either for energy production or nuclear weaponry." (Tellis 2005: 18). While it is not correct to say that India cannot use all the plutonium for energy production, Tellis does bring out the main concern of the USA. Resulting from this concern, the US considered, tightening India's export controls far more important for its national security than capping and rolling back of India's strategic programme. Regarding export controls, India had been following and continues to follow a virtuous path, but the outside world was imposing costs by denying international trade. Ashley Tellis (2005) writes, "Beyond a certain point,

virtue cannot remain its own reward….." (P. 26). Having a deal with India was a way to ensure sustainability of India's export controls.

Growing technological capability of India was a source of concern for the USA as no treaty-based obligation to enforce export controls was applicable to India. It was realised that the pursuit of a closed fuel cycle will leave India "awash with more plutonium than it could ever use either for energy production or nuclear weaponry."

3.5 Dialogue with Other Countries

The joint statement paved the way for a dialogue to finalise the text of a Nuclear Cooperation Agreement with the USA and also other countries, notably France and Russia.

4. Core Issues Important for India and the Separation Plan

4.1 The Debate on the Likely Contour of the Separation Plan

Between July 2005 and March 2006, there was an intense debate in India on the likely contours of the initiative to resume international civil nuclear trade. Debates included views regarding economics and safety of nuclear power, perceived opaqueness of India's nuclear programme, and how conditions associated with various agreements then under negotiations would influence ongoing strategic programme. It is relevant to discuss issues related to influence on the strategic programme. Expanding and paraphrasing the list proposed by Itty Abraham (Abraham 2007: 28), the issues were the following:

The joint statement paved the way for a dialogue to finalise the text of a Nuclear Cooperation Agreement with the USA and also other countries, notably France and Russia.

- Finalising a credible and defensible separation plan;
- Ensuring that there is no moratorium on fissile material production for the strategic programme;
- Moratorium on testing;
- Costs imposed by safeguards;
- Intrusiveness of the Additional Protocol;
- Ensuring reciprocity of actions by India and the U.S.

First criticism came from a former Director of Bhabha Atomic Research Centre, AN Prasad (Varadarajan 2005), when he opined that the course of actions proposed by the Joint Statement was against national interest, fast reactor programme would be undermined and the cost of nuclear weapon programme would dramatically escalate. Note the concerns related to fast reactor programme going under safeguards. He was also concerned about the additional protocol.

Kakodkar clarified important issues early in the debate (Subramanian 2005) when he said that (i) classification of facilities as civilian would be an Indian decision, (ii) only the facilities having no strategic significance would be classified as civilian, (iii) fissile material production would not be capped,

(iv) no developmental programme including Prototype Fast Breeder Reactor (PFBR) would be offered for safeguards, (v) in view of large energy needs, India would look at imports as externalities and facilities would be offered for safeguards in a phased manner.

Concerns regarding offering facilities for safeguards were dispelled based on the fact that several Indian facilities were already under safeguards and this never posed any concern. This includes two reactors at Tarapur and two reactors at Rawatbhata in Rajasthan. Two reactors then under construction at Kudankulam in technical collaboration with Russia were also under safeguards.

The debate regarding influence on strategic programme continued and amidst this debate Kakodkar, in an interview to Pallava Bagla, provided explanation on all important issues (Bagla 2006). He extensively referred to the DAE study about electricity needs covering a 50-year horizon (Grover and Chandra, 2006) and spoke about (i) not compromising country's strategic interest and energy interest, (ii) the criticality of the fast breeder programme for the nuclear power programme of the country as well as for maintaining minimum credible deterrent, (iii) the linkage of the fast reactor programme to the indigenous PHWR programme through the fuel cycle[13] and hence, the impossibility of offering fast reactors for safeguards and also keeping sufficient PHWR capacity outside safeguards. He clarified that research and development could be done only in an autonomous manner and that India needed additional capacity [for production of fissile material from power reactors] for strategic needs.

This interview became a decisive factor in defining the contours of the Separation Plan. Commenting on the interview, Siddharth Varadarajan wrote, "Indeed, the supporters of Indo-U.S. nuclear cooperation should realise that by drawing a thick red line out in the open, Dr. Kakodkar has done the only thing which can still salvage the deal: telling the Americans that if they don't blink, the agreement will die a natural death" (Varadarajan 2006).

Between July 2005 and March 2006, there was an intense debate in India on the likely contours of the initiative to resume international civil nuclear trade. Debates included views regarding economics and safety of nuclear power, perceived opaqueness of India's nuclear programme, and how conditions associated with various agreements then under negotiations would influence ongoing strategic programme.

Concerns regarding offering facilities for safeguards were dispelled based on the fact that several Indian facilities were already under safeguards and this never posed any concern. This includes two reactors at Tarapur and two reactors at Rawatbhata in Rajasthan. Two reactors then under construction at Kudankulam in technical collaboration with Russia were also under safeguards.

4.2 The Separation Plan

While the debate was ongoing, considering responsibilities and practices identified in the Joint Statement, India started working on a plan to separate facilities. The Separation Plan was drafted to include the list of facilities as well as the basis for doing so. Several core issues important to India were embedded in the Separation Plan and their embedding in the Separation Plan was helpful in negotiating the Nuclear Cooperation Agreement with the USA.

Research and development could be done only in an autonomous manner and that India needed additional capacity [for production of fissile material from power reactors] for strategic needs.

The Separation Plan emphasises the centrality of nuclear energy in meeting the challenges of energy security and environmental sustainability. It refers to India's growing technological prowess and points to the fact that "the strategic programme is an offshoot of research on nuclear power programme and consequently, it is embedded in a larger undifferentiated programme." It brings out strengths of India's research and development efforts and dismisses grid-connectivity as irrelevant to the separation exercise. It states that "India will:

- Include in the civilian list only those facilities offered for safeguards that, after separation, will no longer be engaged in activities of strategic significance.
- The overarching criterion would be a judgement whether subjecting a facility to IAEA safeguards would impact adversely on India's national security.
- However, a facility will be excluded from the civilian list if it is located in a large hub of strategic significance, notwithstanding the fact that it may not be normally engaged in activities of strategic significance.
- A civilian facility would therefore, be one that India has determined not to be relevant to its strategic programme."

Taking the above into account, the Separation Plan offers eight indigenous reactors for safeguards by IAEA in addition to six which were already under safeguards. It offers certain facilities at Nuclear fuel Complex for safeguards and also offers a reprocessing plant for safeguards under campaign mode. It designates only three heavy water plants[14] as civilian. India did not offer Fast Breeder Test Reactor (FBTR) and under construction PFBR for safeguards as the fast breeder programme was still under development and its technology would take time to mature. India reserved its right to determine future facilities as civilian or strategic.

The Separation Plan was drafted to include the list of facilities as well as the basis for doing so. Several core issues important to India were embedded in the Separation Plan and their embedding in the Separation Plan was helpful in negotiating the Nuclear Cooperation Agreement with the USA.

The Separation Plan also addresses the important issue of fuel supply assurances in great detail (See relevant extracts in Annex 2), makes a clear linkage of fuel supplies with IAEA safeguards, and provides for India taking corrective measures to ensure uninterrupted operation of its civilian nuclear power reactors. It also brings in the concept of the safeguards agreement being India-Specific. As explained later, these features were woven into both the Nuclear Cooperation Agreement with the USA and also the India-Specific Safeguards Agreement.

DAE supports several research and development centres, industrial units, and grant-in-aid institutions. All grant-in-aid institutions, one research and development centre, and one industrial unit were declared as civilian (See Annex 2 for

the complete list) and it was specifically stated that they are safeguards irrelevant.[15]

The Separation Plan was discussed and agreed between India and the USA during the visit of President Bush to India during 2nd-4th March 2006. The Prime Minister made a Suo Moto statement in the Lok Sabha on 7 March 2006 on the Separation Plan and the Separation Plan was tabled in Parliament on 11 May 2006, communicated to IAEA and issued by IAEA for information and wider circulation (IAEA 2008a).

The Separation Plan didn't impose any restrictions on the production of fissile material and was defensible from strategic considerations. It also paved the way for addressing the remaining issues that is IAEA safeguards, additional protocol, reciprocity of actions, in documents yet to be negotiated as discussed in subsequent sections.

Separation Plan offers eight indigenous reactors for safeguards by IAEA in addition to six which were already under safeguards. It offers certain facilities at Nuclear fuel Complex for safeguards and also offers a reprocessing plant for safeguards under campaign mode. It designates only three heavy water plants as civilian.

5. Adjustment of Laws and Policies by the US

5.1 India joins ITER

In the Joint Statement, the US agreed to seek agreement from Congress to adjust U.S. laws and policies. The first action by the US was to agree[16] to India joining the ITER project. India was invited to present its case before the meeting of negotiators held at Cadarache, France on 12 September 2005. India's case was well received and after that there was a flurry of activity and Indian delegation joined the meeting of the Negotiators' Senior Support Group (NSSG) at Jeju, South Korea during early in December 2005. During NSSG meeting, the list of in-kind contributions to be made by India was negotiated and finalised. NSSG meeting was followed by a meeting of Negotiators and on 6 December 2005, where the then ITER partner countries unanimously approved Indian accession to the negotiations for finalizing the Joint Implementation Agreement (JIA) for setting up ITER. India signed the JIA along with others during a Ministerial meeting held in Paris on 21 November 2006.

India did not offer Fast Breeder Test Reactor (FBTR) and under construction PFBR for safeguards as the fast breeder programme was still under development and its technology would take time to mature. India reserved its right to determine future facilities as civilian or strategic.

5.2 The Hyde Act

Chapter 11 of the US Atomic Energy Act deals with international activities including international and export controls. Nuclear cooperation agreements are governed by the provisions of section 123 and that is why, colloquially, US refers to cooperation agreements as '123' agreements. Being not a signatory to NPT and having an active nuclear weapons programme, India didn't satisfy many provisions of the Act. To proceed with a nuclear cooperation agreement, the US Congress passed "Henry J. Hyde United States-India Peaceful Atomic Energy Cooperation Act of 2006" and it was signed into law by the President with his signature on 8 December 2006.

The Separation Plan didn't impose any restrictions on the production of fissile material and was defensible from strategic considerations. It also paved the way for addressing the remaining issues that is IAEA safeguards, additional protocol, reciprocity of actions, to name a few.

Being not a signatory to NPT and having an active nuclear weapons programme, India didn't satisfy many provisions of the Hyde Act. […] For Indian negotiators, core issues were very clear and for the US side clarity was provided by the Hyde Act. To start with, reconciling the two seemed like a challenge. It was clear that standard US templates for cooperation agreement would not work and real hard work lay ahead.

For Indian negotiators, core issues were very clear and for the US side clarity was provided by the Hyde Act. To start with, reconciling the two seemed like a challenge. It was clear that standard US templates for cooperation agreement would not work and real hard work lay ahead. The next section provides details as to how the core issues have been addressed in the Nuclear Cooperation Agreement.

6. The Nuclear Cooperation Agreements

6.1 Addressing the Core Issues

Finalisation of the Separation Plan and the passage of the Hyde Act paved the way for negotiations to arrive at the text of the Nuclear Cooperation Agreement. In view of the intense debate in the Indian media, core issues to be addressed were clear to the Indian negotiating team and as a result, the finalisation of the text took long time. Text of the Agreement was finalised after several rounds of internal discussions involving scientists from DAE and officials from MEA and PMO, and negotiations with the US team. While one has to look at the Agreement as a whole, the core issues have been addressed in the Nuclear Cooperation Agreement in the following manner.

6.1.1 Maintaining Integrity and Reliability of India's Strategic Programme, Continued Pursuit of Closed Fuel Cycle (Three-stage Nuclear Power Programme), and R&D

- Article 2.4 affirms that the Nuclear Cooperation Agreement will not affect the un-safeguarded programme thereby precluding any moratorium on production of fissile material. It mandates implementation of the agreement without any hindrance or interference with independent development activities.
- Article 1 (H) defines 'information' and is worded to provide precision in understanding. Lack of precision could have led to misunderstanding or even disputes over issues such as replication or infringement of confidentiality.
- Article 12 makes it clear that the Agreement is to be implemented in a manner that avoids hampering or delaying or interference in other activities and the provisions of the Agreement are not to be used to interfere with nuclear policy or programmes.

6.1.2 IAEA Safeguards Agreement and Fuel Supply Assurances

- Article 2.2e. includes development of a strategic reserve of nuclear fuel in the scope of cooperation.
- Article 5.6 reproduces verbatim para 15 of the 'separation plan' related to fuel supply assurances, safeguards and corrective measures (reproduced in Annex 2).

- Article 10.2 on IAEA safeguards starts with the words, "Taking into account Article 5.6 of this Agreement", thereby provides a linkage between safeguards and fuel supply assurances.
- Article 14 on 'Termination and cessation of cooperation' in its para 8 states: "It is not the purpose of the provisions of this article regarding cessation of cooperation and right of return to derogate from the rights of the Parties under Article 5.6." Article 5.6 deals with important issues related to fuel supply, safeguards and corrective measures.
- Article 14.5 refers to uninterrupted operation of nuclear reactors.

Intensity of the dialogue between the Indian and the US negotiators can be judged from the following quote, "A U.S. official who had negotiated 123 Agreements with other countries told me he had never had to explain, discuss, and argue about text and language to the degree he had done with the Indians".

6.1.3 IAEA Safeguards

- Title of the article related to safeguards is "IAEA safeguards." This is further reiterated in Article 10.2, which refers to "India-Specific Safeguards Agreement between India and the IAEA".
- Article 10.4 states, "If the IAEA decides that the application of IAEA safeguards is no longer possible, the supplier and recipient should consult and agree on appropriate verification measures."
- Article 14.3 says that finding of safeguards non-compliance has to be made only by the Board of Governors of IAEA.
- The definition of the by-product material (Article 1(A)) was chosen to ensure precision in understanding.

6.1.4 Full Civil Nuclear Cooperation

- Article 2.2 covers scope of cooperation and Article 2.2d says that scope includes nuclear reactors and aspects of associated nuclear fuel cycle including technology transfer.
- Article 5.2 provides for cooperation in sensitive nuclear technologies only after amendment.[17] This article does permit transfer of dual use items for use in sensitive facilities as per applicable laws, regulations and policies.
- Article 6 iii) grants consent for reprocessing[18] of spent fuel and provides a time limit (one year) for agreeing on arrangements and procedures for reprocessing.

The US nuclear cooperation policies are embedded deep in their domestic law, while it is not so in case of France and Russia. As a result, there was a difference in the intensity of negotiations. […] Being not a part of policy of France and Russia, reprocessing consent was not an issue for negotiations.

6.1.5 Cessation of Cooperation and Right of Return

- Article 14 is a long article and it provides a layered approach for termination. Article 14.2 provides for consultation to carefully consider the circumstances and agree to take into account concerns arising out of security environment.
- Article 16.4 provides for implementation of the Agreement in good faith and in accordance with the principles of international law.

- Article 14.3 refers to 'material violation' as per the definition of Vienna Convention on the Law of Treaties and violation of non-compliance of safeguards agreement to be determined by the Board of Governors of the IAEA.
- Article 14.5 provides for caution before exercising right of return and refers to negative consequences on on-going projects.

Countries who have not signed NPT have to accept safeguards on nuclear facilities that are established by them based on international cooperation. […] India had no obligation to offer facilities designed and constructed indigenously for safeguards by IAEA as long as it used indigenously produced fuel to operate them.

6.2 Intensity of Negotiations

It is to be noted is that core issues have been addressed in multiple ways and it can be argued that the Agreement has become verbose. However, such verbosity drives home the importance of core issues and is an insurance to protect the interests of the country in case of any dispute in the future. Intensity of the dialogue between the Indian and the US negotiators can be judged from the following quote, "A U.S. official who had negotiated 123 Agreements with other countries told me he had never had to explain, discuss, and argue about text and language to the degree he had done with the Indians" (Menon 2016: 85).

6.3 Negotiations with France and Russia

In parallel to negotiations with the USA for arriving at the text of a nuclear cooperation agreement, negotiations were also held with France and Russia. The US nuclear cooperation policies are embedded deep in their domestic law, while it is not so in case of France and Russia. As a result, there was a difference in the intensity of negotiations. Negotiating team felt that negotiators from the three countries might have been in touch and so once an agreement on an issue was reached with the US team, it was easier to deal with the other teams. Being not a part of policy of France and Russia, reprocessing consent was not an issue for negotiations.

One can categorically state that the approach followed by India is consistent with its legal obligations. All safeguards agreement referred to in paragraph 7.1 have since been subsumed in ISSA. All facilities identified as civilian in the Separation Plan have been offered for safeguards.

7. India-Specific Safeguards Agreement (ISSA)

7.1 Origin of Safeguards

Article 3 of the NPT begins with, "Each Non-nuclear-weapon State (NWS) Party to the Treaty undertakes to accept safeguards, as set forth in an agreement to be negotiated and concluded with the International Atomic Energy Agency in accordance with the Statute of the International Atomic Energy Agency and the Agency's safeguards system, for the exclusive purpose of verification of the fulfilment of its obligations assumed under this Treaty with a view to preventing diversion of nuclear energy from peaceful uses to nuclear weapons or other nuclear explosive devices." Safeguards for NWS Parties to NPT are thus

binding and comprehensive[19] in scope as all of their nuclear facilities come under safeguards. NWS offer some of their civilian facilities for safeguards to IAEA on voluntary basis. Countries who have not signed NPT have to accept safeguards on nuclear facilities that are established by them based on international cooperation.[20] There is no obligation on non-signatories to the NPT, to offer their indigenous facilities for safeguards. India had no obligation to offer facilities designed and constructed indigenously for safeguards by IAEA as long as it used indigenously produced fuel to operate them.

7.2 India's Prior Safeguards Agreements

When negotiations for ISSA began, several facilities set up based on international collaboration were already under safeguards. In addition, India had negotiated safeguards agreements for supply of nuclear fuel and heavy water. The list of all such safeguards agreements follows.

- INFCIRC/154, 3 September 1971, for safeguards for the reactors TAPS-1 & 2. (This provided for safeguards for a period of 25 years).
- INFCIRC/211, 6 November 1974, for safeguards for the reactors RAPS 1 & 2.
- INFCIRC/260, 1 July 1978, for safeguards on heavy water supplied by Soviet Union for use in RAPS-1 & 2.
- INFCIRC/374, 1 January 1990, for safeguards on nuclear fuel supplied by France for use in TAPS-1 & 2.
- INFCIRC/433, 1 May 1994 and amended on 1 November 1994 for extension of safeguards to TAPS-1 & 2 beyond the initial 25 years.[21]

Thus, India had enough prior experience in implementing safeguards on nuclear facilities and material and was never in violation of any safeguards commitment.

7.3 Important Issues for India

While negotiating ISSA, the following were the important issues:

- There should be an acknowledgement of the fact that a part of India's programme is not civilian. Applicability of ISSA should be limited to civilian facilities and there should be no interference in activities not offered for safeguards.
- ISSA must reflect India-US understandings and facilitate cooperation with all other member-states of IAEA.
- It should be an umbrella agreement and all facilities offered by India should come under it. This was necessary to avoid spending of time and efforts in negotiating safeguards agreement time and again and to preclude the possibility

The international law principle of *rebus sic stantibus*, which suggests that the duration of a binding commitment is linked to a continuation of the circumstances extant at the time of the original binding commitment. This tends to support an interpretation that India would be able legally to insist on the removal of safeguards from its civil nuclear program if there is a disruption in fuel supply."

of addition of any more conditions in future agreements. It should also provide for subsuming all earlier safeguards agreements. This was to facilitate implementation.

- Reciprocity should be built into it to ensure that India should not end up offering facilities for safeguards without reciprocal actions by the world community.

7.4 An Analysis of ISSA

As a part of the Joint Statement, India undertook to ensure that the necessary steps have been taken to secure nuclear materials and technology through comprehensive export control legislation and through harmonization and adherence to MTCR and NSG guidelines. In fact, even at the time of issuing the Joint Statement, India was implementing export controls based on a law-based regime and the regime is under continuous review for possible strengthening.

Using the guidance provided by the standard template for item-specific safeguards agreement as included in INFCIRC/66, Rev 2 (IAEA 1968), ISSA was negotiated and all the above issues were addressed. The following is an analysis of ISSA (IAEA 2009) with respect to the issues above.

- The title of the agreement includes the word 'civilian', making a subtle reference to the fact that a part of India's programme is not civilian. Third pre-ambular paragraph refers to "nuclear research and development for the welfare of its people and other peaceful purposes." This phrase was borrowed from Atomic Energy Act, 1961 and one should note that "welfare of its people" has a broad meaning including defence of the country. Article 5 states that implementation of safeguards shall not hinder or otherwise interfere with India's independent activities.
- Ninth pre-ambular paragraph refers to India-US understandings as reflected in the Joint Statement: (i) access to fuel supplies from the international market as an essential basis for India going in for ISSA, (ii) facilitating development of a strategic reserve of nuclear fuel, and (iii) India's right to take corrective measures. Twelfth pre-ambular paragraph refers to "the need to avoid hampering the peaceful uses of nuclear energy, economic and technological development or international cooperation in the peaceful uses of nuclear energy". The preamble is tightly coupled to the main agreement by the text, "NOW THEREFORE, taking into account the above, India and the Agency have agreed as follows:". Article 3 states: "The application of safeguards under this agreement is intended to facilitate implementation of relevant bilateral or multilateral arrangements to which India is a party, which are essential to the accomplishment of the objective of this Agreement."
- To make it an umbrella agreement, it provides for bringing the following three kind of facilities under its scope:
  - » Facilities identified by India in the separation plan as civilian (Article 14 (a)).
  - » Any further facilities identified by India as civilian (Article (14(a)), and

» All existing safeguarded facilities (Article 22).

Inclusion of all above makes ISSA an umbrella Agreement.

- According to ISSA, safeguards on a facility shall come into force after the following steps have been taken:
  - » Approval of the Agreement by the Board of Governors of the IAEA.
  - » Signatures on the Agreement.
  - » Notification by India about its entry into force
  - » Filing a declaration by India about its civilian facilities after all conditions conducive to the accomplishment of the objectives of the Agreement have been met (Article 13), and
  - » Notification by India offering civilian facilities for safeguards (Article 14) according to a timeframe decided by India.

All these steps ensure reciprocity.

ISSA has an Annex and to start with it had no entry. India submitted first notification on 16 October 2009 after all conditions conducive to implementation of international cooperation were met. Inclusion of all the above points makes the text of the agreement India-specific.[22]

7.5 Some Remarks about ISSA

IAEA doesn't have any template for implementing safeguards on research centres not handling nuclear material and, therefore, all such institutes remain outside the purview of ISSA. Civilian heavy water plants also remain outside ISSA. Text of the ISSA did create an unease in the non-proliferation lobby as can be seen from the following quote from Dean Rust: "According to IAEA documents and published reports, IAEA Director-General Mohammed ElBaradei asserted that the safeguards agreement meets all IAEA legal requirements, is of indefinite duration, and that the safeguards termination provisions in the operative sections of the agreement conform to longstanding IAEA principles. However, he also cited the preamble as an integral part of the agreement; the preamble contains India's position that fuel assurances are an "essential basis" for India's acceptance of safeguards under the agreement. ElBaradei also noted the international law principle of rebus sic stantibus, which suggests that the duration of a binding commitment is linked to a continuation of the circumstances extant at the time of the original binding commitment. This tends to support an interpretation that India would be able legally to insist on the removal of safeguards from its civil nuclear program if there is a disruption in fuel supply." (Rust 2008).

Essentially India follows a unique nomenclature, but items on the control lists of India are the same as in the NSG lists. Based on the list of items and guidelines for export existing in 2008, it was possible for India to convince the Nuclear Suppliers Group that India's has a robust export control system.

Nuclear fuel would remain under safeguards for all times as Article 29 of ISSA refers to termination according to GOV/1621 (20 August 1973). This is a decision of the Board of Governors of IAEA and ties duration of safeguards to actual use of the supplied items and material by the recipient state (Rockwood 2013: 17). Statement by Rust is, thus, a wrong interpretation of IAEA principles.

Approval of ISSA, with all the features referred to in para 7.4 embedded in it, by the Board of Governors was indeed a diplomatic triumph for India. Articles by academics criticising ISSA (Robertson and Carlson 2016; Carlson 2018), and response thereto (Ramakumar 2019) keep appearing. One can categorically state that the approach followed by India is consistent with its legal obligations. All safeguards agreement referred to in paragraph 7.1 have since been subsumed in ISSA. All facilities identified as civilian in the Separation Plan have been offered for safeguards. Additional facilities which came up afterwards have also been offered for safeguards. An updated list of facilities included in the annex to ISSA is given in Annex 3. This will keep getting updated as and when India offers facilities for safeguards.

Acceptable technical conditions include the fact that any nuclear reactor constructed in India has to pass India's regulatory approval. Viable tariff regime will result from a low capital cost and a financial package based on soft loan. Affordability re-emphasises viable tariff regime.

8. Export Controls

8.1 India's Export Controls in 2008

As a part of the Joint Statement, India undertook to ensure that the necessary steps have been taken to secure nuclear materials and technology through comprehensive export control legislation and through harmonization and adherence to MTCR and NSG guidelines. In fact, even at the time of issuing the Joint Statement, India was implementing export controls based on a law-based regime and the regime is under continuous review for possible strengthening. Export controls are implemented in India based on the Atomic Energy Act, 1961 (AE Act) and the Foreign Trade (Development & Regulations) Act, 1992 (FTDR Act). The FTDR Act and the rules and policies formulated thereunder constitute the basic laws governing foreign trade, whereas the AE Act provides legal framework for controlling the export of nuclear items. The Weapons of Mass Destruction and their Delivery Systems (Prohibition of Unlawful Activities) Act, 2005 (WMD Act), was enacted to prevent unlawful activities pertaining to Weapons of Mass Destruction and further strengthens India's export controls. It provides specific provisions to prohibit the export, transfer or re-transfer of nuclear weapons, nuclear explosive devices or any item usable for such purpose.

Under FTDR, a national export control list known as Special Chemicals, Organisms, Materials, Equipment and Technologies (SCOMET) has been notified. The SCOMET list has eight categories. Category '0' in the SCOMET list is defined as 'Nuclear materials, nuclear-related other materials, equipment and technology' and corresponds to Trigger List items of the Nuclear Suppliers Group. These items are notified under the AE Act as prescribed substances, prescribed equipment and related technology. Nuclear-related dual-use equipment, materials, software, and related technology are covered by Category '4' of SCOMET and corresponds to the list of dual-use items by NSG. Essentially India follows a unique nomenclature, but items on the control lists of India are the same as in the NSG lists. Based on the list of items and guidelines for export existing in 2008, it was possible for India to convince the Nuclear Suppliers Group that India's has a robust export control system.

8.2 Updating India's Export Controls

The export controls have been further updated and according to information on Directorate General of Foreign Trade (DGFT) portal, "India is a signatory to the major multilateral export control regimes, namely, Missile Technology Control Regime (MTCR), Wassenaar Arrangement (WA) and Australia Group (AG) and adherent to Nuclear Suppliers Group (NSG). India is also a signatory to international conventions on non-proliferation, namely, Chemical Weapons Convention (CWC) and Biological and Toxic Weapons Convention (BWC). Accordingly, the SCOMET control list is aligned to the control lists of the all the multilateral export control regimes and conventions" (DGFT 2018a). Various categories and licensing authorities in SOCMET are given in Annex 4. Details regarding lists and guidelines issued by the DGFT and DAE can be seen in the following documents:

- India's Export Control System, dated 4 September 2018 (DGFT 2018a).
- Special Chemicals, Organisms, Materials, Equipment and Technologies (SCOMET) export of which is regulated dated 31 January 2018 (DGFT 2018b).
- Guidelines for Nuclear Transfers (Exports) dated 28 April 2016 (DAE 2016a), and
- Updation of List of Prescribed Substances, Prescribed Equipment and Technology Under Atomic Energy Act 1962 dated 28 April 2016 (DAE 2016b).

There was a debate in India whether the waiver granted to India is 'clean' that is free of conditionalities. It appears to be free of conditionalities, but paragraphs 6 and 7 of the NSG guidelines were amended by NSG in 2011 and NPT membership was made a criterion for the transfer of enrichment and reprocessing equipment and technology. Indian press was quick to pick up the amendment and call it as converting a 'clean' waiver to an 'unclean' waiver.

9. Relaxation of NSG Guidelines

9.1 Steps Completed by India

By the middle of 2008, India had completed the following steps:

- A dialogue with the USA, France and Russia to negotiate nuclear cooperation agreements and all three texts were ready for signature;
- Formulation of a separation plan and its wider circulation as an information circular (IAEA 2008a);
- Negotiating an India-Specific Safeguards Agreement with the secretariat of the IAEA and its approval by the Board of Governors on 1 August 2008;
- Dialogue at the diplomatic level[23] with all the members of the NSG with a view to explain India's credentials including law-based export controls and seek support for a waiver from NSG for resumption of civil nuclear trade; and
- Release of a statement by EAM on 5th September 2008 to reiterate India's stand on disarmament and non-proliferation.
- WMD Act, 2005, SCOMET list and licensing guidelines issued by the DGFT, Guidelines issued by the DAE in 2006 were taken as harmonization of export controls by India.

India-Specific safeguards Agreement and the NSG waiver resulting into resumption of international civil nuclear trade have been described as the biggest foreign policy triumph by India in recent decades.

9.2 The Decision by Relaxation

As a result of this massive effort, guidelines for civil nuclear trade with India were relaxed by NSG on 6th September 2008 and have been issued by the IAEA as INFCIRC 734c (IAEA 2008b) and is reproduced here as Annex 5. The statement issued by the External Affairs Minister is a crucial link for this waiver and is reproduced as Annex 6.

It may be noted that the Statement by NSG,

- doesn't direct India to sign CTBT;
- doesn't direct India to stop producing fissile material for strategic use;
- does refer to paragraphs 6 and 7 of the NSG guidelines, as revised, which cover the transfer of enrichment and reprocessing technology; and
- doesn't ask Participating Governments to terminate cooperation in the event of a nuclear test by India, but does refer to the statement by India's External Affairs Minister (Annex 6).

9.3 Debate on the Text of the Waiver in India

There was a debate in India whether the waiver granted to India is 'clean' that is free of conditionalities. Based on the above, it

appears to be free of conditionalities, but paragraphs 6 and 7 of the NSG guidelines were amended by NSG in 2011 and NPT membership was made a criterion for the transfer of enrichment and reprocessing equipment and technology. Indian press was quick to pick up the amendment and call it as converting a 'clean' waiver to an 'unclean' waiver (Varadarajan 2011). This has the potential to add to capital costs of reprocessing and enrichment plants as explained earlier.[24]

10. Signing of Nuclear Cooperation Agreements

10.1 Intent to Purchase Reactors from the USA

NSG waiver cleared the deck for the signing of the Nuclear Cooperation Agreement with the USA as well as other countries. The US nuclear industry was looking to benefit from this initiative and therefore, an assurance was given by India to the USA about the intent to buy reactors from companies in the USA. In his testimony to the Senate Foreign Relations Committee on 18 September 2008, the Undersecretary of State William Burns said, "The Indian government has provided the United States with a strong Letter of Intent, stating its intention to purchase reactors with at least 10,000 Mega Watts (MWe) worth of new power generation capacity from U.S. firms." India, he added, "has committed to devote at least two sites to U.S. firms" (Varadarajan 2008). Negotiations with one US company for purchase of reactors.[25] are proceeding. The main issue of negotiations with companies in the USA and other countries is to arrive at "mutually acceptable technical and commercial terms and conditions that enable a viable tariff regime for electricity generated" (MEA 2010).

Indian negotiators had taken care to commit to sign only an Additional Protocol in the Joint Statement. Accordingly, India negotiated with the IAEA an Additional Protocol that limits additional obligations to provision of additional information regarding nuclear exports from India and grants no extra physical access to the IAEA

Acceptable technical conditions include the fact that any nuclear reactor constructed in India has to pass India's regulatory approval. Viable tariff regime will result from a low capital cost and a financial package based on soft loan. Affordability re-emphasises viable tariff regime. The process of negotiations for setting up reactors is likely to take more time as India insists on seeing a reference plant and the plant which India can consider as a reference plant is still under construction in the USA.

10.2 Concerns about Statements from Washington

Concerns were expressed in India about certain statements included in the documents coming from Washington, but the Government of India was very clear about what is applicable as can be seen in the following statement issued by New Delhi on 12 September 2008: "The text of the 123 agreement has

been agreed upon by the Governments of India and the US. It is a public document. The rights and obligations of the two governments are clearly spelt out in the terms and provisions of the 123 agreement. Once this Inter-governmental agreement enters into force, the agreement would become a legal document in accordance with well recognised principles of international law and the Law of Treaties" (MEA 2008).

IAEA approach to safeguards and physical security have evolved over the years, but the USA continues to follow policies that were initiated by it before the evolution and, therefore, the document on Arrangements and Procedures in a mere re-iteration of what is included in IAEA documents.

10.3 Signing of Agreements

The nuclear cooperation agreement between India and the USA was signed on 10 October 2008.

As stated earlier, negotiations with France and the USA were going on in parallel and texts of the agreements had been finalised prior to NSG granting its waiver. Agreements were signed one by one; with France on 30 September 2008 and with Russia on 5 December 2008. India-Specific safeguards Agreement and the NSG waiver resulting into resumption of international civil nuclear trade have been described as the biggest foreign policy triumph by India in recent decades with two former foreign secretaries writing extensively about it in their respective books (Menon 2016; Saran 2017).

11. India's Additional Protocol

As a part of the Joint Statement, India had agreed to sign and adhere to an Additional Protocol with respect to civilian nuclear facilities. IAEA applies safeguards to nuclear facilities declared by its Member States. Following discovery of certain undeclared programme in an NNWS, "In February 1992, the Board of Governors [of IAEA] affirmed that the scope of comprehensive safeguards agreements was not limited to verification of the non-diversion of nuclear material actually declared by a State, but included verifying the absence of undeclared nuclear material and activities in the State" (Rockwood 2013: 19). In 1997, IAEA approved the text of a new document called the model additional protocol, INFCIRC/540, with the objective to detect undeclared nuclear activities. It is not a standalone document, but an addition to the safeguards agreement between a State and IAEA. All NNWS are expected to use the standard template, while NWS can choose measures identified by them as capable of contributing to the non-proliferation and efficiency aims of the protocol. The Board also authorised the DG to negotiate protocols with "other states" "that are prepared to accept measures provided for in the Model Protocol in pursuance of safeguards effectiveness and efficiency objectives," (IAEA 1997).

Indian negotiators had taken care to commit to sign only an Additional Protocol in the Joint Statement. Accordingly, India negotiated with the IAEA an Additional Protocol that limits

additional obligations to provision of additional information regarding nuclear exports from India and grants no extra physical access to the IAEA. India's Additional Protocol also adds to safeguards obligation a commitment to provide IAEA inspectors multiple-entry visas and to allow the IAEA "free communications … including attended and unattended transmission of information generated by Agency containment and/or surveillance or measurement devices" from India's safeguarded facilities (IAEA 2014).

Overall, India's additional protocol do not impose any onerous obligations on India.

12. Arrangements and Procedures for Reprocessing

As stated earlier, the Nuclear Cooperation Agreement signed by India with the USA grants India consent for reprocessing of spent fuel, but it also states that: "To bring these rights into effect, India will establish a new national reprocessing facility dedicated to reprocessing safeguarded nuclear material under IAEA safeguards and the Parties will agree on arrangements and procedures under which reprocessing or other alteration in form or content will take place in this new facility." The Agreement goes on to prescribe time limits for arriving at arrangements and procedures. Negotiations to arrive at arrangements and procedures were started soon after signing the nuclear cooperation agreement and concluded as per the time limit.

The Nuclear Cooperation Agreement goes on to prescribe time limits for arriving at arrangements and procedures. Negotiations to arrive at arrangements and procedures were started soon after signing the nuclear cooperation agreement and concluded as per the time limit.

Article 1 of the document deals with notification and effectiveness of right to reprocess. The most difficult part in negotiating this article was regarding number of facilities India would establish. Difficulty arose because the Nuclear Cooperation Agreement uses the words 'a new national reprocessing facility'. After negotiations, this was interpreted broadly and the agreed document states 'two new reprocessing facilities' including their 'future expansion, modifications, renovations or additions'.

Article 2 relates to IAEA safeguards and includes provisions as are applied by IAEA. Article 3 provides for consultation and Article 4 covers physical protection and storage. For implementing physical protection, it refers to IAEA documents, which have been developed with the participation of experts from India. Article 5 covers environment protection and Article 6 deals with management of plutonium. In pursuit of a fast reactor programme, India will at a future date have plutonium stocks to ensure uninterrupted operation of reactors. Therefore, Article 6 begins with: "India is committed to the management

of separated plutonium in ways that are consistent with its national decisions on the civil nuclear fuel cycle."

Article 7 deals with suspension of arrangements and procedures, and provides a layered approach as in Article 14 of the Nuclear Cooperation Agreement for termination and cessation of cooperation. The remaining are routine articles.

IAEA approach to safeguards and physical security have evolved over the years, but the USA continues to follow policies that were initiated by it before the evolution and, therefore, the document on Arrangements and Procedures in a mere re-iteration of what is included in IAEA documents.

While the work on drafting a liability legislation was ongoing, the Government of India launched the initiative to resume civil nuclear cooperation as described in this paper. USA was looking forward to India joining an international liability regime, particularly the Convention on Supplementary Compensation.

13. Administrative Arrangements

Article 17 of the Nuclear Cooperation Agreement provides for establishing administrative arrangements. This also has been done. The only contentious issue in this regard was 'flagging'. India might import nuclear material from several countries and one of the countries could be say 'X'. Spent fuel might arise after irradiation in a reactor constructed in collaboration with a country say 'Y'. Nuclear material becomes obligated to both X and Y. USA requires complete material accounting for any material that is obligated to it. IAEA maintains accounting of nuclear material as a whole. India didn't want to maintain country-wise accounts of materials. This issue was resolved by making the USA agree to get material accounting reports from the IAEA. To satisfy their legal requirements, US will do calculation of their share of material based on operating record of nuclear reactors constructed by India in collaboration with companies from the USA.

Similar approach was adopted for solving the issue of flagging with two other countries who had raised this issue.

14. Civil Liability for Nuclear Damage

Nuclear liability is an area of intense debate amongst policy makers, legal community and nuclear industry professionals. India started looking at this issue towards the end of 1990s when it was decided to set up two large reactors at Kudankulam in collaboration with Russia (Grover, 2017b). To understand the issue, a study-project was awarded and the recommendation of the study included that India should enact a liability legislation and join the Convention on Supplementary Compensation. While the work on drafting a liability legislation was ongoing, the Government of India launched the initiative to resume civil nuclear cooperation as described in this paper. USA was looking forward to India joining an international liability regime, particularly the Convention on Supplementary

Compensation.[26] Ongoing drafting work was completed and "The Civil Liability for Nuclear Damage Act, 2010" was enacted. Rules under the Act were also notified.

As explained by Grover (2017b), the legislation follows established international norms regarding no-fault and strict liability regime, channelling of liability to the operator, a single court of jurisdiction etc. In addition, Indian legislation has a unique feature and that is right of recourse as included in section 17 of the Act and elaborated by Rule 24. Section 17 reads as follows:

"17. The operator of a nuclear installation, after paying the compensation for nuclear damage in accordance with section 6, shall have a right of recourse where –

i) such right is expressly provided for in a contract in writing;

ii) the nuclear incident has resulted as a consequence of an act of supplier or his employee, which includes supply of equipment or material with patent or latent defects or sub-standard services;

iii) the nuclear incident has resulted from the act of commission or omission of an individual done with the intent to cause nuclear damage."

While NPCIL is going ahead with the programme to construct PHWRs based on its own technology and Indian industry is participating in the new-build programme, an agreement for setting up reactors in collaboration with companies from France or the USA is yet be concluded.

While the section cited above uses the word 'supplier', the Act doesn't define it or use it elsewhere in the Act. It is explained by Rule 24 as follows:

""Supplier" shall include a person who –

(i) manufactures and supplies, either directly or through an agent, a system, equipment or component or builds a structure on the basis of functional specification, or

(ii) provides build to print or detailed design specifications to a vendor for manufacturing a system, equipment or component or building a structure and is responsible to the operator for design and quality assurance; or

(iii) provides quality assurance or design services."

Being a unique feature, section 17 has been debated by the industry. This debate has not taken full account of the explanation included in Rule 24. Grover (2017b) has explained it in detail and concludes: "Rule 24 provides explanation about a supplier and when one looks at the explanation along with practices in the nuclear industry, one concludes that it is the plant system designer, who is the supplier. The 'explanation' has been formulated based on an analysis of engineering practices followed by the nuclear industry. Any examination of the Indian law without studying the 'explanation' will lead to erroneous conclusions. To protect his interest against the Right of Recourse, the system designer can avail of the insurance policy that is meant for the suppliers. Sub-vendors have no liability under the Right of Recourse."

Based on the legislation, India has joined and ratified CSC. While NPCIL is going ahead with the programme to construct PHWRs based on its own technology and Indian industry is participating in the new-build programme, an agreement for setting up reactors in collaboration with companies from France or the USA is yet be concluded.

15. Outcome of the initiative

15.1 Tangible Outcomes

India started its nuclear programme in cooperation with other countries and resumed cooperation after a lapse of more than three decades in 2008. Science diplomacy was at play in the beginning of the programme when India benefited from international cooperation.

One can look at the outcome of the initiative in different[27] ways and one is to identify tangible benefits India has derived as a result on the waiver from NSG. This has been examined by Grover (2017c) and is further elaborated here:

- At the time of approval of the waiver by NSG in 2008, two reactors were under construction at Kudankulam in technical collaboration with Russia based on an inter-governmental agreement signed in 1988. The agreement was limited to construction of two reactors. Subsequent to 2008, agreements to construct four more reactors at Kudankulam have been signed with Russia. In addition, an Action Plan to construct six more reactors, 1200 MW each, at a new site was signed in October 2018. Action plan also proposes to increase the level of participation of Indian industry. The Action Plan also envisages implementation of projects jointly in third countries (Chaudhury 2018).
- Due to a mismatch between uranium production and its requirements for the operating requirements, it was not possible to operate the reactors at high capacity factors and average capacity factor dropped to 50 percent in 2008-09. Since then, India has been able to import uranium to fuel its PHWRs and their capacity factors have gone up as can be seen from Table 1. Assured of availability of imported uranium, the government has approved construction of additional PHWRs. Table 2 provides details.
- Dialogue with companies in France and USA is ongoing and one can expect agreements on more reactors to be set up in India. Sites where these reactors are proposed to be set up have already been identified.
- Post PNE in 1974, scientific collaboration between India and advanced countries slowed down not only in nuclear field, but in other fields as well. Of course, based on their personal contacts and expertise, some Indian scientists continued to collaborate with their counterparts abroad in various fields. This started changing in mid-1990s when India accepted to collaborate with CERN. Joint Statement of 2005 made a major change when the US committed to

consult with partners to secure India's participation in the ITER[28] project. India joined ITER[29] and requests to join mega-science projects keep pouring in. As written by Grover (2017c), "India has joined Joule Horowitz Reactor (JHR) being constructed by France, Facility for Antiproton and Ion Research (FAIR), Square Kilometre Array (SKA), Thirty Meter Telescope (TMT), High Intensity Proton Accelerator of Fermi Lab and recently Laser Interferometer Gravitational-Wave Observatory (LIGO)." India has also become an associate member of CERN.

- India has started building a strategic reserve of uranium. Building a reserve needs investment and cannot be done in one step. Therefore, a stepwise approach is being followed to build the reserve.

Table 1: Improvement in Capacity factors of Indian reactors over the past few years.

Total installed capacity: 6780 MW

S. No.	Year	Capacity factor %	Generation in million kW-hours
1	2008-09	50	14927
2	2009-10	61	18831
3	2010-11	71	26473
4	2011-12	79	32455
5	2012-13	80	32863
6	2013-14	83	35333
7	2014-15	82	37835
8	2015-16	75	37456
9	2016-17	80	37674
10	2017-18	70	38336
11	2018-19	70[30]	37813

Source: Grover (2017c) for 2008-09 and 2009-10; NPCIL (2019) for subsequent years.
Note: 2008-09 means the year from 1 April 2008 to 31 March 2009.

Table 2: PHWRs under construction, approved for construction, and planned (site has been approved)

Unit size: 700 MWe

S. No.	Location	No of units	Status
1	Rawatbhata, Rajastahn	2	Under construction
2	Kakrapar, Gujarat	2	Under construction
3	Gorakhpur, Haryana	4	Approved
4	Mahi Banswara	4	Approved
5	Kaiga, Karnatka	2	Approved
6	Chutka, Madhya Pradesh	2	Approved
7	Bhimpur, Madhya Pradesh	4	Planned

Source: DAE, 2018

Though India never signed NPT, yet it became a reason for denying all cooperation to India when India demonstrated its scientific strength by going in for a PNE in 1974.

15.2 Intangible Outcomes

Alternate way of looking at the outcome is to examine all intangibles, which have their own importance.

- India has joined three export control regimes namely Missile Technology Control Regime (MTCR), Wassenaar Arrangement (WA) and Australia Group (AG). It enhances India's prestige and also provides possibilities to access high technology.
- This initiative has transformed the relationship between India and the USA. To quote Menon (2016:82), "While both countries have always fought shy of saying that their partnership is to balance China, it is clear that the rise of China was one of the major spurs." He continues:"There is a creative tension at the heart of India-U.S. relations. India clearly needs U.S. technology, markets, and support to transform itself and create the stable and peaceful environment that the country needs to grow. The United State finds a stronger and more active India useful since there is a clear strategic congruence between the countries' goal in the Asia-Pacific."

India has been able to import uranium, signed contracts for building additional four reactors at Kudankulam, announced construction of several PHWRs as uranium availability is not a constraint and joined three export control regimes. India has been trying to join the fourth regime that is NSG, but due to opposition by China, NSG has not been able to come to a consensus about India's membership

16. Concluding Thoughts

Resumption of international civil nuclear trade has been transformative in several ways as described in the previous section. All steps taken by the Government were subjected to a detailed scrutiny by the country as well as international non-proliferation experts, and the government of the day had to even go through a vote of confidence in Parliament. All technical steps have been already described earlier, but to recap, Annex 7 lists important milestones (including political milestones); details of political steps have been covered by other authors (Tellis 2005; Abraham 2007; Chari (ed) 2009, Menon 2016; Saran 2017).

India started its nuclear programme in cooperation with other countries and resumed cooperation after a lapse of more than three decades in 2008. Science diplomacy was at play in the beginning of the programme when India benefited from international cooperation. It was at play when the world was divided into 'haves' and 'have nots' by NPT. Though India never signed NPT, yet it became a reason for denying all cooperation to India when India demonstrated its scientific strength by going in for a PNE in 1974. Subsequent emergence of technology control regimes is also a demonstration of how diplomacy and science are intertwined.

Despite disruption in scientific cooperation and denial of advanced technologies, India continued to grow scientifically and industrially and this was noticed by the global community.

The result was resumption of international civil nuclear trade and this was achieved without influencing ongoing strategic programme.

Beginning from the unveiling of the Joint Statement on 18 July 2005, and until the grant of the waiver by the NSG on 6 September 2008, there was an intense debate in India advising the Government about the steps it should take, red lines that should not be crossed, and even the desirability of the initiative itself. Global non-proliferation experts were worried about crumbling of the narrative built by them ever since NPT came into force.

Prior to the finalisation of the Separation Plan, differing views were expressed by individuals involved in drafting the Separation Plan, but the differences changed into a unity of purpose once the Separation Plan was finalised. The Separation Plan maintains strategic autonomy and provided clarity about what was expected by the nation from the negotiating teams. Several teams were involved in various negotiations: agreements for cooperation with the USA, France and Russia, ISSA and an additional Protocol with the Secretariat of the IAEA, arrangements and procedures for reprocessing with the US, administrative arrangements for implementing the cooperation agreements and so on.

The Separation Plan maintains strategic autonomy and provided clarity about what was expected by the nation from the negotiating teams. Several teams were involved in various negotiations: agreements for cooperation with the USA, France and Russia, ISSA and an additional Protocol with the Secretariat of the IAEA, arrangements and procedures for reprocessing with the US, administrative arrangements for implementing the cooperation agreements and so on.

Membership of the negotiating teams was chosen to ensure that scientific and diplomatic expertise is available in the team. While the author had the privilege of being associated with all the teams, to provide specific scientific expertise, domain-specific subject experts were also added to the teams. For example, an expert in reprocessing was a member of the team for discussing arrangements and procedures for reprocessing, and a nuclear chemist having experience in implementing nuclear material accounting and control was a member for discussing ISSA and India's additional protocol. Embedding experts having knowledge of nuclear science and engineering in the negotiating teams, prior experience by India in implementing safeguards on several nuclear facilities, and clarity about the goal to be achieved were important factors that helped in shaping the text of the agreements. Paying attention to domain specificity along with general science diplomacy is necessary for achieving success.

This was a landmark initiative and was achieved because of scientific strength of the research establishment, industrial capability, and clear guidelines provided by the leadership of the country, and by the joint efforts of scientists and diplomats. India has been able to import uranium, signed contracts for building additional four reactors at Kudankulam, announced construction of several PHWRs as uranium availability is not

a constraint and joined three export control regimes. India has been trying to join the fourth regime that is NSG, but due to opposition by China, NSG has not been able to come to a consensus about India's membership. While continuing diplomatic efforts regarding the membership of NSG, India should work to increase exports of components, equipment and systems for the nuclear island of reactors and provide services for construction of reactors. Efforts should also be made to export PHWRs as many developing countries may be interested in small and medium size reactors currently being built and operated by India. Countries going for a nuclear power plant for the first time also need regulatory services and this should also be explored. Increasing nuclear exports would require joint efforts by diplomats and the nuclear industry in India. The tripartite memorandum of understanding for cooperation in the construction of the Rooppur nuclear power plant in Bangladesh signed by India, Bangladesh and Russia on 1 March 2018 is a step in this direction. Overall objective should be to become a de facto nuclear supplier prior to becoming a de jure member of NSG.

Efforts should also be made to export PHWRs as many developing countries may be interested in small and medium size reactors currently being built and operated by India. Countries going for a nuclear power plant for the first time also need regulatory services and this should also be explored.

Endnotes

1. In October 2017, speaking at the UN General assembly committee, on behalf of New Agenda Coalition comprising Brazil, Egypt, Ireland, New Zealand, South Africa and Mexico, Mexico's Alternate Permanent Representative said that they were introducing a resolution urging "India, Israel and Pakistan to accede to the (Non-Proliferation) Treaty as non-nuclear-weapon states promptly and without conditions, and to place all their nuclear facilities under International Atomic Energy Agency safeguards." Accessed on 03 February 2019 at https://www.news18.com/news/india/at-un-india-denounces-call-to-give-up-nuclear-arsenal-and-sign-npt-1545073.html.
2. "On November 29, 2009, Manmohan Singh, India's Prime Minister, stated on Fareed Zakaria's GPS show that India wants to join the NPT as a nuclear-weapons state (NWS) and become the sixth NPT-recognized nuclear power." (Fidler and Ganguly, 2010)
3. The word 'Unit' is colloquial equivalent of the scientific term 'kWhr' and is used throughout the text.
4. This may be considered as a low estimate. This was chosen to give a lower value for the estimate of projected electricity requirements.
5. As mentioned in the Introduction, an aggressive campaign was launched to locate more reserves and increase domestic production of uranium and that is now showing results. More mines have been opened in Jharkhand and mining in Andhra Pradesh has also been started. Even after all these efforts, reserves continue to be modest in view of large energy requirements of the country. However, this paper is focused on resumption of international trade.
6. Excerpts from the statement by the Prime Minister Atal Bihari Vajpayee in Lok Sabha on 27 May 1998, "India is now a nuclear weapon state. This is a reality that cannot be denied. It is not a conferment that we seek; nor is it a status for others to grant. It is an endowment to the nation by our scientists and engineers."
7. Emphasis added.
8. Ashley writes, "Moreover, as successive U.S. chairmen [of Nuclear Regulatory Commission] who have visited India have come to appreciate, Indian nuclear science could actually contribute to the success of these research efforts, thanks, paradoxically, to the enormous repository of indigenous theoretical and engineering capability that has been developed as a result of decades of forced isolation." (Tellis, 2005: 21)
9. When first proposed, ITER was an abbreviation for International Thermonuclear Experimental Reactor. Iter is Latin for 'the way' and the earlier detail has been dropped. ITER began in 1985 as a Reagan–Gorbachev initiative with the equal participation of the Soviet Union, European Union, the United States, and Japan. Detailed design was completed in 2001 under the auspices of IAEA. Negotiation for the launch of the project started thereafter and over the years more Parties joined the original four.
10. As per the information on its website, the Generation IV International Forum (GIF) is a co-operative international endeavour which was set up to carry out the research and development needed to establish the feasibility and performance capabilities of the next generation nuclear energy systems. At present, the forum has 14 members. Based on goals adopted by GIF, six nuclear energy systems have been selected by it for further development. Selected systems do include systems based on closed fuel cycle approach, which India has been pursuing right from the beginning of the nuclear power programme. After a careful analysis of pros and cons, it was decided not to join GIF.
11. When nuclear cooperation agreement was being negotiated, Anish Goel was working in the White House.
12. Siegfried S. Hecker served as the Director of the Los Alamos National Laboratory from 1986 till 1997 and is now with Stanford University. He visited India several times and interacted with the scientific community.
13. Spent fuel arising from PHWRs is reprocessed to recover Plutonium, which is fabricated into fuel for the Fast Breeder Reactors. To keep fast breeder reactors outside safeguards, it is necessary to keep PHWRs also outside safeguards.
14. Heavy water plants at Thal, Tuticorin and Hazira were designated as civilian. All three are ammonia based, are co-located with fertilizer plants and draw feed synthesis gas from the fertilizer plant. Hydrogen-sulphide-based plants at Kota and Manuguru are independent plants. The sixth plant at Baroda is ammonia based, and is collocated with a fertilizer plant. Baroda plant has been modified by adding an Ammonia Water Exchange Front-End unit making it independent of the co-located fertilizer plant.
15. IAEA has standard templates for implementing safeguards: one template for implementing comprehensive safeguards (see information circular INFCIRC/153, (IAEA1972)) which is applicable to a NNWS as a whole; and the other for item-specific safeguards (see information circular INFCIRC/66, (IAEA 1966)). Non-signatories to the NPT have to offer facilities established based on international cooperation for item-specific safeguards. Both these templates are meant for safeguarding nuclear reactors and fuel cycle facilities. While comprehensive safeguards cover everything by definition, there are no guidelines for safeguarding research institutes or heavy water plants in the item-specific template that is INFCIRC/66.
16. Section 57 of the US Atomic Energy Act prohibits transfer of technologies for the production of special nuclear materials. Informal dialogue with colleagues from USA gives an indication that a determination was made that ITER has no such possibilities.
17. India has developed all three sensitive nuclear technologies namely heavy water production, spent fuel reprocessing and uranium enrichment. While the complete process flow sheet has been designed by the units of DAE based on indigenous

efforts, equipment for setting up plants are sourced from industry. From past experience, one finds that cost of setting up plants can be brought down if equipment are sourced from global market based on open global tenders. Indian negotiators wanted to include cooperation in sensitive nuclear technology with the objective of sourcing equipment from the global market, but the US side responded that it was not possible and so opted for a forward-looking language which says that transfer of sensitive technologies would become possible only after an amendment.

18. Granting consent for reprocessing on a programmatical basis is not a normal part of the US policy. Such a consent is included only in the nuclear cooperation agreement concluded by the US with Japan. Pursuit of closed fuel cycle is an integral part of India's policy and it has been explicitly included in the Agreement. However, as per the US law, for giving consent it is necessary to have arrangements and procedures for implementing the consent and these were negotiated subsequently within the time limit specified. More on arrangements and procedures is covered in a later section.
19. Please see endnote xv.
20. Please see endnote xv.
21. The agreement of cooperation between India and the United States for setting up TAPS-1 & 2 expired n 24 October 1993. According to the terms of the agreement, the safeguards agreement also expired on that date. A new safeguards agreement was negotiated to continue safeguards on TAPS-1 & 2 on voluntary basis by India.
22. In the initial stages, there was some criticism from experts in India about the agreement not being India specific. This subsided after all questions were answered in a Press Conference on 12 July 2008 (Smith and Chakravarthi, 2008).
23. Shivshankar Menon has described in detail about the position of various countries during the debate and how they were brought on board (Menon 2016:72). Shyam Saran lobbied with various NSG countries (Saran 2017: 227). China's opposition was formidable. It looked at the US-India nuclear agreement as a major blow to the international non-proliferation regime (Kumara and Jayasekera, 2008).
24. Please see endnote xvii.
25. An incident narrated by Shivshankar Menon provides an interesting insight, "On September 25, 2008, President Bush hosted a small private dinner for Prime Minister Manmohan Singh in the white house for about ten persons, when the 123 Agreement was awaiting approval by the U.S. senate. Secretary of State Rice leaned over and asked Singh when India would be ordering reactors from Westinghouse. Bush cut her off immediately and said that this was not about reactor sales but about much bigger things. Singh did not have to reply" Menon 2016: 82)
26. HYDE Act, SEC 103 (b) (3) (E) says, "Secure India's ratification of or accession to the Convention on Supplementary Compensation for Nuclear damage, done at Vienna on September 12, 1997."
27. Itty Abraham has approached the issue of outcomes from a different perspective. He poses several hypothetical questions and gives his response. My short response to questions posed by him is that this deal is about civilian programme, India has been and will continue to follow a law-based export control regime, research collaborations and nuclear installed capacity in India will grow as result of this deal.
28. EU was looking at India to join ITER even prior to the US taking the initiative. However, without the support by the US, it would not have been possible for India to join ITER.
29. In view of its high cost, ITER project came in for criticism in the USA, but a recent report from National Academies Press has strongly supported the project (USNAP 2018). High project cost remains a concern for all Parties.
30. Low capacity factors for the years 2017-18 and 2018-19 are due to initial problems with reactors at Kudankulum.
31. This has indeed happened. Institute for Plasma Research is the lead institute for India's participation in ITER.
32. Adapted and expanded from the version published by "India Today", Indo-Asian News Service, Washington, 2 October 2008.

References

Abraham I. 2007. "Origins of the Unites States-India Nuclear Agreement", Working Paper No 9, East-West Centre, Washington.

Bagla, P. 2006. "US shifting goalpost on n-deal: Atomic Energy chief," Indian Express, Feb. 6, 2006. It has been reproduced at and was accessed on 9 February 2019 at http://www.anilkakodkar.in/nuclear_pdf/kakokdar_interview_with_pallava_bagla_in_the_indian_express_2012_08_07_10_51_56_718.pdf.

Carlson J. 2018. "India's Nuclear Safeguards: Not fit for purpose", Belfer centre for Science and International Studies, Harvard Kennedy School, Cambridge, USA.

Chari P R (ed). 2009. Indo-US Nuclear Deal: Seeking Synergy in Bilateralism, Routledge.

Chaudhury D. R. 2018. "India, Russia sign civil nuclear Action Plan for second plant; third country project", Economic Times. 6 October 2018.

DAE. 2016a. "Guidelines for Nuclear Transfers (Exports)", accessed on 11 February 2019 at http://www.dae.nic.in/writereaddata/nucl_tr_0516.pdf.

DAE. 2016b. "Updation of List of Prescribed Substances, Prescribed Equipment and Technology Under Atomic Energy Act 1962", accessed on 12 February 2019 at http://www.dae.nic.in/writereaddata/pres_subs_0516.pdf.

DAE. 2018. "Rajya Sabha Unstarred Question No 2243", accessed on 12 February 2019 at http://dae.nic.in/writereaddata/parl/budget2018/rsus2243.pdf

DGFT. 2018a. "India's Export Control System", accessed on 12 February 2019 at http://dgft.gov.in/sites/default/files/SCOMETHelp04092018.pdf.

DGFT. 2018b. "Special Chemicals, Organisms, Materials, Equipment and Technologies (SCOMET) export of which is regulated", accessed on 12 February 2019 at https://www.dgft.gov.in/sites/default/files/SCOMETLIST31012018.pdf.

ElBaradei, M. 2003. "Towards a safer world", The Economist, 16 October 2003.

Fidler. P. F and Ganguly. S. 2010. "India Wants to Join the Non-Proliferation Treaty as a Weapon State", 27 January 2010, Yale Global Online, accessed on 03 February 2019 at https://yaleglobal.yale.edu/content/india-wants-join-non-proliferation-treaty-weapon-state.

Goel, A. 2014. "Nuclear Negotiations, Scientific Literacy, and U.S.-India Relations", Science and Diplomacy, December.

Grover, R. B. and Chandra, S. 2006. "Scenario for growth of electricity in India," Energy Policy, 34, pp 2834-2847.

Grover, R. B. 2014. "Nuclear power for India's energy security: external and internal challenges", in India's National Security: Annual Review, edited by Satish Kumar, Routledge, ISBN 978-1-138-79638-6, pp 421-444.

Grover, R. B. 2017a. "Powering aspirational India", The Hindu. 31 August 2017.

Grover, R. B. 2017b. "The Civil Liability for Nuclear Damage Act of India: An Engineering Perspective regarding Supplier's Liability", Progress in Nuclear Energy, 101 (2017), pp.168-175.

Grover, R. B. 2017c. "Opening up of International civil nuclear cooperation with India and related developments", Progress in Nuclear Energy, 101 (2017); pp.160-167.

IAEA. 1968. "The Agency's safeguards system", INFCIRC/66 Rev 2.

IAEA. 1972. "The Structure and content of Agreements Between the Agency and States required with the Treaty on the Non-Proliferation of Nuclear Weapons", INFCIRC/153.

IAEA. 1997. "Model Protocol Additional to the Agreements between State(s) and the International Atomic Energy Agency for the Application of Safeguards," INFCIRC/540.

IAEA. 2008a. "Communication dated 25 July 2008 received from the Permanent Mission of India concerning a document entitled "Implementation of the India-United States Joint Statement of July 18, 2005: India's Separation Plan", INFCIRC/731.

IAEA. 2008b. "Communication dated 10 September 2008 received from the Permanent Mission of Germany to the Agency regarding a "Statement on Civil Nuclear Cooperation with India", INFCIRC/734 (Corrected).

IAEA. 2009. "Agreement between the Government of India and the International Atomic Energy Agency for the Application of Safeguards to Civilian Nuclear Facilities", INFCIRC/754.

IAEA. 2014. "Protocol Additional to the Agreement between the Government of India and the International Atomic Energy Agency for the Application of Safeguards to Civilian Nuclear Facilities", INFCIRC/754/Add 6.

IAEA. 2018.. "Agreement between the Government of India and the International Atomic Energy Agency for the Application of Safeguards to Civilian Nuclear Facilities", INFCIRC/754/Add 9.

IEA. 2018. Key World Energy Statistics. Accessed on 12 February 2019 at https://webstore.iea.org/

Katz, A. 1992. "Fusion Power: Is ITER the way," Bulletin of the Atomic Scientists, October 1992; pp. 8-10.

Kumara, K. and Jayasekera, D. 2008, Nuclear Supplier Group gives India unique "waiver," but only after row between Delhi and Beijing, World Socialist Web Site, accessed on 29 April 2019 at https://www.wsws.org/en/articles/2008/09/nucl-s17.html.

MEA. 2008. "Response by Official Spokesperson to news report regarding the Civil Nuclear Initiative", September 12, 2008, accessed on 12 February 2019 at https://www.mea.gov.in/media-briefings.htm?dtl/3030/Response+by+Official+Spokesperson+to+news+report+regarding+the+Civil+Nuclear+Initiative.

MEA. 2010. "Joint Statement of the Prime Minister Dr. Manmohan Singh of India and the President Barak Obama of the USA", 8 November 2010, accessed on 2 March 2019 at http://mea.gov.in/bilateral-documents.htm?dtl/5081/Joint+Statement+of+Prime+Minister+Dr+Manmohan+Singh+and+President+Barack+Obama.

Menon, S. 2016. Choices: Inside the Making of India's Foreign Policy. Penguin Books.

NPCIL. 2019. Nuclear Power Generation (2010-11 to 2019-20), accessed on 28 April 2019 at https://www.npcil.nic.in/content/302_1_AllPlants.aspx

Ramakumar, K L. 2019. "Critical Analysis of India's Safeguards Agreement INFCIRC/754 with the International Atomic Energy Agency (IAEA)", Strategic Analysis, 43:2, pp. 114-125.

Condoleezza, R. 2000. "Campaign 2000: Promoting the National Interest", Foreign Affairs, accessed on 05 February, 2019 at https://www.foreignaffairs.com/articles/2000-01-01/campaign-2000-promoting-national-interest

Robertson, K. A. and Carlson, J. 2016. "Three overlapping streams of India's nuclear program", Belfer centre for Science and International Studies, Harvard Kennedy School, Cambridge, USA.

Rockwood, L. 2013. "Legal Framework for IAEA Safeguards"; IAEA.

Rust, D. 2008. "Doubts linger about India's safeguards", Carnegie Endowment for International Peace, August 14, accessed on 10 February at https://carnegieendowment.org/2008/08/14/questions-linger-about-india-s-safeguards-pub-20390.

Saran S. 2017. “How India sees the World: Kautilya to the 21st Century”, Juggernaut.

Smith, G. and Chakravarthi. R. 2008. “India Specific Safeguards Agreement”, 30 July 2008, Mainstream, Vol XLVI No 32.

Talbott, S. 2004, Engaging India: Diplomacy, Democracy and the Bomb, Penguin: Viking.

Tellis, A. 2005. “India as a New Global Power: An Action Agenda for the United States”, Carnegie Endowment for International Peace.

Subramanian. T.S. 2005. “Identifying a civilian nuclear facility is India’s decision”, The Hindu, 12 August,.

UNDP. 2018. “Human development Indices and Indicators”, 2018 Statistical Update, United Nations Development Programme.

USNAP. 2018. “Final Report of the Committee on a Strategic Plan for U.S. Burning Plasma Research”, a consensus report of the National Academies of Sciences, Engineering and Medicine.

Varadarajan, S. 2005. “Nuclear bargain may prove costly in long run”, 20 July, 2005, The Hindu.

Varadarajan, S. 2006. “Questions mark over Indo-U.S. nuclear deal”, 07 February 2006, The Hindu.

Varadrajan, S. 2008. “India offers 10,000 MW of nuclear contracts to U.S.”, The Hindu, 20 September.

Varadarajan, S. 2011. “NSG ends India’s ‘clean’ waiver”, The Hindu. 24 June.

Annexure 1

Extracts from the Joint Statement

Recognizing the significance of civilian nuclear energy for meeting growing global energy demands in a cleaner and more efficient manner, the two leaders discussed India's plans to develop its civilian nuclear energy programme.

President Bush conveyed his appreciation to the Prime Minister over India's strong commitment to preventing WMD proliferation and stated that as a responsible state with advanced nuclear technology, India should acquire the same benefits and advantages as other such states. The President told the Prime Minister that he will work to achieve full civil nuclear energy cooperation with India as it realizes its goals of promoting nuclear power and achieving energy security. The President would also seek agreement from Congress to adjust U.S. laws and policies, and the United States will work with friends and allies to adjust international regimes to enable full civil nuclear energy cooperation and trade with India, including but not limited to expeditious consideration of fuel supplies for safeguarded nuclear reactors at Tarapur. In the meantime, the United States will encourage its partners to also consider this request expeditiously. India has expressed its interest in ITER and a willingness to contribute. The United States will consult with its partners considering India's participation. The United States will consult with the other participants in the Generation IV International Forum with a view toward India's inclusion.

The Prime Minister conveyed that for his part, India would reciprocally agree that it would be ready to assume the same responsibilities and practices and acquire the same benefits and advantages as other leading countries with advanced nuclear technology, such as the United States. These responsibilities and practices consist of identifying and separating civilian and military nuclear facilities and programmes in a phased manner and filing a declaration regarding its civilians facilities with the International Atomic Energy Agency (IAEA); taking a decision to place voluntarily its civilian nuclear facilities under IAEA safeguards; signing and adhering to an Additional Protocol with respect to civilian nuclear facilities; continuing India's unilateral moratorium on nuclear testing; working with the United States for the conclusion of a multilateral Fissile Material Cut Off Treaty; refraining from transfer of enrichment and reprocessing technologies to states that do not have them and supporting international efforts to limit their spread; and ensuring that the necessary steps have been taken to secure nuclear materials and technology through comprehensive export control legislation and through harmonization and adherence to Missile Technology Control Regime (MTCR) and Nuclear Suppliers Group (NSG) guidelines.

The President welcomed the Prime Minister's assurance. The two leaders agreed to establish a working group to undertake on a phased basis in the months ahead the necessary actions mentioned above to fulfill these commitments. The President and Prime Minister also agreed that they would review this progress when the President visits India in 2006.

Annexure 2

Extracts from the Separation Plan

14 vii) **Research Facilities**: India will declare the following facilities as civilians:

α) Tata Institute of Fundamental Research

β) Variable Energy Cyclotron Centre

χ) Saha Institute of Nuclear Physics

δ) Institute for Plasma Research

ε) Institute of Mathematical Sciences

ϕ) Institute of Physics

γ) Tata Memorial Centre

η) Board of Radiation and Isotope Technology

ι) Harish-Chandra Research Institute

These facilities are safeguards irrelevant. It is our expectation that they will play a prominent role in international cooperation.[31]

15. Safeguards:

a) The United States has conveyed its commitment to the reliable supply of fuel to India. Consistent with the July 18, 2005, Joint Statement, the United States has also reaffirmed its assurance to create the necessary conditions for India to have assured and full access to fuel for its reactors. As part of its implementation of the July 18, 2005, Joint Statement the United States is committed to seeking agreement from the U.S. Congress to amend its domestic laws and to work with friends and allies to adjust the practices of the Nuclear Suppliers Group to create the necessary conditions for India to obtain full access to the international fuel market, including reliable, uninterrupted and continual access to fuel supplies from firms in several nations.

b) To further guard against any disruption of fuel supplies, the United States is prepared to take the following additional steps:

ι. The United States is willing to incorporate assurances regarding fuel supply in the bilateral U.S.-India agreement on peaceful uses of nuclear energy under Section 123 of the U.S. Atomic Energy Act, which would be submitted to the U.S. Congress.

ιι. The United States will join India in seeking to negotiate with the IAEA an India-specific fuel supply agreement.

ιιι. The United States will support an Indian effort to develop a strategic reserve of nuclear fuel to guard against any disruption of supply over the lifetime of India's reactors.

ιϖ. If despite these arrangements, a disruption of fuel supplies to India occurs, the United States and India would jointly convene a group of friendly supplier countries to include countries such as Russia, France and the United Kingdom to pursue such measures as would restore fuel supply to India.

c) In light of the above understandings with the United States, an India-specific safeguards agreement will be negotiated between India and the IAEA providing for safeguards to guard against withdrawal of safeguarded nuclear material from civilian use at any time as well as providing for corrective measures that India may take to ensure uninterrupted operation of its civilian nuclear reactors in the event of disruption of foreign fuel supplies. Taking this into account, India will place its civilian nuclear facilities under India-specific safeguards in perpetuity and negotiate an appropriate safeguards agreement to this end with the IAEA.

Annexure 3

List of facilities subject to safeguards under ISSA

	Facility offered for Safeguards by India	Date of receipt of notification
1.	Uranium Oxide Plant (Block A), Nuclear Fuel Complex, Hyderabad	16 October 2009
2.	Ceramic Fuel Fabrication Plant (Pelletizing)(Block A), Nuclear Fuel Complex, Hyderabad	16 October 2009
3.	Ceramic Fuel Fabrication Plant (Assembly)(Block A), Nuclear Fuel Complex, Hyderabad	16 October 2009
4.	Enriched Uranium Oxide Plant, Nuclear Fuel Complex, Hyderabad	16 October 2009
5.	Enriched Fuel Fabrication Plant, Nuclear Fuel Complex, Hyderabad	16 October 2009
6.	Gadolinia Facility, Nuclear Fuel Complex, Hyderabad	16 October 2009
7.	TAPS 1 -Tarapur Atomic Power Station, Unit 1	16 October 2009
8.	TAPS 2 - Tarapur Atomic Power Station, Unit 2	16 October 2009
9.	RAPS 1 - Rajasthan Atomic Power Station, Unit 1	16 October 2009
10.	RAPS 2 - Rajasthan Atomic Power Station, Unit 2	16 October 2009
11.	KK 1 -Kundankulam Nuclear Power Plant, Unit 1	16 October 2009
12.	KK 2 - Kundankulam Nuclear Power Plant, Unit 2	16 October 2009
13.	RAPS 5 - Rajasthan Atomic Power Station, Unit 5	19 October 2009
14.	RAPS 6 - Rajasthan Atomic Power Station, Unit 6	19 October 2009
15.	RAPS 3 - Rajasthan Atomic Power Station, Unit 3	9 March 2010
16.	RAPS 4 - Rajasthan Atomic Power Station, Unit 4	9 March 2010
17.	KAPS 1 - Kakrapar Atomic Power Station	3 December 2010
18.	KAPS 2 - Kakrapar Atomic Power Station	3 December 2010
19	Away from Reactor (AFR) Fuel Storage Facility, Tarapur	11 December 2012
20.	Nuclear Material Store at Tarapur	11 March 2014
21.	NAPS 1 - Narora Atomic Power Station, Unit 1	12 December 2014
22.	NAPS 2 - Narora Atomic Power Station, Unit 2	12 December 2014
23	Kakrapar Atomic Power Station Unit 3 (KAPS-3)	11 September 2017
24	Kakrapar Atomic Power Station Unit 4 (KAPS-4)	11 September 2017
25	KK-3 – Kudankulam Nuclear Power Plant, Unit 3	7 May 2018
26	KK-4 – Kudankulam Nuclear Power Plant, Unit 4	7 May 2018

Source: IAEA, 2018

Annexure 4

Various categories and licensing authorities in SCOMET

SCOMET Category	SCOMET items	Licensing Jurisdiction
0	Nuclear materials, nuclear-related other materials, equipment and technology	DAE
1	Toxic chemical agents and other chemicals	DGFT
2	Micro-organisms, Toxins	DGFT
3	Materials, Materials Processing Equipment and related Technologies	DGFT
4	Nuclear-related other equipment and technology, not controlled under Category '0'	DGFT
5	Aerospace systems, equipment, including production and test equipment, and related technology	DGFT
6	Munitions List	DD
7	Reserved	
8	Special Materials and Related Equipment, Material Processing, Electronics, Computers, Telecommunications, Information Security, Sensors and Lasers, Navigation and Avionics, Marine, Aerospace and Propulsion.	DGFT

Source: DGFT, 2018a

Note: DAE: Department of Atomic Energy

DGFT: Directorate General of Foreign Trade, Ministry of Commerce and Industry

DD: Department of Defence Production, Ministry of Defence

Annexure 5

Nuclear Suppliers Group's Statement on Civil Nuclear Cooperation with India (IAEA, 2008b)

1. At the Extraordinary Plenary Meeting on 6 September 2008 the Participating Governments of the Nuclear Suppliers Group decided that they:
   a. Desire to contribute to the effectiveness and integrity of the global nonproliferation regime, and to the widest possible implementation of the provisions and objectives of the Treaty on the Non-Proliferation of Nuclear Weapons;
   b. Seek to avert the further spread of nuclear weapons;
   c. Wish to pursue mechanisms to affect positively the non-proliferation commitments and actions of all states;
   d. Seek to promote fundamental principles of safeguards and export controls for nuclear transfers for peaceful purposes; and
   e. Note the energy needs of India.
2. Participating Governments have taken note of steps that India has voluntarily taken with respect to the following commitments and actions:
   a. Deciding to separate civilian nuclear facilities in a phased manner and to file a declaration regarding its civilian nuclear facilities with the IAEA, in accordance with its Separation Plan (circulated as INFCIRC/731);
   b. Concluding negotiations with the IAEA and obtaining approval by the Board of Governors on 1 August 2008 for an "Agreement between the Government of India and the IAEA for the Application of Safeguards to Civilian Nuclear Facilities," in accordance with IAEA standards, principles, and practices (including IAEA Board of Governors Document GOV/1621;)
   c. Committing to sign and adhere to an Additional Protocol with respect to India's civil nuclear facilities;
   d. Refraining from transfer of enrichment and reprocessing technologies to states that do not have them and supporting international efforts to limit their spread;
   e. Instituting a national export control system capable of effectively controlling transfers of multilaterally controlled nuclear and nuclear-related material, equipment and technology;
   f. Harmonizing its export control lists and guidelines with those of the Nuclear Suppliers Group and committing to adhere to the Nuclear Suppliers Group Guidelines; and
   g. Continuing its unilateral moratorium on nuclear testing, and its readiness to work with others towards the conclusion of a multilateral Fissile Material Cutoff Treaty.
3. Based on the commitments and actions mentioned above, as reiterated by India on September 5, 2008, and without prejudice to national positions thereon, Participating Governments have adopted and will implement the following policy on civil nuclear cooperation by Participating Governments with the IAEA-safeguarded Indian civil nuclear programme:
   a. Notwithstanding paragraphs 4(a), 4(b) and 4(c) of INFCIRC/254/Rev.9/Part 1, Participating Governments may transfer trigger list items and/or related technology to India for peaceful purposes and for use in IAEA safeguarded civil nuclear facilities, provided that the transfer satisfies all other provisions of INFCIRC/254/Part 1, as revised, and provided that transfers of sensitive exports remain subject to paragraphs 6 and 7 of the Guidelines.

b. Notwithstanding paragraphs 4(a) and 4(b) of INFCIRC/254/Rev.7/Part 2, Participating Governments may transfer nuclear-related dual-use equipment, materials, software, and related technology to India for peaceful purposes and for use in IAEA safeguarded civil nuclear facilities, provided that the transfer satisfies all other provisions of INFCIRC/254/Part 2, as revised.

c. At each Plenary, Participating Governments shall notify each other of approved transfers to India of Annex A and B items listed in INFCIRC/254/Part 1, as revised. Participating Governments are also invited to exchange information, including about their own bilateral agreements with India.

d. With a view to intensification of dialogue and cooperation with India, the Chairman is requested to confer and consult with India and keep the Plenary informed of these consultations.

e. Participating Governments will maintain contact and consult through regular channels, including the Consultative Group and Plenary, for the purpose of considering matters connected with the implementation of all aspects of this Statement taking into account relevant international commitments or bilateral agreements with India. In the event that one or more Participating Governments consider that circumstances have arisen which require consultations, Participating Governments will meet, and then act in accordance with paragraph 16 of the Guidelines.

4. In order to facilitate India's adherence to INFCIRC/254/Parts 1 and 2 and to remain current in its implementation of the Guidelines, the NSG Chair is requested to consult with India regarding changes to and implementation of the Guidelines and inform the Plenary of the outcome of the dialogue with India. Consultations with India regarding proposed amendments will facilitate their effective implementation by India.

5. Upon request by Participating Governments, the Chairman is requested to submit this Statement to the IAEA Director General with a request that it be circulated to all Member States.

Annexure 6

Statement by External Affairs Minister of India Shri Pranab Mukherjee on the Civil Nuclear Initiative

September 05, 2008

To reiterate India's stand on disarmament and nonproliferation, EAM has made the following statement:

A Plenary meeting of the Nuclear Suppliers Group to consider an exception for India from its guidelines to allow for full civil nuclear cooperation with India is being held in Vienna from September 4-5, 2008.

India has a long-standing and steadfast commitment to universal, non-discriminatory and total elimination of nuclear weapons. The vision of a world free of nuclear weapons which Shri Rajiv Gandhi put before the UN in 1988 still has universal resonance.

We approach our dialogue with the Nuclear Suppliers Group and all its members in a spirit of cooperation that allows for an ongoing frank exchange of views on subjects of mutual interest and concern. Such a dialogue will strengthen our relationship in the years to come.

Our civil nuclear initiative will strengthen the international non-proliferation regime. India believes that the opening of full civil nuclear cooperation will be good for India and for the world. It will have a profound positive impact on global energy security and international efforts to combat climate change.

India has recently submitted a Working Paper on Nuclear Disarmament to the UN General Assembly, containing initiatives on nuclear disarmament. These include the reaffirmation of the unequivocal commitment of all nuclear weapon States to the goal of complete elimination of nuclear weapons; negotiation of a Convention on the complete prohibition of the use or threat of use of nuclear weapons; and negotiation of a Nuclear Weapons Convention prohibiting the development, production, stockpiling and use of nuclear weapons and on their destruction, leading to the global, non-discriminatory and verifiable elimination of nuclear weapons within a specified timeframe.

We remain committed to a voluntary, unilateral moratorium on nuclear testing. We do not subscribe to any arms race, including a nuclear arms race. We have always tempered the exercise of our strategic autonomy with a sense of global responsibility. We affirm our policy of no-first-use of nuclear weapons.

We are committed to work with others towards the conclusion of a multilateral Fissile Material Cut-off Treaty in the Conference on Disarmament that is universal, non-discriminatory and verifiable.

India has an impeccable non-proliferation record. We have in place an effective and comprehensive system of national export controls, which has been constantly updated to meet the highest international standards. This is manifested in the enactment of the Weapons of Mass Destruction and their Delivery Systems Act in 2005. India has taken the necessary steps to secure nuclear materials and technology through comprehensive export control legislation and through harmonization and committing to adhere to Missile Technology Control Regime and Nuclear Suppliers Group guidelines.

India will not be the source of proliferation of sensitive technologies, including enrichment and reprocessing transfers. We stand for the strengthening of the non-proliferation regime. We support international efforts to limit the spread of ENR equipment or technologies to states that do not have them. We will work together with the international community to advance our common objective of non-proliferation. In this regard, India is interested in participating as a supplier nation, particularly for Thorium-based fuel and in establishment of international fuel banks, which also benefit India.

India places great value on the role played by the IAEA's nuclear safeguards system. We look forward to working with the IAEA in implementing the India-specific Safeguards Agreement concluded with the IAEA. In keeping with our commitment to sign and adhere to an Additional Protocol with respect to India's civil nuclear facilities, we are working closely with the IAEA to ensure early conclusion of an Additional Protocol to the Safeguards Agreement.

New Delhi

5th September 2008

Annexure 7

Milestones[32] leading to resumption of international civil nuclear trade

13 January 2004: The release of a matched set of statements (Next Steps in Strategic Partnership) by the PM of India in New Delhi and the President of the USA in Washington.

18 July 2005: Prime Minister Manmohan Singh and President George W. Bush sign a joint statement in Washington on India-US civil nuclear cooperation.

6 December 2005: India is admitted to the negotiations for finalising the Joint Implementation Agreement for setting up ITER.

2 March 2006: India and the US agree on a Separation Plan during President Bush's visit to Delhi. In exchange, US agrees to work towards full civil nuclear cooperation with India.

21 November 2006: India signs Joint Implementation Agreement to set up ITER.

9 Dec 2006: Both houses of US Congress approve the Henry J. Hyde US-India Peaceful Atomic Energy Cooperation Act of 2006.

18 Dec 2006: Bush signs the enabling legislation permitting full civilian nuclear cooperation with India into law.

27 July 2007: India and the US announce the finalization of bilateral Nuclear Cooperation Agreement.

October 2007: Left allies of the UPA government declare that the Nuclear Cooperation Agreement is against the country's interests and compromising to its sovereignty.

November 2007-June 2008: United Progressive Alliance (UPA) and its Left allies set up a joint panel and agree to allow the government to carry on negotiations with IAEA on finalising the text of ISSA.

May 2008: Indian negotiators complete negotiations with the secretariat of the IAEA for the text of the ISSA.

25 June 2008: The Left parties give an ultimatum to withdraw support if the government goes ahead with signing the ISSA with IAEA.

7 July 2008: While on way to Japan to attend the G8 summit, Prime Minister Manmohan Singh informs accompanying media persons about the plan of the government to go ahead with finalising the ISSA with IAEA.

9 July 2008: The Left parties withdraw support to the UPA government and call for a vote of confidence.

22 July 2008: The UPA government wins the vote of confidence.

1 Aug 2008: Board of Governors of IAEA approves ISSA.

6 Sep 2008: The 45-member Nuclear Suppliers Group (NSG) grants waiver to India facilitating international civil nuclear trade with India.

10 Sep 2008: President Bush notifies Congress to take up the India-US nuclear deal and give it the approval. Also, India's Foreign Secretary issues a letter to his counterpart regarding India intention to conclude agreements for the construction of nuclear power units at least at two sites capable of generation a minimum of 10,000 MWe.

23 Sep 2008: The US Senate Foreign Relations Committee approves the agreement but its new provisions stating that the deal will be subject to the Hyde Act create some uncertainty.

25 Sep 2008: President Bush tells Prime Minister Manmohan Singh in Washington that the administration is working to get the Nuclear Cooperation Agreement approved by the US Congress.

26 Sep 2008: US House of Representatives Foreign Affairs Committee chairman Howard Berman introduces the approval resolution to Agreement.

27 Sep 2008: House of Representatives passes the 123 Agreement for civil nuclear cooperation between India and the US by a margin of 298-117.

30 September 2008: India and France sign the Nuclear Cooperation Agreement.

1 Oct 2008: Senate passes the approval bill 86-13.

10 October 2008: India and the USA sign the Nuclear Cooperation Agreement.

5 Dec 2008: India and Russia sign the Nuclear Cooperation Agreement.

30 Jul 2010: India and the USA sign Arrangements and Procedures for Reprocessing

15 May 2009: India signs the Additional Protocol

16 Oct 2009: India notifies first set of facilities to IAEA for implementing safeguards under ISSA

11 Nov 2011: India notifies the Civil Liability for Nuclear Damage Act, 2010.

12 Dec 2014: India completes separation plan by offering last set of facilities identified in the Separation Plan as per the agreed time table.

4 Feb 2016: India deposits its instrument of ratification for the Convention on Supplementary Compensation for Nuclear Damage (CSC).

Science Diplomacy Case Study Series

Forum for Indian Science Diplomacy

As part of its ongoing research studies on Science &Technology and Innovation (STI), RIS together with the National Institute of Advanced Studies (NIAS), Bangalore has endeavoured a major project for Science Diplomacy this year, supported by the Department of Science and Technology. The programme was launched on 7 May 2018 at New Delhi. Forum for Indian Science Diplomacy (FISD), under the RIS–NIAS Science Diplomacy Programme envisages harnessing science diplomacy in areas of critical importance for national development and S&T cooperation.

The key objective of the Forum is to realize the potential of Science Diplomacy by various means, including Capacity building in science diplomacy, developing networks and Science diplomacy for strategic thinking. It aims for leveraging the strengths and expertise of Indian Diaspora working in the field of S&T to help the nation meet its agenda in some select S&T sectors.

OTHER FISD PUBLICATIONS

Case Studies

- A New Lexicon of Science Diplomacy *by V. Siddhartha*
- Building Capacity in Electronics Sector and Science Diplomacy A Case Study of Semiconductor Complex Ltd (SCL) *by Ashok Parthasarathi*

Science Diplomacy Review

Vol. 1; No. 1 to 3

Science Diplomacy News Alerts

Issues 1 to 22

RIS
Research and Information System for Developing Countries
विकासशील देशों की अनुसंधान एवं सूचना प्रणाली

Research and Information System for Developing Countries (RIS) is a New Delhi-based autonomous policy research institute that specialises in issues related to international economic development, trade, investment and technology. RIS is envisioned as a forum for fostering effective policy dialogue and capacity-building among developing countries on global and regional economic issues.

Core IV-B, Fourth Floor, India Habitat Centre,
Lodhi Road, New Delhi-110 003, India, Ph.: +91-11-24682177-80, Fax: +91-11-24682173-74
E-mail: dgoffice@ris.org.in, Website: www.ris.org.in; www.fisd.in

Follow us on:

/FISD-Forum-for-Indian-Science-Diplomacy

@fisd19

/RISNewDelhi

MultiCraft

International Journal of Engineering, Science and Technology
Vol. 11, No. 4, 2019, pp. 49-57

INTERNATIONAL JOURNAL OF ENGINEERING, SCIENCE AND TECHNOLOGY

www.ijest-ng.com
www.ajol.info/index.php/ijest
ðÓ 2019 MultiCraft Limited. All rights reserved

Short-term PV power forecasting based on sky conditions using intelligent modelling techniques

G. Perveen[1]*, M. Rizwan[2], N. Goel[3]

[1*]Department of Electronics & Communication Engineering, Delhi Technological University, Delhi-110042, INDIA
[2]Department of Electrical Engineering, Delhi Technological University, Delhi-110042, INDIA
[3]Department of Electronics & Communication Engineering, IGDTUW, Delhi-110006, INDIA
*Corresponding Author: e-mail: gulegulnar@gmail.com, Tel +91-9891788761

Abstract

The work in this paper involves intelligent modelling techniques i.e. fuzzy logic, Artificial Neural Network (ANN) Adaptive Neural Fuzzy Inference System (ANFIS) methodologies for estimating the power in a solar photovoltai system. Since, the generation of power is subjective to environmental factors such as ambient temperature, variatio conditions and solar insolation, therefore, an intelligent modelling techniques have been proposed for forecasting the p solar photovoltaic system employing 210 W Heterojunction with Intrinsic Thin layer (HIT) photovoltaic modules for diff sky-conditions such as clear sky, hazy sky, partially foggy/cloudy sky and fully foggy/cloudy sky conditions respecti composite climate zone and performance has been evaluated using statistical indicators

Keywords: ANFIS (adaptive neural fuzzy inference system); ANN (artificial neural network); forecasting; fuzzy logic; s condition; SPV system

DOI: http://dx.doi.org/10.4314/ijest.v11i4.5

1. Introduction

Last few years have seen tremendous growth in the field of renewable power generation especially in the field of solar energ which employs PV system comprising a number of solar cells (Chauhan and Saini, 2004; Singh, 2003; Zhou et al., 2010). Its advantage is that it generates green house gas emissions and simple scalability in terms of power needs while its disadvantage is that the output power diminishes due to dust, clouds and other obstructions in the atmosphere (Al-Lawati et al., 2015; Jimenez-Perez and Mora-Lopez, 2016). Therefore, an intelligent modelling technique has been introduced to accurately estimate the power generation in SPV system based on sky conditions.

Detailed literature analysis reveals that the intelligent model available in the literature have been defined for clear sky-condition, however, very few models are available in the literature that discussed about variation in sky condition. The objective of the present research is to do the comparative analysis of Fuzzy logic, ANN and ANFIS models in forecasting short-term PV power of SPV systems employing HIT PV modules for different sky conditions i.e. sunny sky, hazy sky, partially foggy/cloudy sky and fully foggy/cloudy sky conditions and for composite climate zone. The performance has been evaluated based on statistica indicators.

The work has been arranged as follows: The data and methodology are presented in Section 2. Section 3 discusses the performance characterization of SPV system. Section 4 is developed to statistical error tests. Results and discussions are prese in Section 5. Conclusion followed by references is presented in Section 6.

2. Data and Methodology

2.1 Meteorological data: In this work, the parameters namely solar irradiance and cell temperature have been achieved from the collaboration of Indian Meteorological Department and National Institute of Solar Energy for composite climate zone (Ajit, 2009; Bansal and Minke, 1988). Further, the normalization of the parameters in 0.9-1.1 range has been done to avoid convergence issues.

2.2 Intelligent modelling for estimating the power of SPV system: Intelligent modelling methods namely fuzzy logic, ANN and ANFIS have been employed to forecast the power in a solar photovoltaic system (Perveen, 2019). The details of methodologies are discussed section-wise as follows:

2.2.1 Fuzzy logic: The first section discusses the fuzzy logic modelling where the membership functions are defined in 0.1-0.9 range and a set of rules are explained in fuzzy logic toolbox of MATLAB. The model has been developed using Mamdani-Sugeno in fuzzy inference system to forecast the power of SPV system for input parameters namely solar cell temperature and solar insolation while the PV power is the output and presented in Fig. 1 (Chen, 2013; Rizwan et al., 2014; Perveen et al., 2018).

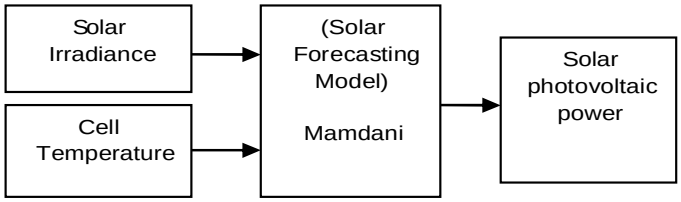


Figure 1. Fuzzy logic based model for estimating power in a solar photovoltaic system.

2.2.2 Artificial Neural Network (ANN): The second section presents an artificial intelligent modelling technique artificial Neural Network (ANN) employing a feed-forward neural network. Model-based on ANN are planned and designed in such manner that output is simulated from the input variables for power forecasting in the SPV system and are presented in Fig. 2 (Azadeh et al., 2013; Azimi et al., 2016; Al-Waeli et al., 2019; Notton et al., 2019).

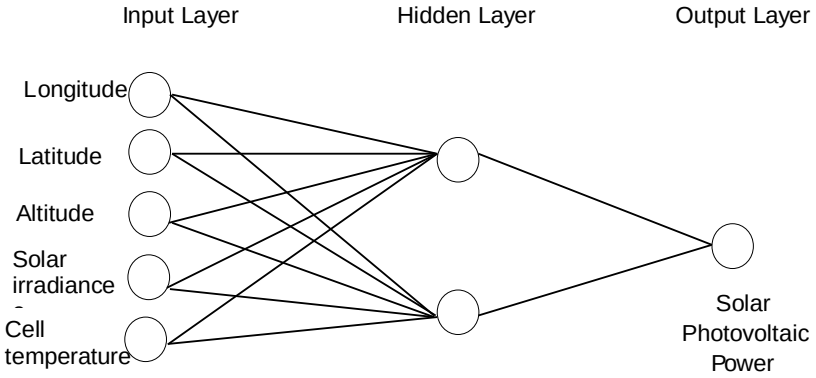


Figure 2. ANN architecture for estimating power in the SPV system.

2.2.3 Adaptive Neural Fuzzy Inference System (ANFIS): The third section presents the combination of ANN and fuzzy logic modelling. ANFIS applies the graphical analysis of Fuzzy-Sugeno system lying within the framework of adaptive networks and relates back-propagation and least squares method using MATLAB software for training and testing the data. The main advantage of using ANFIS methodology is that the convergence rate is much faster (Abbod et al., 2013; Jang and Sun, 1995).

3. Performance Characterization of SPV System

In the present work, 210 W HIT solar PV module is chosen and operated at Maximum Power Point (MPP) conditions. Influenced by parameters namely solar irradiance and cell temperature and based on the standard test condition, the power generation can be defined by Eq. (1) as shown below (Riffonneau et al., 2011):

$$Q_{PV} = \quad * \qquad (1)$$

and - 20) (2)

where $Q_{PV,STC}$ is the photovoltaic system rated power output at standard test conditions, Q_{PV} is the power output of photovoltaic array at MPP, G_T is solar irradiance in W/m^2 at standard test condition, N_{PVS} is the photovoltaic arrays in series in number, is a temperature parameter at maximum power point, N_{PVP} is photovoltaic arrays in parallel in number of, T_j is the junction temperature of solar panel in °C. Given the efficiency of the module is 16.7%, T_m is ambient operating temperature range between - 4°F to 115°F and N_{OCT} (Normal operating cell temperature) is 114.8°F.

4. Statistical Error Tests

For evaluating the performance of the models, statistical error tests namely MAPE, NMAE, and nRMSE respectively have been used for analysis (Dolara et al., 2018).

4.1 Mean absolute percentage error (MAPE):

$$MAPE_{\%} = \qquad (3)$$

4.2 Normalized mean absolute error (NMAE):

$$NMAE_{\%} = \qquad (4)$$

4.3 Normalized root mean square error (nRMSE):

$$nRMSE_{\%} = \qquad (5)$$

where n is the number of observed data, a_i and e_i are the i_{th} measured and estimated data respectively, and $E_i = e_i - a_i$ is the absolute error. This definition of error is normalized over the maximum hourly measured data.

5. Results and Discussions

Intelligent models have been developed for SPV system which uses 210 W HIT solar PV module operated at Maximum Power Point Tracking (MPPT) conditions. The data at the input layer comprises cell temperature, solar irradiance and the output layer is the PV power. Fuzzy logic, artificial neural network and ANFIS methodologies have been employed in forecasting power of SPV systems for composite climate zone and are presented in Table 1.

It has been observed from Table 1 that the average mean absolute percentage error by using fuzzy logic methodology is 0.10 with ANN methodology the mean absolute percentage error reduced to 0.04%; and with ANFIS methodology the mean absolute percentage error further reduced to 0.01% which reveals that the obtained results are precise and accurate. This is due to reason that it employs both fuzzy logic approach and artificial neural network. Further, the graphical representation between the measured and forecasted power employing different methodologies has been presented in Fig. 3.

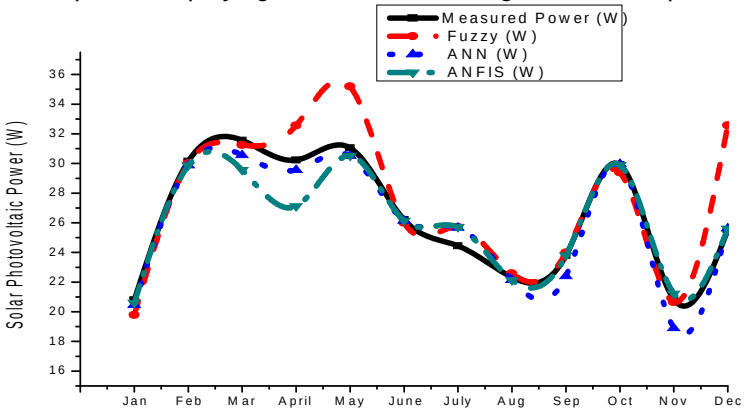


Figure 3. Graphical analysis of measured and forecasted PV power employing fuzzy logic, ANN and ANFIS methodologies.

Table 1. Intelligent methodologies for forecasting power in a solar PV system employing HIT solar PV modules under composite climate zone

Month	Solar Irradiance (W/m^2)	V_{oc} (V)	I_{sc} (A)	Cell Temp (^0C)	Measured Power (W)	Fuzzy				ANN				ANFIS			
						Forecasted Power (W)	MAPE (%)	NMAE (%)	nRMSE (%)	Forecasted Power (W)	MAPE (%)	NMAE (%)	nRMSE (%)	Forecasted Power (W)	MAPE (%)	NMAE (%)	nRMSE (%)
Jan	361.15	81.47	0.42	28.14	20.80	19.79	0.18	0.09	0.11	20.46	0.01	0.01	0.02	20.61	0.00	0.00	0.00
Feb	461.52	82.03	0.61	35.69	30.15	30.02	0.07	0.05	0.07	29.87	0.00	0.00	0.01	29.87	0.00	0.00	0.00
Mar	548.24	83.64	0.62	40.12	31.56	31.25	0.01	0.03	0.03	30.56	0.01	0.02	0.00	29.56	0.00	0.00	0.00
April	575.12	79.62	0.60	42.53	30.23	32.56	0.15	0.06	0.04	29.56	0.00	0.02	0.00	27.12	0.00	0.00	0.00
May	559.67	77.35	0.59	46.22	31.05	35.19	0.12	0.05	0.06	30.52	0.00	0.00	0.00	30.53	0.00	0.00	0.00
June	537.17	76.92	0.55	45.24	26.20	26.00	0.07	0.06	0.07	26.14	0.00	0.00	0.00	26.14	0.01	0.00	0.01
July	537.81	76.66	0.05	48.28	24.45	25.72	0.05	0.04	0.05	25.67	0.02	0.02	0.03	25.71	0.00	0.00	0.00
Aug	428.80	76.90	0.47	53.27	22.33	22.59	0.06	0.04	0.05	22.14	0.03	0.01	0.04	22.13	0.00	0.00	0.00
Sep	437.51	77.78	0.50	51.62	23.80	23.99	0.13	0.07	0.09	22.40	0.27	0.12	0.19	23.80	0.07	0.03	0.04
Oct	466.57	78.74	0.63	53.21	29.92	29.40	0.07	0.06	0.07	29.97	0.01	0.00	0.01	29.90	0.00	0.00	0.00
Nov	369.82	78.48	0.44	45.08	20.85	20.65	0.13	0.05	0.06	18.90	0.03	0.01	0.02	21.23	0.00	0.00	0.00
Dec	370.84	80.66	0.47	43.36	25.53	32.57	0.12	0.01	0.09	25.62	0.10	0.08	0.43	25.61	0.00	0.00	0.01
Avg.	471.19	79.19	0.50	44.40	26.41	27.48	0.10	0.05	0.07	25.98	0.04	0.03	0.06	26.02	0.01	0.00	0.01

From Figure 3, it has been revealed that the forecasted power employing ANFIS methodology closely follows the measured power as compared to others methods such as fuzzy logic and artificial neural network. Furthermore, the power generation in SPV system is greatly influenced by external environmental factors namely sky conditions, ambient temperature and solar insolation. Therefore, sunny, hazy, partially foggy/cloudy and fully foggy/cloudy sky conditions are considered and performance has been evaluated based on statistical error tests and are presented in Table 2.

Table 2. Short-term PV power forecasting employing HIT solar PV module under composite climate zone

Sky conditions	Time (hr)	Cell Temp. (^oC)	Solar Irradiance (W/m^2)	Measured Power (W)	Fuzzy		ANN		ANFIS	
					Forecasted Power (W)	MAPE (%)	Forecasted Power (W)	MAPE (%)	Forecasted Power (W)	MAPE (%)
Sunny sky	7:00	34.11	140.93	7.00	7.02	0.0900	7.02	0.0027	7.03	0.0036
	8:00	40.55	273.76	13.83	15.00	0.0840	13.84	0.0015	12.85	0.0001
	9:00	48.14	486.12	28.00	30.34	0.0830	28.02	0.0009	28.01	0.0004
	10:00	52.23	625.25	39.83	41.66	0.0460	39.82	0.0002	39.83	0.0002
	11:00	57.79	783.13	52.33	54.17	0.0350	52.31	0.0005	52.34	0.0002
	12:00	62.86	875.34	59.33	60.50	0.0200	59.35	0.0003	59.06	0.0040
	13:00	64.69	888.95	62.00	61.17	0.0900	61.88	0.0018	61.99	0.0005
	14:00	63.91	744.96	43.17	42.00	0.0270	43.16	0.0000	43.46	0.0052
	15:00	61.55	726.73	48.33	46.32	0.0420	48.32	0.0001	48.33	0.0007
	16:00	58.04	549.15	34.33	31.51	0.0820	34.34	0.0004	34.35	0.0006
	17:00	52.72	361.6	20.83	19.32	0.0720	20.79	0.0024	20.82	0.0004
	Avg.	54.24	586.90	37.18	37.18	0.0610	37.17	0.0014	37.10	0.0010
Hazy sky	10:00	40.83	123.1	8.00	7.12	0.0100	8.39	0.0492	8.03	0.0031
	11:00	44.49	146.12	8.50	9.70	0.0020	9.45	0.1408	8.49	0.0011
	12:00	43.56	307.56	19.67	24.20	0.2300	20.25	0.0123	19.66	0.0004
	13:00	52.55	519.54	45.00	45.46	0.0100	44.82	0.0031	44.98	0.0003
	14:00	42.4	467.65	39.00	37.40	0.0410	39.25	0.0060	39.01	0.0001
	15:00	49.36	313.06	24.50	21.50	0.0170	24.54	0.0012	24.51	0.0004
	16:00	41.05	185.35	11.00	10.09	0.0830	11.68	0.0828	11.00	0.0001
	Avg.	44.89	294.63	22.24	22.21	0.0561	22.63	0.0422	22.24	0.0008
Partially foggy/ cloudy sky	8:00	45.79	134.08	9.67	10.26	0.0610	9.72	0.0045	9.66	0.0008
	9:00	47.49	179.69	27.33	12.99	0.5250	27.32	0.0038	12.36	0.0670
	10:00	52.07	355.98	32.50	30.14	0.0730	33.33	0.0125	27.53	0.0031
	11:00	55.57	463.45	44.00	40.78	0.0730	44.01	0.0010	32.56	0.0108
	12:00	58.38	547.32	39.50	44.39	0.1240	39.24	0.0013	43.92	0.0005
	13:00	59.96	519.74	39.83	31.40	0.2120	39.95	0.0130	39.49	0.0026
	14:00	55.52	492.69	50.17	40.93	0.1840	50.28	0.0024	39.78	0.0009
	15:00	61.3	647.1	37.33	51.69	0.3850	37.60	0.0178	50.22	0.0010
	16:00	59.64	562.02	21.17	35.16	0.6610	21.20	0.0114	37.59	0.0139
	17:00	50.88	299.99	15.33	18.22	0.1880	15.39	0.0034	20.88	0.0096
	18:00	50.73	235.3	7.33	12.65	0.7250	7.53	0.0473	15.30	0.0076
	19:00	47.99	156.43	7.33	7.23	0.0130	7.75	0.0858	8.81	0.2480
	Avg.	53.78	382.82	27.62	27.99	0.2687	27.78	0.0305	28.18	0.0170
Fully foggy/ cloudy sky	9:00	19.08	170.77	12.50	10.75	0.1400	12.60	0.0066	12.48	0.0011
	10:00	19.02	74.87	8.33	10.62	0.2750	8.35	0.0030	8.40	0.0099
	11:00	23.29	96.49	11.17	8.72	0.2190	12.78	0.1733	10.98	0.0442
	12:00	18.5	41.2	7.83	8.10	0.0340	7.82	0.0018	7.81	0.0036
	13:00	18.57	140.77	10.67	10.46	0.0200	10.66	0.0002	10.63	0.0056
	14:00	18.59	87.31	9.83	11.14	0.1330	10.11	0.0243	10.01	0.0735
	15:00	17.86	164.25	12.00	10.83	0.0970	12.01	0.0016	12.00	0.0004
	Avg.	19.27	110.81	10.33	10.09	0.1311	10.62	0.0301	10.33	0.0198

Following inferences can be drawn from Table 2 shown as:

Sunny/Clear sky: 1st June 2015 is considered as a sunny day based on the annual analysis of solar radiation data and availability of sunshine hours. Further, the graphical representation between measured and forecasted power for sunny sky condition has been shown in Figure 4 from which it has been concluded that the forecasted power employing ANFIS methodology represented by cyan colored dash-dot line on hour basis closely follows the measured power represented by the black solid line, whereas some deviation can be seen in terms of fuzzy logic approach represented by red colored dashed line and ANN approach represented by blue colored dot line. The maximum power output is observed to be 62 W during the day with average MAPE of 0.0610% using the fuzzy logic methodology; and the error reduced to 0.0014% by employing ANN methodology; and with ANFIS methodology the mean absolute percentage error has further reduced and is observed to be 0.0010% respectively.

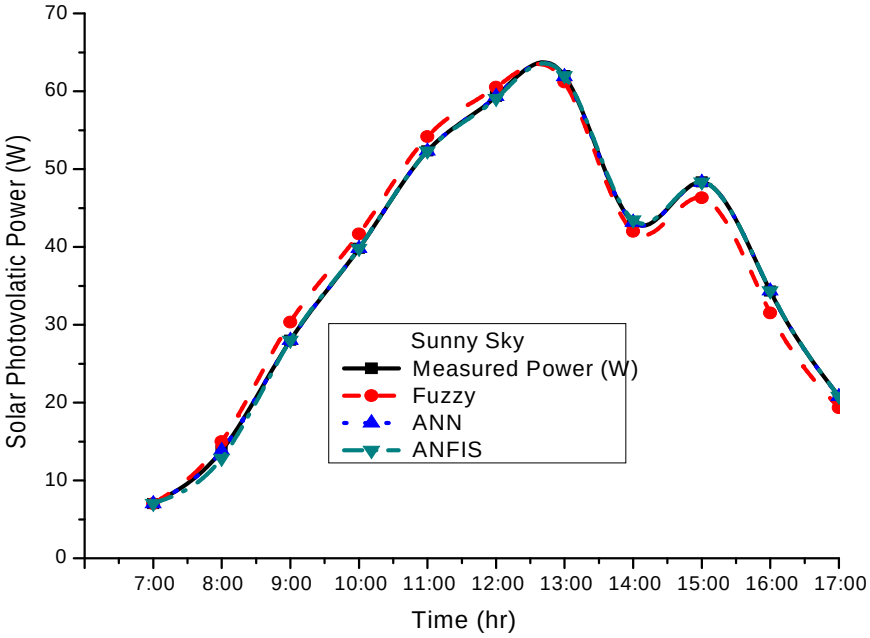


Figure 4. Graphical analysis of measured power and forecasted power employing fuzzy logic, ANN and ANFIS methodologies for sunny sky condition.

Hazy sky: 26th December 2015 is considered as a hazy day based on the annual analysis of solar radiation data and availability of sunshine hours. Further, the graphical representation between measured and forecasted power for hazy sky condition has been shown in Figure 5 from which it has been concluded that the forecasted power employing ANFIS methodology represented by cyan dash-dot line on hour basis closely follows the measured power represented by the black solid line, whereas some deviation can be seen in terms of fuzzy logic approach represented by red dashed line and ANN represented by blue dot line. The maximum power output is observed to be 50 W during the day with averaged mean absolute percentage error by employing fuzzy logic methodology is 0.0561%; by employing ANN methodology the mean absolute percentage error is 0.0422%; and by employing ANFIS methodology the error is 0.0068% respectively.

Partially foggy/cloudy sky: 3rd August 2015 is considered as partially foggy/cloudy day based on the annual analysis of solar radiation data and availability of sunshine hours. Further, the graphical representation between measured and forecasted power for partially foggy/cloudy sky condition have been shown in Figure 6 from which it has been concluded that the forecasted power employing ANFIS methodology represented by cyan dash-dot line on hour basis closely follows the measured power represented by the black solid line, whereas some deviation can be seen in terms of fuzzy logic approach represented by red dashed line and ANN represented by blue dot line. The maximum power output is observed to be 44 W during the day with averaged mean percentage error by employing fuzzy logic methodology is 0.2687%; by employing ANN methodology the mean absolute percentage error is 0.0305%; and by employing ANFIS methodology the error is 0.0170%.

Fully foggy/cloudy sky: 3rd January 2015 is considered as fully foggy/cloudy day based on the annual analysis of solar radiation data and availability of sunshine hours. Further, the graphical representation between measured and forecasted power for fully foggy/cloudy sky condition have been shown in Figure 7 from which it has been concluded that the forecasted power employing ANFIS methodology represented by cyan dash-dot line on hour basis closely follows the measured power represented by the black solid line, whereas some deviation can be seen in terms of fuzzy approach represented by red

dashed line and ANN represented by the dot line. The maximum power output is observed to be 12 W during the day with averaged mean absolute percentage error by employing fuzzy logic methodology is 0.1311%; by using ANN methodology the error is 0.0301%; and by employing ANFIS methodology the mean absolute percentage error is 0.0198% respectively.

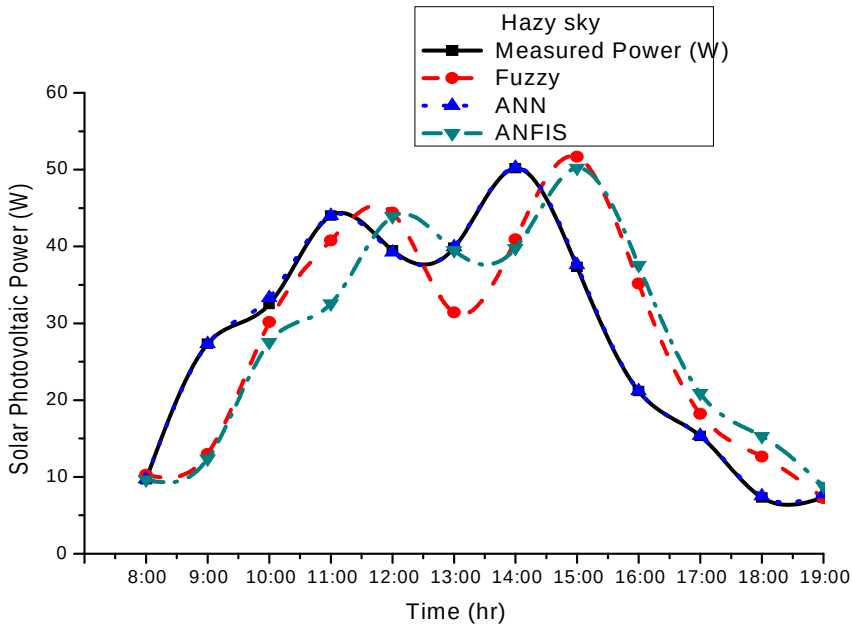


Figure 5. Graphical analysis of measured power and forecasted power employing fuzzy logic, ANN and ANFIS methodologies for hazy sky condition.

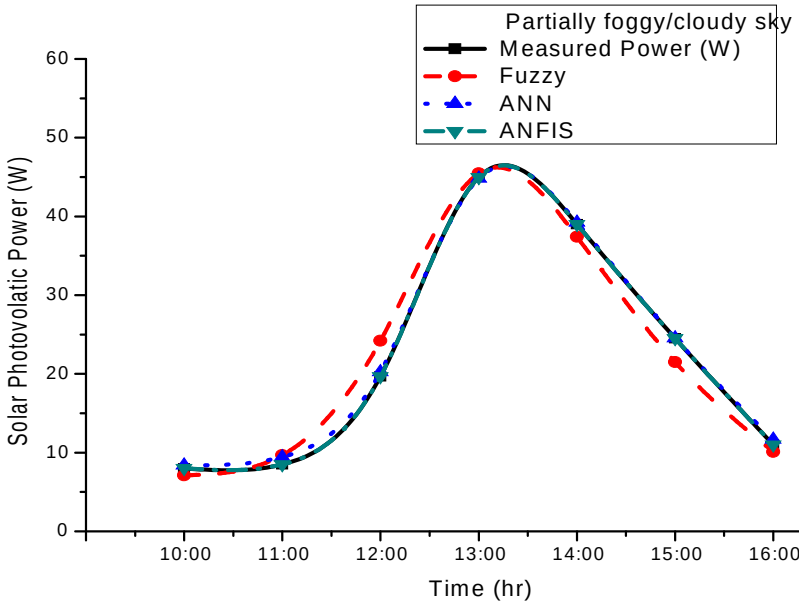


Figure 6.Graphical analysis of measured power and forecasted power employing fuzzy logic, ANN and ANFIS methodologies for partially foggy/cloudy sky condition.

From Table 2, it has been observed that out of the four models, especially the hazy sky model perform well with mean absolute mean percentage error of 0.0008% in forecasting the power of SPV system followed by sunny sky model with mean absolute percentage error of 0.0010%, partially foggy/cloudy sky model with mean absolute percentage error of 0.0170%, and fully foggy/cloudy sky model with mean absolute percentage error of 0.0198% respectively by using ANFIS methodology.

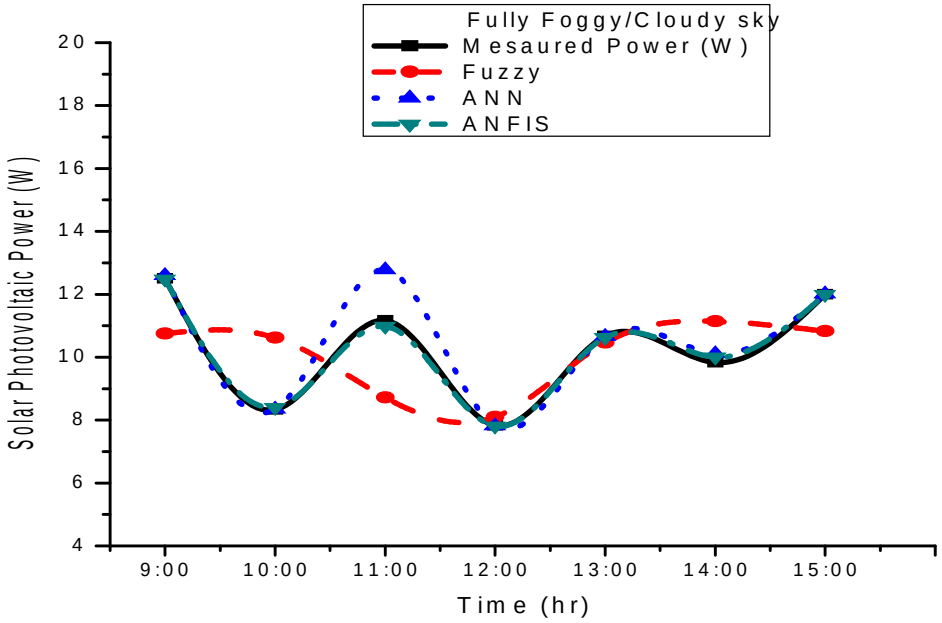


Figure 7. Graphical analysis of measured power and forecasted power employing fuzzy logic, ANN and ANFIS methodologies for fully foggy/cloudy sky condition.

6. Conclusion

In the present work, intelligent modelling techniques have been employed for estimating the power of a SPV system for composite climate zone. Simulations have been carried out using fuzzy logic, ANN and ANFIS methodologies for varying sky-conditions i.e. sunny sky, hazy sky, partially foggy/cloudy sky and fully foggy/cloudy sky conditions. Three criteria namely mean absolute percentage error, normalized mean absolute error and normalized root mean square error verifies the forecasting errors of the model. Obtained results reveal the accuracy of the ANFIS model which is far accurate and better than the artificial neural network and fuzzy logic models. ANFIS model has certain advantages such as the ease of design, robustness and adaptability with the non-linearity associated with the data. The ANFIS methodology integrates the features of both fuzzy logic and ANN which increases the system accuracy and makes the system response much faster. The result reveals the model implementation for a broad series of applications. The prediction of solar energy makes it suitable for installation of a monitoring station for a remote place and furthermore, can be extended for sizing of standalone PV system as a part of the future work.

Nomenclature

e_i	i_{th} estimated data (dimensionless)
MAPE	mean absolute percentage error (%)
m_i	i_{th} measured data (dimensionless)
nMAE	normalized mean absolute error (%)
nRMSE	normalized root mean square error (%)
N_{PVP}	photovoltaic arrays in parallel (number)
N_{PVS}	photovoltaic arrays in series (number)
O_{PV}	power output of photovoltaic array at MPP (W)
$O_{PV,STC}$	rated power output of photovoltaic array (W)
T_m	measured ambient temperature (^{0}C)
T_j	temperature of solar panel (^{0}C)
T_o	maximum ambient temperature (^{0}C)
	temperature parameter at MPP

References

Abu-rub H., Iqbal A., Ahmed M., Peng. Z., Li Y., and Baoming G. 2013. Quasi-Z-Source Inverter based Photovoltaic generation system with maximum power tracking control using ANFIS. IEEE Transactions on Sustainable Energy, Vol. 4, No. 1, pp. 11-20.

Ajit, P. T. 2009. Solar radiant energy over India, Ministry of New and Renewable Energy and India Meteorological Department.

Al-Waeli A. H. A., Sopian K., Yousif J. H., Kazem H. A., Boland J., and Chaichan M. T. 2019. Artificial neural network modeling and analysis of photovoltaic/thermal system based on the experimental study, Energy conversion and Management, Vol. 186, pp. 368-379.

Azadeh A., Babazadeh R. and Asadzadeh M. S. 2013. Optimum estimation and forecasting of renewable energy consumption by artificial neural networks, Renewable and Sustainable Energy Reviews, Vol. 27, pp. 605-612.

Azimi R., Ghayekhloo M. and Ghofrani M. 2016. A hybrid method based on a new clustering technique and multilayer perceptron neural network for hourly solar radiation forecasting, Energy Conversion and Management, Vol. 118, pp. 331-344.

Bansal N. K. and Minke G. 1988. Climatic zones and rural housing in India, Kernforschungsanlage Julich GmbH Zentralbibliothek, Julich.

Chauhan A. and Saini R. P. 2014. A review on Integrated Renewable Energy system based power generation for stand-alone applications: Configurations, storage options, sizing methodologies, and control, Renewable and Sustainable Energy Reviews, Vol. 3, pp. 99-120.

Chen S. X., Gooi H. B. and Wang M. Q. 2013. Solar radiation forecast based on fuzzy logic and neural networks. Renewable Energy, Vol. 60, No. (c), pp. 195-201.

Dolara, A., Grimacci, F., Lev, S., Mussett, M. and Ogliari E. 2018. Comparison of training approaches for photovoltaic forecasts by means of machine learning, Applied Sciences, Vol. 8, pp. 1-16.

Jang J. R. and Sun C. 1995. Neuro-fuzzy modelling and control, Proceedings of the IEEE, Vol. 83, No. 3, pp. 378-406.

Jang R. 1993. ANFIS: Adaptive-Network-Based Fuzzy Inference System, IEEE Transactions on System, MAN and Cybernetics, Vol. 23, No. 3, pp. 665-685.

Jimenez-Perez P. F. and Mora-Lopez L. 2016. Modeling and forecasting hourly global solar radiation using clustering and classification techniques, Solar Energy, Vol. 135, pp. 682-691.

Liu J., Fang W., Zhang X. and Yang C., 2015. An improved photovoltaic power forecasting model with the assistance of Aerosol Index data, IEEE Transaction on Sustainable Energy, Vol. 6, No. 2, pp. 434-442.

Notton B., Fouilloy A., Duchaud J. L., Voyant C. and Dizene R. 2019. Solar radiation forecasting using artificial neural network and random forest methods: application to normal beam, horizontal diffuse and global components, Renewable Energy, Vol. 132, pp. 871-884.

Perveen G., Rizwan M. and Goel N. 2019. Comparison of intelligent modelling techniques for solar energy forecasting and its application in solar PV systems, IET Energy Systems Integration, Vol. 1, No. 1, pp. 3-51.

Perveen G., Rizwan M. and Goel N. 2018. Intelligent model for solar energy forecasting and its implementation for solar photovoltaic applications, Journal of Renewable and Sustainable Energy, Vol. 10, No. 6, Article ID 063702, pp. 1-23.

Riffonneau Y., Bacha S., Barruel F. and Ploix, S. 2011. Optimal power flow management for grid-connected PV systems with batteries, IEEE Transaction of Sustainable Energy, Vol. 2, No. 3, pp. 309-320.

Rizwan M., Jamil M., Kirmani S. and Kothari D. P. 2014. Fuzzy logic based modelling and estimation of global solar radiation using meteorological parameters, Energy, Vol. 70, pp. 685-691.

Singh G. K. 2013. Solar power generation by PV (photovoltaic) technology: A review, Energy, Vol. 53, pp. 1-13.

Zhou W., Lou C., Li Z., Lu L., and Yang H. 2010. Current status of research on optimum sizing of stand-alone hybrid solar-wind power generation systems, Applied Energy, Vol. 87, No. 2, pp. 380-389.

Bibliographic Notes

Dr. Gulnar Perveen is presently associated with research and development activities in Defence Research and Development Organization (DRDO), M/o Defence, India. Prior to joining DRDO she has served National Institute of Solar Energy (NISE), M/o New and Renewable Energy, Delhi, India wherein she was involved in R&D of solar PV systems and it applications. She has obtained her Ph.D. from Delhi Technological University, Delhi, India. She has a research, teaching and industrial experience of 9 years wherein she has published more than 15 research papers in reputed journal and conference proceedings.

Dr. M. Rizwan is presently working as an Associate Professor in the Department of Electrical Engineering, Delhi Technological University (DTU), Delhi, India. He has more than 16 years of teaching and research experience wherein he has published more than 95 research papers in reputed journals and conference proceedings. He is the recipient of the Raman Fellowships for Post-Doctoral research for Indian scholars in the USA, DST Start-Up Grants (Young Scientists). His areas of interest include power system engineering, renewable energy systems particularly solar PV and soft computing applications. He is a senior member of IEEE, life member of ISTE, life member of SSI, member in International Association of Engineers (IAENG) and many other reputed societies.

Dr. Nidhi Goel is working as an Associate Professor in the Department of Electronics & Communication Engineering at Indira Gandhi Delhi Technical University (IGDTUW), Delhi, India. Prior to joining IGDTUW, she was associated with DTU, Delhi, India. She has teaching and research experience of approximately 17 years wherein she has authored 2 book chapters and published over 45 research papers in reputed journals and conferences proceedings. Her area of interest includes Embedded Systems, Biomedical Image Processing, Renewable Energy and Multimedia Security.

Received July 2019
Accepted September 2019
Final acceptance in revised form September 2019

Sign Language Interpretation and Conversion to Text

Mrs. Pallavi Ghaste, Mrs.Soniya Bastwade, Rajvi Khandelwal, Shubhangi Ambapkar, Ziaurrehman Ansari and Chirag Singh Salaria.
Department of Computer Engineering
DYPCOE Akurdi, Pune

ABSTRACT

In our day to day life, communication plays important role for conveying information from one person to another person. But it becomes very difficult for the people who are deaf and dumb to communicate with normal people. Sign language is the only one way to communicate with them. But normal people are unaware of sign language. So there is only one way and that is to covert sign language into text & speech & vice versa. That is known as sign recognition. Sign language is a combination of body languages, hand gestures and facial expressions. Among those hand gestures are provides majority of the information

In this paper we will discuss the method proposed by author for Sign Language Interpretation and its conversion to text.

Index Terms- Pattern Recognition, Media pipe Framework, Image Pre-Processing, Feature Extraction.

I. INTRODUCTION

Over 5% of the total population on Planet are deaf people. 5% may seem like a small number, but that totals over 360 million-! Being specially abled of this type, it is very difficult to communicate with ordinary people. To help them communicate among themselves and others, there is a sign language through which they communicate. However, ordinary people find it hard to learn this language and hence, interaction with specially abled people is difficult.

To overcome this problem, several methods had been proposed. However, some are expensive and others rather inaccurate. Here we propose an accurate, affordable and easily accessible method to solve the problem. In the proposed method, we make use of the image recognition and deep learning, to make machine act as a translator. Through camera image of the sign alphabet will be captured and the interface (machine) will output the sign alphabet in text.

II. LITERATURE REVIEW ON DIFFERENT SYSTEMS.

The hand gesture recognition system builds a natural, interactive and the most favorable environment for the Human-Computer Interaction. Due to its utmost importance in the present scenario, and more importantlyin the future, many techniques have been thought about, developed and implemented till now. Mostly sign language recognition is done by two approaches: Image based [1].sensor-based approach.

Truong*et al.* [2] presents a work on applying AdaBoost and Haarlike classifiers for recognizing American Sign Language with complex background from live videos and translates into text and speech. The key factor of the high accuracy result was the big dataset in the training process.

Fu*et al.* [3] proposed a wavelet enhanced image preprocessing and neural networks for hand gesture recognition approach. Collected images are processed to extract the contour of the hand gestures also provide the neural networks for hand gesture learning with wavelet-based feature extraction.

Since many proposed methods does not works on geometric properties and require any extracted features. Vishwakarma*et al.* [4] implemented method which works on geometric properties and does not require any extracted features also works well on low-resolution images which is a great advantage for developers regarding data storage.

Vishwakarma et al. [5] presents a novel hand gesture recognition system, for the efficient human-computer interaction. Also proposed algorithm for hand gesture recognition in the complex background in proposed based on shape and texture feature.

Kumar et al. [6] proposed an idea for feasible communication between hearing impaired and normal person with the help of machine learning approach. Instead of preprocessing techniques, like filtering and segmentation of hand patch images, features are extracted from dataset by discrete wavelet transform.

Anup Kumar et al. [7] propose the design for a basic yet extensible system that is able to recognize static and dynamic gestures of American Sign Language. Mapping of the spoken alphabet to text is done by proposed algorithm.

III. THE PROPOSED SYSTEM

Communication plays a major role in human life, but deaf and dumb people find it difficult to communicate with normal people. To minimize this communication gap, we introduce a handgesture recognition system to recognize the alphabets of Sign Language. We propose an easy-to-use and inexpensive approach to recognize single and double handed gestures accurately. We intend to develop the said project in total of three modules namely:

1. Image Input Module:
This Module will consist of capturing the image of hand to recognize the gesture within the given resolution from camera.

2. Image Processing Module:
This module will give 21 3D key points of the given image of hand. It will be essentially a hand tracking module implemented using Media pipe framework.

3. Classification Module:
With provided points from the Image Processing Module, the classifier will decide thesign alphabet by taking reference from the training set by the help of classificationalgorithm. The output will be generated as text corresponding to the Sign alphabet decided by Classifier.

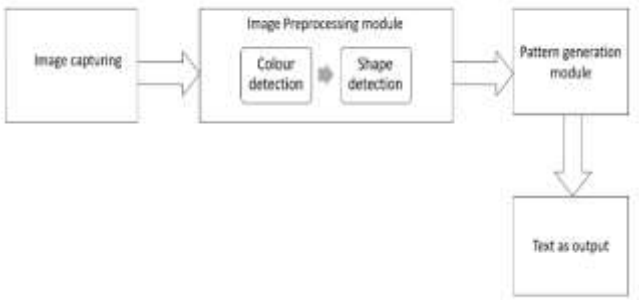


Fig 1. Architecture of the Proposed System.

Fig1 demonstrates the working of sign language interpretation system. Initially the image is captured which is then pre-pre-processed. The pre-processing consists of colour detection and shape detection module. The processed image is then given as input to pattern generation module further to which the generated pattern is converted to its equivalent text.

IV. ALGORITHM AND MATHEMATICAL MODEL

I. Sign Language Detection.
1. Capture image.
2. Retrieve key points of the hand.
3. Calculate distances between the focus of the given image and the clusters of different sign alphabets.
4. Whichever is nearest to the given image, give it as an output text.

CNN Algorithm:
Convolution operation: Prepare feature detector for the input image. Feature detector will be images of sign language. Make a corresponding feature map for it.

Rectified Linear Unit (ReLU): Apply Rectifier Function to increase the non-linearity of image.

Pooling: Apply max pooling on the image to make the network realize whichever the condition of the image, CNN must be able to identify it. Roughly 75% of information is filtered after this step.

Flattening: Flatten the 2*2 matrix to single dimensional array.

Full connection step : After Flattening feed the flattened array to the fully connected trained network and whichever class has the highest value as output, return the label of that class.(Here class represents alphabets of sign language) Mathematical model for Hand Gesture Recognition System as follows:
Let us consider S be the system such that

S= {I, F, O}

Where,
I = Set of Input
as a images captured by camera or by video such as raw video.
F = Set of Functions
F = {I, TD, IF, SS, G}
where,

I = Input images extracted from video or captured by camera.
TD = {T1, T2, T3....T n| T is the train data used to compare the captured images.
IF = {IF1, IF2, IF3.... IFn| IF is the important frames calculated by score function

$$F[V, S, w] = w^t f(V, S)$$

where F is the score function, V and S stands for the feature matrix of images and its candidate summary, w is the property weight

SS = Storage server
SS is used to store the extracted imported frames.

G = Generate the recognized images of Hand Gesture.

O = Output
Textual or Verbal result

V. ACTIVITIES AND EVENTS:

EVENT 1
User will make registration on SYSTEM and Storage Server. Let f(U) be a function of User.
Thus, f(U) →{U Ds}

EVENT 2
Let f(C) be a capture function of System.
Thus, f(C) → {I1,I2, I3In} ε C

EVENT 3
Pre Processing is performed on detected images. Let f(I) be a image processing function.
Thus, f(I) →{I1,I2,I3In} ε T
EVENT 4
Let f(FT) be an function used to extract the features on basis of train dataset
Thus, f(FT) →{T,I} ε FT

EVENT5
Let f(E) be be the Character Detection on extracted features using CNN .

Thus, f(E) → {FT} ε T

FLOWCHART
Object Detection using media pipe.

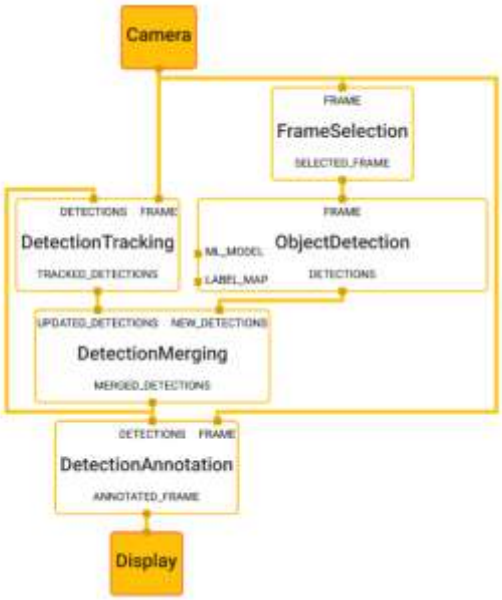


Fig 2. Object Detection using mediapipe.[8]

VI. CONCLUSION

In this paper we have proposed a sign language interpretation system and its conversion to text for deaf and dumb people. The system will make use of camera to capture hand gestures further media pipe framework will detect the hand gesture and the this gesture will be classified by matching it with the sample dataset which is used for the conversion of gesture into its equivalent text.

VII. REFERENCES

[1] Umang Patel, Arti G Ambekar,**"***Moment Based Sign Language Recognition for Indian Languages*"Third International Conference on Computing, Communication, Control and Automation (ICCUBEA),2017

[2] Xingang Fu, Ting Zhang, Chadwell H. Bonair, and Marvin L. Coats*,"Wavelet Enhanced Image Preprocessing and Neural Networks for Hand Gesture Recognition"* The University of Alabama Tuscaloosa, Alabama, USA,2015

[3] D.K. Vishwakarma, Sahib Majithiax and Nikhil Kumar Mishray,*"A Static Hand Gesture Recognition System to Recognize the Total number of Fingers"*Department of Electronics and Communication Engineering, Delhi TechnologicalUniversity New Delhi, India, 2017

[4] Dinesh Kumar Vishwakarma,*" Hand Gesture Recognition using Shape and Texture evidences in Complex Background"*Department of Electronics and Communication Engineering, Delhi Technological University Delhi, India, 2017.

[5] Naresh Kumar,*" Sign Language Recognition for Hearing Impaired People based on Hands Symbols Classification"*PhD Research Scholar Department of Mathematics Indian Institute of Technology, Roorkee India,2017.

[6] Anup Kumar et al*" Sign language recognition"* 3rd International Conference on Recent Advances in Information Technology (RAIT), 2016.

[7] Anup Kumar et al*" Sign language recognition"* 3rd International Conference on Recent Advances in Information Technology (RAIT), 2016.

[8] Flowchart Available on: https://www.groundai.com/project/mediapipe-a-framework-for-building-perception-pipelines/1

Sliding Mode Control with RBF Neural Network for Two Link Robot Manipulator

Ankita Yadav
Department of Electrical Engineering
Delhi Technological University

Ajit Kumar Sharma
Department of Electrical Engineering
Delhi Technological University

Bharat Bhushan
Department of Electrical Engineering
Delhi Technological University

ABSTRACT
Nonlinear control techniques are applied on mechanical systems namely two link robot manipulator to study the effect of the controllers on the tracking performance of the two system. A design of sliding mode control(SMC) for the position tracking of two link robot manipulator based on the sliding mode control technique and the Lyapunov stability theory is carried out to eliminate the perturbation and asymptotical stability can be achieved when the system is subjected to the sliding mode. A sliding mode control method based on RBF(radial basis function) neural network is addressed which has the capability of learning uncertain control actions shown by the several industrial robots. In RBFNN-SMC method the algorithm for tuning the parameters are extracted from the RBF function. The comparative study is done based on the evaluated parameters for the system

Keywords
Sliding mode control, RBF Neural Network, Two-link robot manipulator.

1. INTRODUCTION
Sliding mode control has been the vast topic for research in the field of control engineering. This section discuss Extensive research work has been done on sliding mode control of the system as well as the integration of compensation technique with this control methodology.[1] It is shown that the adaptive sliding mode control to control a 2 link robot manipulator having uncertain parameters. An adaptive algorithm is used which is based on the sliding mode control concept to ease the chattering phenomena. Lyapunov stability is applied to achieve the stability of robot manipulator system. Controller scheme proposed in that assures robustness and attain good trajectory tracking performance.[2] Conventional design methods for controllers for a MIMO plant like a multi-joint robot require minimum knowledge of the structure and accurate mathematical model of the plant. [3] The dynamic model and the terminal sliding mode control for two link robot manipulators. Lagrangian method is used to derive the closed loop dynamic equations for the robot manipulators. This technique has been used to design the controller. This technique gives the output tracking error to converge at zero in a given period of time. [4] A radial basis function network with SMC is designed for the link positions of two link robot manipulator. The learning techniques has degrades its transient performance also This neural technique uses the curve fitting mode to get nonlinear mapping. A sliding mode effectively used to eliminate thenonlinearities and has given a fast response but this control created chattering. The Lyapunov and the back propagation algorithm creates switch gain and an update law respectively.[5] Computed torque control is one of such schemes, which is based on the exact cancellation of the nonlinear dynamics of the manipulator system.[6]]Robot manipulators are very important in the field of flexible automation as they have been widely used for increasing productivity and reduce losses. A lot of research effort is been made to design their controller.[7]]Robot manipulators are very important in the field of flexible automation as they have been widely used for increasing productivity and reduce losses. A lot of research effort is been made to design their controller.[8] Neural Networks are being applied in many fields such as learning, pattern recognition, signal processing and system control. The major advantages this parallel structure, learning ability, non- linear function approximation, fault tolerance. This benefits make a wide range of applications..[9] Robot manipulators are very important in the field of flexible automation as they have been widely used for increasing productivity and reduce losses. A lot of research effort is been made to design their controller.[10] The payload of the robot manipulator varies during its operation and is not known advance. To face such problems, adaptive control strategies for robot manipulators have been developed.[11] Research from the past shows that universal approximation theorems on RBF has shown that any nonlinear function over a compact set with arbitrary accuracy can be approximated by RBF neural network.[12] The RBF network adaptation improves the control performance against large uncertainty of the system. The adaptation law can be derived by Lyapunov method which guarantees the stability of the system. RBF neural networks. A schematic diagram of two link robot manipulator is shown in figure 1.

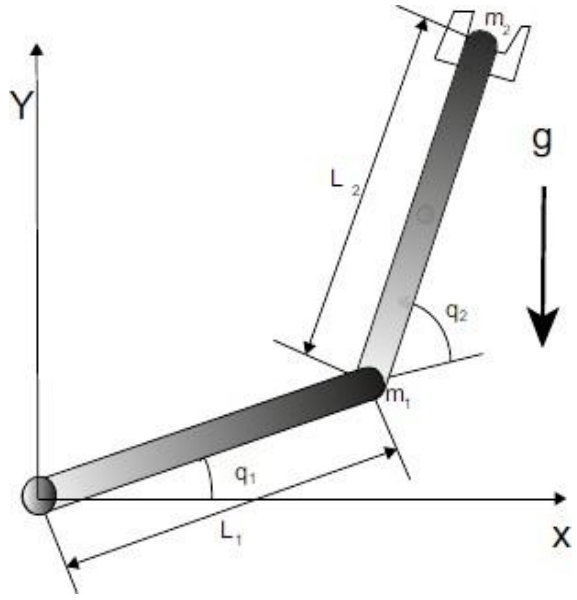


Figure 1: A two link robot manipulator

2. SLIDING MODE CONTROL

Sliding mode control deals with the problem of model uncertainties as these uncertainties can have strong adverse effects on nonlinear control systems. A major approach to deal with the model uncertainly is adaptive control and another approaches which can be used to solve the control problems includes the sliding mode techniques. These techniques are generating greater interest nowadays. Discrepancies may occur between the actual plant and the mathematical model established for the controller design. Various factors may be responsible for this mismatch. The engineer's role is to ensure required performance levels for the system instead of such mismatches. A set of robust control methods have been developed to eliminate any error. One such method is sliding mode control methodology(SMC). This is a specific type of variable structure control system (VSCS). SMC has been used for several systems including nonlinear system, multi-input multi- output(MIMO) systems, discrete-time models, large-scale and infinite-dimension systems, and stochastic systems. SMC is completely insensitive to parametric imprecisions and external disturbances during sliding mode. VSC uses a high-speed switching control law to accomplish two objectives. Firstly, the nonlinear plant's state trajectory is taken onto a specified and user-chosen surface in the state space which is called the sliding or switching surface. This is called as the switching surface because a control path has a unique gain if the state trajectory of the plant is "above" the surface and a different gain if the trajectory falls "below" the surface. Secondly, it keeps the plant's state trajectory on this surface for all consequent times. During this process, the control system's structure changes from one to another and thus given the name variable structure control(VSC). The control is also named as the sliding mode control to accentuate the importance of the sliding mode. Sliding mode controller can stabilize the trajectory of a system. Control structures are designed so as to ensure that trajectories will always move towards a switching condition. Therefore, the ultimate trajectory will not exist completely within one control structure. Instead, the ultimate trajectory will slide along the boundaries of the control structures. The motion of the system as it slides along these boundaries is termed as a sliding mode and the geometrical locus consisting of the boundaries is called the sliding (hyper) surface. The sliding mode phases are shown in figure 2

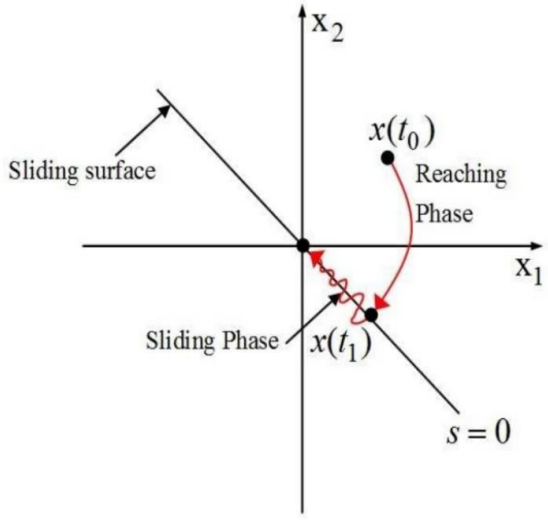


Figure 2: Sliding mode phases

3. LYAPUNOV STABILITY

The Lyapunov stability analysis is based upon the concept of energy and also depends on the relation of stored energy with the stability of system. But in general there is no way of relating the energy function with the given set of equations which describes the system. There is no uniqueness about the total energy of the system which associates with the system stability. There are scalar functions present which describe this relation and this ides is given by the mathematician A.M Lyapunov. The scalar functions are known as the Lyapunov Function and the method used for evaluating the stability of the control system is known as the Lyapunov's direct method. Following are some of the scalar functions which are described as the follows by the Lyapunov direct method. Based on these functions, stability can be evaluated.

(a)Positive Definiteness of Scalar Functions :V(x) which is a scalar function is said to be positive definite in the region $\|x\| \leq k$ if $V(x)>0$ at all the points of the region except origin since it is zero at origin.

(b)Negative Definiteness of Scalar Functions: V(x) which is a scalar function is said to be negative definite if [-V(x)] is positive definite.

(c)Positive Semi definiteness of Scalar Functions: V(x) which is a scalar function is said to be positive semidefinite in the region $\|x\| \leq k$ if its value is positive at all points of the region and at origin too.

(d)Negative Semi definiteness of Scalar Functions: V(x) which is a scalar function is said to be negative definite if [-V(x)] is positive semidefinite.

4. RBF NEURAL NETWORK

This is an alternative approach for multilayer perceptron which is trained by the backpropagation algorithm. It is the basic component of feedforward network. A gaussian function has two parameters: center for defining its position and for determining its shape, parameter called spread is there. In one dimensional gaussian function spread is same as the standard deviation σ.

The input vector is given by

$$x = [x_1 \;\; x_2 \ldots\ldots\ldots x_n] \quad (1)$$

and the output of the RBF network is given below

$$x(x, c, \sigma) = \exp\frac{-||x-c||^2}{2\sigma^2} \quad (2)$$

there is no weights between the input and the RBF unit. c and σ represents the weight which is shown in figure 3.

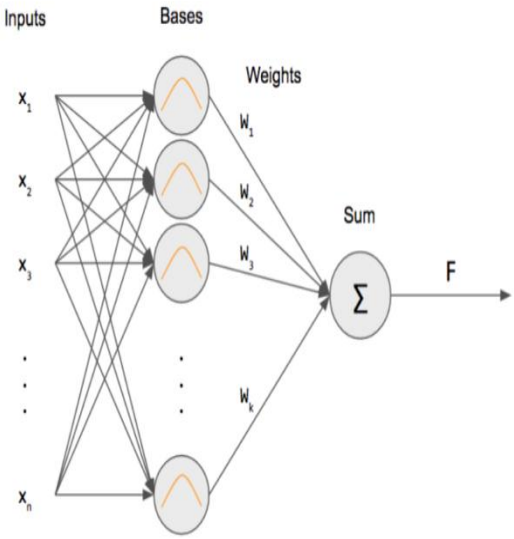


Figure 3: **Gaussian function in RBF neuron model**

5. SLIDING MODE CONTROLLER FOR A TWO LINK ROBOT

The plant model is given by equation (3)

$$\dot{x_1} = \dot{\theta} = x_2$$

$$\dot{x_2} = \ddot{\theta} = f(x) + g(x)\tau \quad (3)$$

Where

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_{11} \\ x_{12} \\ x_{21} \\ x_{22} \end{bmatrix} = \begin{bmatrix} \theta \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix}; \tau = \begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} \quad (4)$$

θ_1 and θ_2 are the angles of the two links defined in figure, and τ_1 and τ_2 are the torques applied by the actuators to control the angles θ_1 and θ_2 respectively.

Variable structure sliding mode controller design consists of the following two phases:

(a)Designing of sliding surface to obtain the desired system behaviour, when restricted to the surface. The system will remain at that surface for further times. The plant dynamics will always be limited to the surface and is robust towards the disturbances.

(b)Selection of the feedback gains of the controller, so that the closed loop system is stable to the sliding surface.

For a two link robot manipulator, our aim is to track the desired motion trajectory $\theta_d(t)$

The tracking error is defined a

$$e(t) = \theta_d(t) - \theta(t)$$

$$\dot{e(t)} = \dot{\theta_d}(t) - \dot{\theta}(t); \ddot{e}(t) = \ddot{\theta_d}(t) - \ddot{\theta}(t)$$

Defining $\widetilde{x_1} = e$ and $\widetilde{x_2} = \dot{e}$ robot dynamics can be written in the form:

$$\dot{x_2} = \ddot{\theta} = f(x) + g(x)\tau \quad (5)$$

Now we use linear sliding surface which is defined by the equation

$$\sigma(\tilde{x}) = \lambda\widetilde{x_1} + I\widetilde{x_2} \quad (6)$$

Or

$$\begin{bmatrix} \sigma_1(\tilde{x}) \\ \sigma_2(\tilde{x}) \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} \widetilde{x_{11}} \\ \widetilde{x_{12}} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \widetilde{x_{21}} \\ \widetilde{x_{22}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} (7)$$

$$\sigma_1(\tilde{x}) = \lambda_1\tilde{x}_{11} + \widetilde{x_{21}} = \lambda_1(\theta_{1d} - \theta_1) + (\dot{\theta_{1d}} - \dot{\theta}_1)$$

$$\sigma_2(\tilde{x}) = \lambda_2\tilde{x}_{12} + \widetilde{x_{22}} = \lambda_2(\theta_{2d} - \theta_2) + (\dot{\theta_{2d}} - \dot{\theta}_2) \quad (8)$$

Assume the coefficient of $\widetilde{x_{21}}$ to be unity.

Now combine the equations of the plant and the sliding surface (Eqns (7) and (8))

$$\dot{\widetilde{x_1}} = -\lambda\widetilde{x_1} \quad (9)$$

Therefore

$$\widetilde{x_2} = -\lambda\tilde{x}_1 \quad (10)$$

The above equations describes the system dynamics in sliding. The response is specified by the selection of the parameters λ_1 and λ_2 of the switching surface. In sliding mode, the system is not affected by the model uncertainties. After designing the sliding surface, we build a feedback controller. The objective of the controller is to derive the plant state to the sliding surface and maintain it on the surface for further time. A Lyapunov approach has been used for building the controller.

The Lyapunov function is given below

$$V = \frac{1}{2}\sigma^T\sigma = \frac{1}{2}(\sigma_1^2 + \sigma_2^2) \quad (11)$$

The above function is used as the Lyapunov function and the time derivative of the chosen Lyapunov function comes out to be negative definite with respect to switching surface and ensuring the motion of the state trajectory to the surface. We have to find τ so that

$$\frac{d}{dt}\frac{1}{2}\sigma^T\sigma = \sigma^T\dot{\sigma} < 0$$

$$\sigma^T\dot{\sigma} = \sigma^T[\lambda\dot{\widetilde{x_1}} + I\dot{\widetilde{x_2}}] \quad (12)$$

The controller structure form is given below

$$\tau = g^{-1}(x)\left[-f(x) + \dot{x_{2d}} + \lambda(\dot{x_{1d}} - \dot{x_1}) + \begin{Bmatrix} k_1 sgn(\sigma_1) \\ k_2 sgn(\sigma_2) \end{Bmatrix}\right] \quad (13)$$

where k_1>0 and k_2>0 are the gains which has to be determined so that the condition $\sigma^T\dot{\sigma} < 0$ is satisfied. We substitute τ, given by (3.52) into the expression $\sigma^T\dot{\sigma}$

$$\sigma^T\dot{\sigma} = -[\sigma_1 \quad \sigma_2]\begin{bmatrix} k_1 sgn(\sigma_1) \\ k_2 sgn(\sigma_2) \end{bmatrix} = -\sigma_1 k_1 sgn(\sigma_1) - \sigma_2 k_2 sgn(\sigma_2) = -k_1|\sigma_1| - k_2|\sigma_2| < 0 \quad (14)$$

The sliding surface $\sigma(\tilde{x}) = 0$ is therefore asymptotically attractive. The larger the value of gains, faster the trajectory converges to the sliding surface. The tolerance of the sliding mode control to the model imprecision and disturbances, is high and thus satisfying stability requirement. The Simulink model for the sliding mode control of the two link robot manipulator is given in figure 3.2

Simulation of this controller for a two link robot manipulator was done in MATLAB and results are attached at the end.

The specifications of the robot manipulator is taken as:

$m_1 = 1kg; m_2 = 1kg; l_1 = 1m; l_2 = 1m; g =$

$\frac{9.8kgm}{s^2}; \theta_{d1}(t) = \sin(\pi t); \theta_{d2}(t) = \cos(\pi t)$ (15)

Now the objective is to construct an RBF network sliding mode control for robot manipulator. This is completely opposite to the SMC technique so far used as the unknow disturbances are approximated by RBF neural network.

6. RBF NETWORK ADAPTIVE SLIDING MODE CONTROL FOR ROBOT MANIPULATOR

The dynamic equation of 2-joint manipulator is as follows:

$$H(q)\ddot{q} + C(q,\dot{q})\dot{q} + G(q) = \tau - F(\dot{q}) - \tau_d$$

where H(q) is a 2x2 positive definite integral matrix, G(q) is a 2x1 inertial vector, $C(q,\dot{q})$ is a 2x2 inertial matrix, τ_d is the unknown disturbance, $F(\dot{q})$ is friction force, and τ is the control input

The tracking error is denoted as:

$e(t) = q_d(t) - q(t)$ (17)

Sliding variable is selected as:

$s = \dot{e} + \dot{\Lambda} e$ (18)

where Λ is a symmetric positive definite matrix and $\Lambda = \Lambda^T > 0$, therefore , we have

$\dot{q} = -s + \dot{q_d} + \Lambda e$ (19)

$$H\dot{s} = H(\ddot{q_d} - \ddot{q} + \Lambda\dot{e}) = H(\ddot{q_d} + \Lambda\dot{e}) - H\ddot{q}$$

$$= H(\ddot{q_d} + \Lambda\dot{e}) - Cs + C(\dot{q_d} + \Lambda e) + G + F + \tau_d - \tau$$

$$= H(\ddot{q_d} + \Lambda\dot{e}) + C\dot{q} + G + F + \tau_d - \tau$$

$= -Cs - \tau + f + \tau_d$ (20)

Where f(x) = $H(\ddot{q_d} + \Lambda\dot{e}) + C(q_d + \Lambda e) + G + F$

as f(x) is unknown, hence RBF network is adopted to approximate f(x). input is selected based on expression of f(x):

$x = [e^T \quad \dot{e}^T \quad q_d^T \quad \dot{q}_d^T \quad \ddot{q}_d^T]$ (21)

The controller is designed as follows:

$\tau = \hat{f}(x) + K_v s$ (22)

where Kv is a symmetrical positive definite constant matrix, $\hat{f}(x)$ is the output of RBF

network, approximates f(x). The RBF-NN weights are adjusted using adaptive laws

From equations (21) and (22) we have

$H\dot{s} = -Cs - \hat{f}(x) - Kvs + f(x) + \tau_d$ (23)

Where

$f(x) = -(K_v + C)s + \tilde{f}(x) + \tau_d = -(K_v + C)s + \varsigma_0$

Lyapunov function is selected as:

$L = \frac{1}{2}s^T H s$ (24)

$\dot{L} = s^T H s + \frac{1}{2}s^T \dot{H} = -s^T K_v s + \frac{1}{2}s^T(\dot{H} - 2C)s + s^T\varsigma_0$ (25)

$\dot{L} = s^T\varsigma_0 - s^T K_v s$ (26)

It shows that with Kv, the stability of control system depends on ς_0, the magnitude of τ_d

The algorithm suggested to approximate f(x) is given below

$\phi_i = g(\frac{||x - c_i||^2}{\sigma_i^2})$ (27)

$y = W^{*T}\varphi(x), f(x) = W^{*T}\varphi(x) + \varepsilon$ (28)

Where x is the input state of network,

$$\varphi(x) = [\phi_1 \quad \phi_2 \quad \ldots\ldots \quad \phi_n]^T$$

ε is the approximation error, W^* is the desired weight vector.

7. SIMULATION AND RESULTS

Kinetic equation of two joint robot manipulator is:

$H(q)\ddot{q} + C(q,\dot{q})\dot{q} + G(q) = \tau - F(\dot{q}) - \tau_d$ (29)

Where

$$\boldsymbol{H(q)} = \begin{bmatrix} \boldsymbol{p_1 + p_2 + 2p_3 cosq_2} & \boldsymbol{p_2 + p_3 cosq_2} \\ \boldsymbol{p_2 + 2p_3 cosq_2} & \boldsymbol{p_2} \end{bmatrix}$$ (30)

$$C(q,\dot{q}) = \begin{bmatrix} -p_3 q_2 sinq_2 & -p_3(\dot{q_1} + \dot{q_2})sinq_2 \\ p_3\dot{q_1}sinq_2 & 0 \end{bmatrix}$$ (31)

$$G(q) = \begin{bmatrix} p_4 g cosq_1 & p_5 g cos(q_1 + q_2) \\ p_5 g cos(q_1 + q_2) \end{bmatrix}$$ (32)

$F(\dot{q}) = 0.02sgn(\dot{q})\ ; \tau_d = [0.2sint \quad 0.2sint]^T$ (33)

Let

$$\mathbf{p} = [\boldsymbol{p_1} \quad \boldsymbol{p_2} \quad \boldsymbol{p_3} \quad \boldsymbol{p_4} \quad \boldsymbol{p_5}]$$
$$= [\mathbf{2.9} \quad \mathbf{0.76} \quad \mathbf{0.87} \quad \mathbf{3.04} \quad \mathbf{0.87}]$$

Gaussian function of the RBF network is selected to control the neural network. If the parameter is not suitable then mapping of the gauss function cannot be obtained and the RBF network is unavailable. The model for this is given in the figure 4.

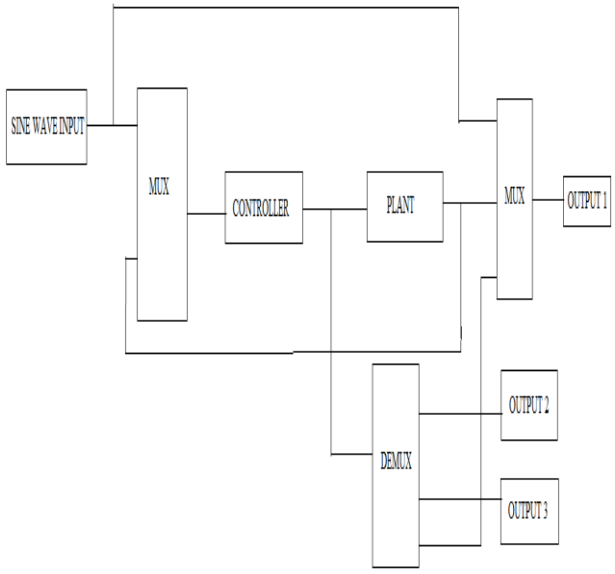


Figure 4: Block diagram for RBF network sliding mode control for two link robot manipulator

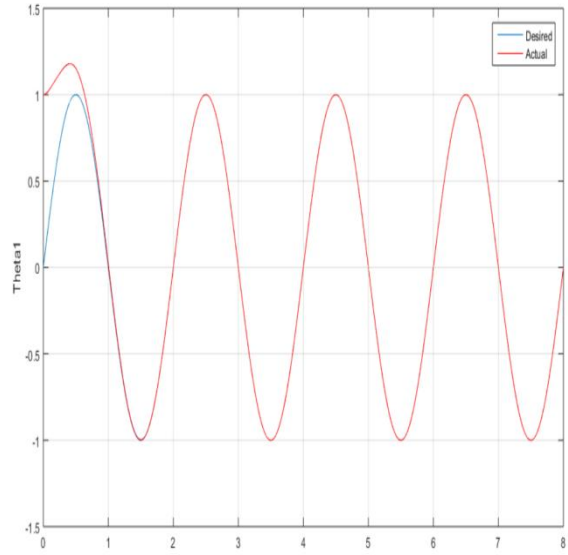


Figure 5: Desired and actual trajectory for link 1(θ_1) using SMC

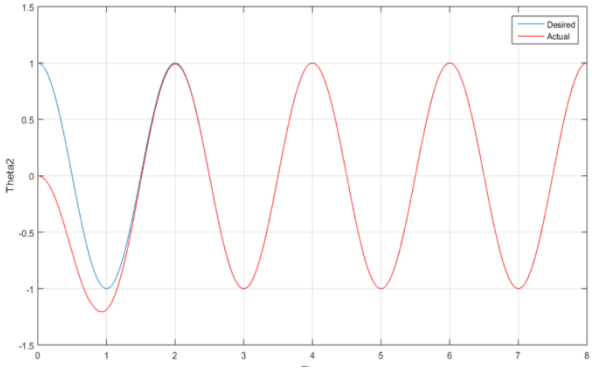


Figure 6: Desired and actual trajectory for link 2 (θ_2)using SMC

Here figure 5 and 6 shows the desired and actual trajectory of the link 1 and link 2 of two link robot manipulator applied with SMC whereas figure 7 and 8 displays the same for RBF-SMC

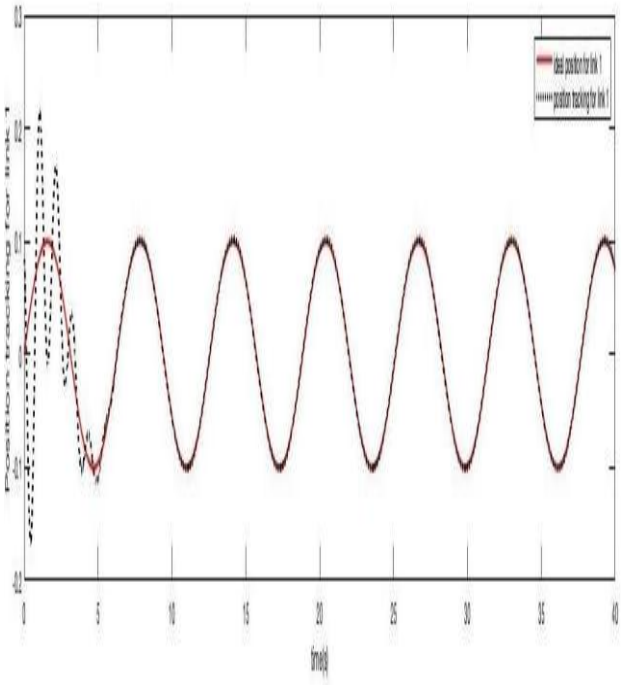


Figure 7: Desired and actual trajectory for link 1(θ_1) using RBFNN-SMC

Table 1. Parameters for position of link 1 and 2

LINKS		DELAY TIME(sec)	RISE TIME (sec)	PEAK TIME (sec)	PEAK OVERSHOOT %
LINK 1	ACTUAL	0.02134	0.1524	0.4113	17.9
	DESIRED	0.03665	0.3618	0.5031	0
LINK 2	ACTUAL	0.1969	0.7317	0.9231	14.53
	DESIRED	0.2018	0.8579	0.9978	0

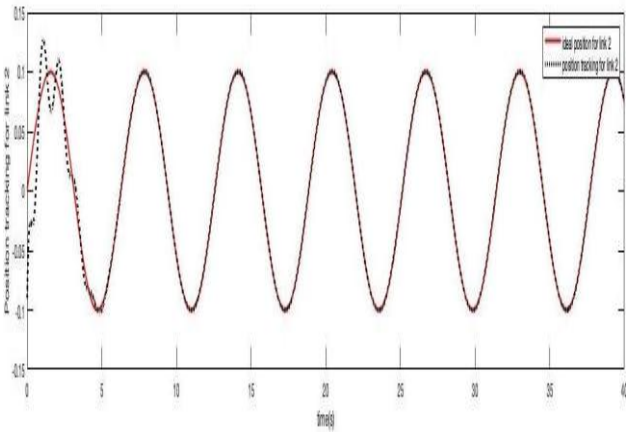


Figure 8: Desired and actual trajectory for link 2 (θ_2)using RBFNN-SMC

8. CONCLUSIONS

The mathematical model of two DOF robot manipulator is designed in the form of differential equations and then feedback linearization technique is applied to obtain the plant model and two control techniques that is adaptive sliding mode control and RBF neural network based sliding mode control have been successfully applied to improve the stability of this system.it is seen that position tracking is instantly and correctly achieved in case of RBF-SMC model, also the system implemented with RBF-SMC is more stable. A comparative study of the parameters for the links of robot manipulator has been carried out.

Table 2. Parameters for position of link 1 and 2

LINKS		DELAY TIME(sec)	RISE TIME (sec)	PEAK TIME (sec)	PEAK OVERSHOOT %
LINK 1	ACTUAL	0.01874	0.1224	0.3216	14.18
	DESIRED	0.02265	0.3103	0.3031	0
LINK 2	ACTUAL	0.1472	0.6283	0.8163	12.44
	DESIRED	0.1524	0.7152	0.8816	0

9. REFERENCES

[1] K.M.Koo, J.H.Kim: Robust Control of Robot Manipulators with Parameters Uncertainty, IEEE Trans.Auto.Contr., Vol.39, No.6, 1230-1233,1994

[2] Utkin V.I., "Variable structure systems with sliding modes", IEEE Transactions on Automatic Control, vol. 22, no. 2, pp. 212-22, 1977.

[3] V. I. Utkin, Sliding Modes and Their Application in Variable Structure Systems. Moscow: MIR, 1978.

[4] Slotine J. and Sastry S., "Tracking control of nonlinear system using sliding surfaces with application to robot manipulators", International Journal of Control, vol. 38, pp.465–492, 1983.

[5] J.J.E.Slotine, W. Li: On the Adaptive Control of Robot Manipu -lators,The Int. J. Robotics Reseach,Vol.6,No.3, 49-59, 1987

[6] Slotine J.E. and Li W.P., "On the adaptive control of robot manipula-tors", Interational Journal of Robotics Research, vol. 6, no. 3, pp.49–59, 1987

[7] J.J.E.Slotine, W. Li: Adaptive Manipulators Control: A Case Study, IEEE Trans. Auto. Contr.,Vol.33,No.11,995 -1003, 1988

[8] R. A. DeCarlo, S. H. Zak, and G. P. Matthews, "Variable Structure Control of Nonlinear Multivariate Systems: A Tutorial," Proceedings of the IEEE, Vol. 76, No. 3, pp. 212-232, March 1988.

[9] Astrom K.J. and Wittenmark B., "Adaptive Control", 2nd edn. Addison-Wesley, NY 1994,1989.

[10] V. I. Utkin and S. Drakunov, "On Discrete-time Sliding Mode Control," Proceedings of IFAC Symposium on Nonlinear Control Systems (NOLCOS), (Capri, Italy), pp. 484-489, 1989.

[11] Hartman EJ, Keeler JD, Kowalski JM. Layered neural networks with Gaussian hidden units as universal approximations. Neural computation, 1990, 2(2): 210 _ 215

[12] Kanellakopoulos I, Kokotovic PV, Morse AS (1991) Systematic design of adaptive controllersfor feedback linearizable systems. IEEE Trans Autom Control 36(11):1241–1253

[13] Park J, Sandberg IW. Universal approximation using radial-basis-function networks. Neural computation, 1991,3: 246 _ 257

[14] Narendra KS (1991) Adaptive control using neural networks. In: Neural networks for control.MIT Press, Cambridge, MA, pp 115–142

Nonlinear Dyn
https://doi.org/10.1007/s11071-019-05276-z

ORIGINAL PAPER

Stability behavior of a nonlinear mathematical epidemic transmission model with time delay

Kanica Goel · Nilam

Received: 28 June 2019 / Accepted: 23 September 2019
© Springer Nature B.V. 2019

Abstract In this article, we study a time-delayed susceptible–infected–recovered mathematical model along with nonlinear incidence rate and Holling functional type II treatment rate for epidemic transmission. The mathematical study of the model demonstrates that the model exhibits two equilibria, to be specific, disease-free equilibrium (DFE) and endemic equilibrium (EE). We obtain the basic reproduction number R_0 and investigate that the model is locally asymptotically stable at DFE if $R_0 < 1$ and unstable if $R_0 > 1$ for the time lag $\nu > 0$. The stability of DFE at $R_0 = 1$ is also investigated for the time lag $\nu \geq 0$, and we show that for $\nu > 0$, the DFE is linearly neutrally stable, whereas for $\nu = 0$, the model exhibits backward bifurcation whereby the DFE will coexists with two endemic equilibria, when $R_0 < 1$. We also investigate the stability of the model at the EE and find that oscillatory solution may appear via Hopf bifurcation, taking the delay as a bifurcation parameter. Further, global stability of the model equilibria has also been analyzed. Finally, numerical simulations have been presented to illustrate the analytical studies.

K. Goel · Nilam (✉)
Department of Applied Mathematics, Delhi Technological University, Delhi 110042, India
e-mail: nilam.research18@gmail.com; rathi.nilam@gmail.com

K. Goel
e-mail: kanica.dtu2016@gmail.com

Keywords Time delay · Nonlinear incidence · Holling functional type II · Basic reproduction number (BRN) · Stability · Bifurcations

Mathematics Subject Classification 34D20 · 92B05 · 37M05

1 Introduction

Mathematical modeling is progressively used in various fields such as biology, engineering, and economics [1–5]. In epidemiology, mathematical modeling is an effective method to study a diverse range of infectious diseases, for a better understanding of transmission dynamics. Many authors have proposed numerous epidemic models, such as susceptible–infected–recovered (SIR) [2,6–8], SEIR [9], SVEIR [10], etc., and studied the spread of infectious diseases mathematically and quantitatively. A major goal of the epidemiological study is to develop an understanding of the transmission of epidemics and interventions that can be taken to prevent and control the spread of infectious diseases. For the most part, an exemplary condition for the elimination of an infection is for the basic reproduction number (BRN) [11] to be less than one whereby the disease-free equilibrium is stable; the unique endemic equilibrium will be asymptotically stable as far as when the BRN is greater than one, implies that the disease-free equilibrium will be unstable. However, many authors have exhibited that it is feasible for the disease-free

Published online: 10 October 2019

equilibrium to exist together with two endemic equilibria even when the BRN is less than one, leading to the phenomenon of backward bifurcation [12,13]. This phenomenon demonstrates that for disease control and destruction, relying on the BRN to be less than one is now only necessary but no longer sufficient. Therefore, the occurrence of backward bifurcation has important public health implications. Many authors have studied epidemic models characterized by backward bifurcation, for both generic and specific diseases [3,14–16]. The phenomenon of backward and forward bifurcation is based on the center manifold theory.

Incidence rate has a major role in the study of transmission dynamics of infectious disease. The term incidence refers to the number of new cases of a disease occurring in a given period. The incidence rate is calculated by dividing the total number of new cases of the disease in a certain period into the total number of people in the population. The type of the incidence rate that is used in the classical Kermack Mckendrick model (1927) [1] is the basic mass action (bilinear) λSI where S and I denote the number of susceptible and infected individuals, respectively, and λ is known as the infection coefficient. But, this bilinear incidence rate is not reasonable for large population because if the number of susceptibles increases, then the incidence rate also increases, i.e., the number of infected individuals per unit time also increases. So this unreasonable bilinear incidence rate is needed to be modified. Due to this unrealistic situation, several authors [6,7,9,17–21] have focused on the nonlinear type of incidence rate. Nonlinear incidence rate permits to incorporate social/behavioral changes and prevent unbounded contact rates. Thus, changing the form of incidence rate can potentially fill the gap in the transmission behavior of the system. There are several types of nonlinear incidence rates like Holling type II, Crowley–Martin type and many others that have already been used by the researchers in their model to understand the transmission dynamics of epidemics. In 1975, Beddington [22] and DeAngelis [23] independently presented a nonlinear incidence rate of the form $\frac{\beta SI}{1+\alpha S+\gamma I}$ $(\beta, \alpha, \gamma > 0)$, called Beddington–DeAngelis functional type incidence rate. After that some authors [24–26] introduced the same incidence rate in their disease transmission models. This incidence rate has a significant role in studying transmission dynamics as it includes inhibition measures taken by both susceptible and infected individuals. If individuals are familiar with the diseases and have knowledge about the transmission modality of infection, then they can take necessary preventive measures to avoid infections. These measures are called the measure of inhibition. In the present article, we take the Beddington–DeAngelis type incidence rate. The Beddington–DeAngelis type incidence rate of susceptible individuals through their contacts with infectious individuals is a more general and realistic form for the study of diseases transmission dynamics as it contains three forms: bilinear incidence rate if $\alpha = \gamma = 0$; saturated incidence rate for the susceptible if $\gamma = 0$; and saturated incidence rate for the infectives if $\alpha = 0$. Preventive anti-epidemic measures represent a set of measures aimed at preventing the occurrence and slowing down the spread of infections in a population such as the decreased contact rate, quarantinization of the infected individuals, introductory examinations before being employed, improvement of the hygienic level of population, treatment for infectives. To study the realistic phenomenon of disease transmission, we must consider the inclusion of the time delay into the model as when a disease emerges, there will be an initial delay in recognizing it. Therefore, present article incorporates the impact of time delay into Beddington–DeAngelis incidence rate and investigate its effect on the transmission dynamics. Inhibitions are effective measures to control the spread of infectious diseases but to control the spread of further infection and to eliminate the infection from the society; there is a need for effective treatment also. Treatment rate has a substantial role in reducing complications and preventing the transmission of infectious diseases to others. Appropriate and timely treatment strategy can significantly lessen the effect of disease on society. The basic idea behind any treatment is to cure a disease and to control its spread. In classical epidemic models, the treatment rate of infectives is considered to be either constant or proportional to the number of infectives. But, as the community has low treatment resources availability, therefore, choosing a suitable treatment rate of disease becomes of grave importance. In the presence of insufficient treatment resources and vaccines availability, the epidemical control systems depend on taking suitable preventive measures. Preventive measures refer to measures that are applied to prevent the occurrence of a disease. In 2004, Wang and Ruan [27] considered an SIR epidemic model with constant treatment rate (i.e., the recovery from infected subpopulation per unit time) as given below:

$$h(I) = \begin{cases} b, & I > 0 \\ 0, & I = 0 \end{cases}$$

where b is a positive constant and I denotes the number of infected individuals. They investigated the stability of the model and proposed that their model could exhibit various bifurcations. This kind of treatment rate is appropriate for the situation when there are small numbers of infectives and treatment resources are sufficient and inappropriate when the number of the infected individuals is large, and treatment resources are limited. Therefore, Zhang and Liu [28] presented the improved continuous, differentiable nonlinear treatment rate function which saturates at its maximum value. This saturated treatment rate is known as the Holling type II treatment rate [9]. For any outbreak of epidemic disease, the treatment capacity of Holling type II is initially very slow which develops gradually with improved treatment facilities. Further, Holling type II treatment rate refers to the condition that when there is a large number of infected people, then the treatment capacity reaches to its maximum because of limited treatment facilities which mimic the more realistic phenomena.

Delay differential equation reveals more complex dynamical behavior than an ordinary differential equation. The time delay is an influential parameter in the dynamical behavior of infectious diseases, and it can change the behavior of the dynamical system. The present paper aims to study the effect of latent period (the period between exposure and infection is called latent period, since the pathogen is present in a latent stage, without clinical symptoms or signs of infection in the host) on the SIR model and analyze the transmission of epidemics into the susceptible host with the impact of inhibitions taken by both susceptible and infected individuals. We also aim to study the impact of treatment rate in which we have considered the fact of the limited availability of resources, which allows us to reach more realistic phenomena. For this, we have considered an SIR epidemic model along with Beddington–DeAngelis functional type incidence rate by incorporating time delay as latent period, and Holling functional type II treatment rate. We derive the basic reproduction number for the proposed model, and we study the dynamical behavior of the model through stability analysis. We give a precise indication of the bifurcation phenomenon using center manifold theory, ensuring that either forward or backward bifurcation occurs. We also discuss the Hopf bifurcation near the endemic equilibrium by considering time delay as a bifurcation parameter.

The organization of the paper is as follows: In Sect. 2, we introduce the model formulation. Basic properties of the solutions are discussed in Sect. 3. In Sect. 4, we investigate the existence of equilibria and present the qualitative analysis through stability at disease-free equilibrium and endemic equilibrium. We also perform the bifurcation analysis whereby both forward and backward scenarios are derived. Further, by considering the time delay as a bifurcation parameter, we obtain the conditions for the occurrence of Hopf bifurcation around the endemic equilibrium. Furthermore, global stability of the disease-free equilibrium and endemic equilibrium has been discussed with the help of Lyapunov direct method. Numerical simulations have been performed in Sect. 5 using MATHEMATICA 11, to demonstrate the theoretical results. Finally, significant findings are discussed in Sect. 6.

2 Mathematical model

The fundamental approach in studying epidemic models is to divide the total population into mutually exclusive compartments according to epidemic status. We assume that the total size of population N is constant. Therefore, the total constant population is divided into three disjoint compartments $S(t)$, $I(t)$, and $R(t)$, where $S(t)$, $I(t)$, and $R(t)$ denotes the size of compartment of susceptible, infected, and recovered at time t, respectively. We assume that each population of SIR is well mixed and interact homogeneously with each other [29]. That is, $N \equiv S(t)+I(t)+R(t)$ which means that N is a fixed population and $S(t)$, $I(t)$ and $R(t)$ may vary at time t.

Latent period has been considered as time delay ν which means that the force of infection at present is determined by the number of infectives in the past.

The model we present is under the framework of the following system of nonlinear delay differential equations:

$$\begin{aligned} \frac{\mathrm{d}S}{\mathrm{d}t} &= A - \vartheta S - \frac{\beta S(t-\nu)I(t-\nu)}{1+\alpha S(t-\nu)+\gamma I(t-\nu)}, \\ \frac{\mathrm{d}I}{\mathrm{d}t} &= \frac{\beta S(t-\nu)I(t-\nu)}{1+\alpha S(t-\nu)+\gamma I(t-\nu)} \end{aligned}$$

Table 1 Summary of symbols

Symbol	Description
N	Total constant population
$S(t)$	Susceptible individuals
$I(t)$	Infected individuals
$R(t)$	Recovered individuals
ν	Latent period (time delay)
A	Recruitment rate of susceptible individuals
α	Inhibitions taken by susceptibles
β	Transmission rate
ϑ	Natural death rate
d	Disease-induced death rate
γ	Inhibitions taken by infectives
ζ	Recovery rate
a	Cure rate
b	Rate of limitation in treatment availability
$\Phi(S(t), I(t))$	Beddington–DeAngelis functional type incidence rate
$h(I(t))$	Holling type II treatment rate of infected individuals

$$- (\vartheta + d + \zeta)I - \frac{aI}{1+bI},$$
$$\frac{\mathrm{d}R}{\mathrm{d}t} = \frac{aI}{1+bI} + \zeta I - \vartheta R. \tag{1}$$

The model (1) is considered with the following initial conditions:

$$S(\theta) = \phi_1(\theta), I(\theta) = \phi_2(\theta), R(\theta) = \phi_3(\theta)\ \phi_i(\theta) \geq 0, \theta \in [-\nu, 0], \phi_i(\theta) > 0\ (i = 1, 2, 3) \tag{2}$$

where $(\phi_1(\theta), \phi_2(\theta), \phi_3(\theta)) \in C([-\nu, 0], R^3_+)$. Here, C denotes the Banach space of continuous functions mapping the interval $[-\nu, 0]$ into R^3_+, and $\nu > 0$ denotes the latent period.

The description of the parameters used in the model (1) is as follows [Table 1]: A represents the recruitment rate of the susceptible population by birth or immigration. ϑ and d are the natural death rate and disease-induced death rate, respectively. ζ is the recovery rate of infectives. The term $\frac{\beta S(t-\nu)I(t-\nu)}{1+\alpha S(t-\nu)+\gamma I(t-\nu)}$ describes the Beddington–DeAngelis type nonlinear incidence rate, where β is the force of infection, α is the measure of inhibitions taken by susceptibles, i.e., social awareness among the susceptibles, and γ is the measure of inhibition taken by infected individuals, i.e., the magnitude of interference among the infectives. This incidence rate represents the rate at time $(t - \nu)$ at which susceptible individual leaves the susceptible compartment and enters in the infectious compartment at time t. The term $h(I) = \frac{aI}{(1+bI)}$ is the Holling functional type II treatment term, where a denotes the cure rate of infectives and b denotes the limitation in the availability of resources.

3 Basic properties

Under the assumption that the population size N is constant, the model (1) may be reduced to a two-dimensional system for analysis purpose only. The equation for $R(t)$ is traditionally omitted; it is easy to see that under the assumption of constant population size, this equation is decoupled from the first and second equations of the model (1). The condition $N = S + I + R =$ constant may be used to find R [29,30]. In this manner, it is equivalent to study the following reduced system:

$$\frac{\mathrm{d}S}{\mathrm{d}t} = A - \vartheta S - \frac{\beta S(t-\nu)I(t-\nu)}{1+\alpha S(t-\nu)+\gamma I(t-\nu)}, \tag{3}$$
$$\frac{\mathrm{d}I}{\mathrm{d}t} = \frac{\beta S(t-\nu)I(t-\nu)}{1+\alpha S(t-\nu)+\gamma I(t-\nu)} - (\vartheta + d + \zeta)I - \frac{aI}{1+bI} \tag{4}$$

with initial conditions

$$S(\theta) = \phi_1(\theta), I(\theta) = \phi_2(\theta), \phi_i(\theta) \geq 0,\ \theta \in [-\nu, 0], \phi_i(\theta) > 0\ (i = 1, 2) \tag{5}$$

where $(\phi_1(\theta), \phi_2(\theta)) \in C([-\nu, 0], R^2_+)$. Here, C denotes the Banach space of continuous functions mapping the interval $[-\nu, 0]$ into R^2_+.

We assume that $\Phi(S(t), I(t)) = \frac{\beta S(t) I(t)}{1+\alpha S(t)+\gamma I(t)}$ and $h(I(t)) = \frac{aI(t)}{1+bI(t)}$ are always positive, continuously differentiable, and monotonically increasing for all $S, I > 0$, i.e., the following postulates are satisfied:

P1. $\Phi(S(t), I(t)) > 0$, $\Phi'_S(S(t), I(t)) > 0$, $\Phi'_I(S(t), I(t)) > 0$ for $S(t) > 0$ and $I(t) > 0$.

P2. $\Phi(S(t), 0) = \Phi(0, I(t)) = 0$, $\Phi'_S(S(t), 0) = 0$, $\Phi'_I(S(t), 0) > 0$ for $S(t) > 0$ and $I(t) > 0$.

P3. $h(0) = 0, h'(0) > 0$ for $I(t) \geq 0$.

For biological reasons, we consider that all the parameters of the system (3)–(4) are positive, i.e., $A, \vartheta, \beta, \alpha,$

$d, \gamma, \zeta, a, b > 0$. The equations of the system (3)–(4) monitor populations; therefore, it is essential to prove the positivity of the state variables, i.e., $(S, I) \in R^2_+$.

Lemma 1 *Any solution of system* (3)–(4) *with* $\phi_i(\theta) > 0$, $\theta \in [-\nu, 0]$ *remains positive whenever it exists and bounded in the region*

$$D = \left\{(S, I) \in R^2_+ : S(t) + I(t) \leq \frac{A}{\vartheta}\right\}.$$

Proof Note that

$$\frac{\mathrm{d}S}{\mathrm{d}t} + \frac{\mathrm{d}I}{\mathrm{d}t} = A - \vartheta S - (\vartheta + d + \zeta)I - \frac{aI}{1 + bI} \leq A - \vartheta(S + I). \tag{6}$$

and so,

$$\lim_{t\to\infty}(S + I) \leq \frac{A}{\vartheta}. \tag{7}$$

It follows that the system is point dissipative. Without loss of generality, we assume that $S(t) + I(t) \leq \frac{2A}{\vartheta}$ for all $t \geq -\nu$. A consequence of this is that we may assume I is bounded above, which in turn implies $\frac{dS}{dt}$ is positive for small S, and so S is positive for $t > 0$ [31]. □

We now prove the positivity of $I(t)$. Using the variation of constants formula and applying step by step integration method, for $0 < t \leq \nu$, we integrate Eq. (4) from 0 to t , we get

$$I(t) = I(0) \cdot e^{(\vartheta+d+\zeta)} \cdot e^{\int_0^t f(S(\omega-\nu), I(\omega-\nu), I(\omega))\mathrm{d}\omega} \tag{8}$$

We see that $I(t) > 0$ for all $\nu \leq t \leq 2\nu$. Proceeding in the same way, this process will carry on. Hence, it is proved that $I(t) > 0$ for all $t > 0$.

Also, from (7), it follows that (S, I) is bounded in the region

$$D = \left\{(S, I) \in R^2_+ : S(t) + I(t) \leq \frac{A}{\vartheta}\right\}.$$

Hence, we state the following lemma:

Lemma 2 *The solutions* $(S(t), I(t))$ *of the system* (3)–(4) *with initial conditions* (5) *uniquely exists in the region* $D = \{(S, I) \in R^2_+ : S(t) + I(t) \leq \frac{A}{\vartheta}\}$.

Proof We already have proved the positivity and the boundedness of solution $(S(t), I(t))$. Now, we will prove the existence and uniqueness of this solution. We notice that the right-hand side of the model (3)–(4) is completely continuous and locally Lipschitzian on D. Then, it follows that the solution of model (3)–(4) exists and is unique. This completes the proof of Lemma 2. □

4 Equilibria and stability analysis

In this section, we show the existence of the equilibria of the system (3)–(4) and investigate the local and global stability behavior of the obtained equilibria. We obtain the equilibrium points of time-delayed system by setting the right-hand side terms of the system (3)–(4) to zero since equilibrium solutions of a system with time delays are equivalent to the corresponding system with zero delay [18].

4.1 The disease-free equilibrium (DFE) and its analysis

The equilibria are acquired by equating the rate of change of all the compartment to zero. Consequently, we see that the system (3)–(4) has a unique DFE of the form $E_0\left(\frac{A}{\vartheta}, 0\right)$.

The characteristic equation of the system (3)–(4), evaluated at E_0 is as follows:

$$(\vartheta + \lambda)\left(\frac{\beta A}{\vartheta + \alpha A}e^{-\lambda\nu} - \vartheta - d - \zeta - a - \lambda\right) = 0. \tag{9}$$

Equation (9) has one negative real root $\lambda_1 = -\vartheta$, and other roots are the solutions of

$$\frac{\beta A}{\vartheta + \alpha A}e^{-\lambda\nu} - \vartheta - d - \zeta - a - \lambda = 0. \tag{10}$$

The term $\frac{\beta A e^{-\lambda\nu}}{(\vartheta+\alpha A)(\vartheta+d+\zeta+a)}$ at $\nu = 0$ is known as the basic reproduction number. The basic reproduction number is defined as "the expected number of secondary cases produced, in a completely susceptible population, by a typical infective individual" [11]. Hence, the basic reproduction number of our model is

$$R_0 = \frac{\beta A}{(\vartheta + \alpha A)(\vartheta + d + \zeta + a)}.$$

4.1.1 Analysis for $\mathbf{R_0 \neq 1}$

As mentioned above, Eq. (9) has one negative real root $\lambda_1 = -\vartheta$ and other roots satisfy the equation

$$\zeta(\lambda) := \lambda + \vartheta + d + \zeta + a - \frac{\beta A}{\vartheta + \alpha A}e^{-\lambda\nu} = 0 \tag{11}$$

I. Assuming that $R_0 > 1$, then

$$\zeta(0) = \vartheta + d + \zeta + a - \frac{\beta A}{\vartheta + \alpha A} < 0$$

Therefore, $\zeta(0) < 0$. Since $\lim_{\lambda\to\infty}\zeta(\lambda) = +\infty$, there exists at least one positive real root of (11) if $R_0 > 1$.

II. Let be $R_0 < 1$.
Our goal is to prove that for any values of the parameters, the characteristic roots cannot reach the imaginary axis. This means that for any values of the parameters and all delays ν, it happens that $\mathrm{Re}\,\lambda < 0$. On the contrary, suppose that $\mathrm{Re}\,\lambda \geq 0$. Note that

$$\mathrm{Re}\,\lambda = \frac{\beta A}{\vartheta+\alpha A}e^{-\mathrm{Re}\,\lambda\nu}\cos(Im\lambda)\nu - \vartheta - d$$
$$-\zeta - a < \frac{\beta A}{\vartheta+\alpha A} - \vartheta - d - \zeta - a < 0.$$

a contradiction to our assumption. Hence, if $R_0 < 1$ then all the roots of Eq. (11) must have the negative real part. Using the Routh–Hurwitz criterion, we have the following result:

Theorem 1 *The DFE E_0 is locally asymptotically stable if $R_0 < 1$ and unstable at $R_0 > 1$ for time lag $\nu \geq 0$.*

4.1.2 Analysis for $R_0 = 1$

We analysis the system (3)–(4) at $R_0 = 1$ for $\nu > 0$ and $\nu = 0$ separately.

$\nu > 0$: If $R_0 = 1$, then $\lambda = 0$ is a simple characteristic root of Eq. (11). Let $\lambda = \alpha + \iota\omega$ any of the other solutions, then (11) turns into:

$$\alpha + \iota\omega + \vartheta + d + \zeta + a - \frac{\beta A}{\vartheta+\alpha A}e^{-(\alpha+\iota\omega)\nu} = 0. \tag{12}$$

Applying Euler's formula and parting real and imaginary parts, we can write

$$\alpha + \vartheta + d + \zeta + a = \frac{\beta A}{\vartheta+\alpha A}\cos(\omega\nu)e^{-\alpha\nu}, \tag{13}$$

$$\omega = -\frac{\beta A}{\vartheta+\alpha A}e^{-\alpha\nu}\sin(\omega\nu). \tag{14}$$

Observing that $R_0 = 1$ implies $\beta A = (\vartheta+\alpha A)(\vartheta + d + \zeta + a)$. Moreover, if there exist roots satisfying both Eqs. (13) and (14), then they also satisfy the equation obtained by squaring and adding them member to member, we obtain

$$(\alpha+\vartheta+d+\zeta+a)^2+\omega^2 = (\vartheta+d+\zeta+a)^2e^{-2\alpha\nu}. \tag{15}$$

For Eq. (15) to be verified, we must have $\alpha \leq 0$. Therefore, DFE E_0 is linearly neutrally stable.

$\nu = 0$: Note that the system (3)–(4) evaluated at $R_0 = 1$ and bifurcation parameter $\beta = \beta^* = \frac{(\vartheta+\alpha A)(\vartheta+d+\zeta+a)}{A}$ have a simple zero (null) eigenvalue, and other eigenvalue is real and negative. Hence, when $R_0 = 1$, the DFE E_0 is a non-hyperbolic equilibrium. To investigate the stability of DFE at $R_0 = 1$, we use the bifurcation theory approach based on the center manifold theory [32]. For this, we redefine $S = x_1$ and $I = x_2$ then the system (3)–(4) can be rewritten as

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = A - \vartheta x_1 - \frac{\beta x_1 x_2}{1+\alpha x_1+\gamma x_2}, \tag{16}$$

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = \frac{\beta x_1 x_2}{1+\alpha x_1+\gamma x_2} - (\vartheta+d+\zeta)x_2 - \frac{ax_2}{1+bx_2} \tag{17}$$

Let $J(E_0, \beta^*)$ denotes the Jacobian matrix of the system (16)–(17) evaluated at $R_0 = 1$, $\beta = \beta^*$. Then

$$J(E_0,\beta^*) = \begin{bmatrix} -\vartheta & -\frac{\beta^* A}{(\vartheta+\alpha A)} \\ 0 & 0 \end{bmatrix}$$

The left eigenvector $u = [u_1, u_2]$ of the Jacobian matrix $J(E_0, \beta^*)$ is given by $u \cdot J(E_0, \beta^*)$. We obtain:

$$u_1 = 0, u_2 = 1.$$

The right eigenvector $w = (w_1, w_2)^{\mathrm{T}}$ of the Jacobian matrix $J(E_0, \beta^*)$ is by $J(E_0, \beta^*) \cdot w$ We obtain:

$$w_1 = -\frac{\beta^* A}{\vartheta(\vartheta+\alpha A)}, w_2 = 1.$$

Let $f_i, i = 1, 2$ denotes the right-hand side of the system (16)–(17). It can be checked that:
$\left(\frac{\partial^2 f_2}{\partial x_1 \partial x_2}\right)_{E_0} = \frac{\beta^*\vartheta^2}{(A\alpha+\vartheta)^2}$, $\left(\frac{\partial^2 f_2}{\partial x_2 \partial x_1}\right)_{E_0} = \frac{\beta^*\vartheta^2}{(A\alpha+\vartheta)^2}$,
$\left(\frac{\partial^2 f_2}{\partial x_2^2}\right)_{E_0} = 2ab - \frac{2A\beta^*\gamma\vartheta}{(A\alpha+\vartheta)^2}$, $\left(\frac{\partial^2 f_2}{\partial x_2 \partial \beta^*}\right)_{E_0} = \frac{A}{A\alpha+\vartheta}$.

Then, by [15], we obtain the bifurcation constants a_1 and b_1 as:

$$a_1 = \sum_{k,i,j=1}^{2} u_k w_i w_j \left(\frac{\partial^2 f_k}{\partial x_i \partial x_j}\right)_{E_0}$$
$$= \frac{2(ab(A\alpha+\vartheta)^3 - A\beta^*\vartheta(\beta^* + \gamma(A\alpha+\vartheta)))}{(A\,\alpha+\vartheta)^3},$$

$$b_1 = \sum_{k,i=1}^{2} u_k w_i \left(\frac{\partial^2 f_k}{\partial x_i \partial \beta^*}\right)_{E_0} = \frac{A}{A\alpha + \vartheta}.$$

It can be seen that b_1 is always positive. Therefore, by theorem 4.1 of Castillo-Chavez and Song [15], the local dynamics of the system (16)–(17) depends on the sign of the bifurcation constant a_1. Evaluating a_1 at bifurcation constant β^*, we get

$$a_1 = -\frac{2(\vartheta a^2 + (-Ab(A\alpha+\vartheta) + \vartheta(2d + A\gamma + 2(\zeta+\vartheta)))a + \vartheta(d+\zeta+\vartheta)(d+A\gamma+\zeta+\vartheta))}{A(A\alpha+\vartheta)}, = -\frac{g(a)}{A(A\alpha+\vartheta)},$$

where,

$$\begin{aligned} g(a) &= \vartheta a^2 + (\vartheta(2d + A\gamma + 2(\zeta+\vartheta)) \\ &\quad - Ab(A\alpha+\vartheta))a \\ &\quad + \vartheta(d+\zeta+\vartheta)(d+A\gamma+\zeta+\vartheta) \\ &= A_1 a^2 + A_2 a + A_3, \end{aligned} \tag{18}$$

with

$$\begin{aligned} A_1 &= \vartheta, \\ A_2 &= \vartheta(2d + A\gamma + 2(\zeta+\vartheta)) - Ab(A\alpha+\vartheta), \\ &= \vartheta(A(-b+\gamma) + 2(d+\zeta+\vartheta)) - A^2 b\alpha, \\ A_3 &= \vartheta(d+\zeta+\vartheta)(d+A\gamma+\zeta+\vartheta). \end{aligned}$$

It can be seen that A_1 and A_3 are always positive. Let D denotes the discriminant of the quadratic $g(a)$ given in Eq. (18). We obtain that

$$\begin{aligned} D &= A_2^2 - 4A_1A_3 \\ &= A^2\gamma^2\vartheta^2 + Ab(A\alpha+\vartheta)(A^2 b\alpha + \vartheta(A(b-2\gamma) \\ &\quad - 4(d+\zeta+\vartheta))). \end{aligned} \tag{19}$$

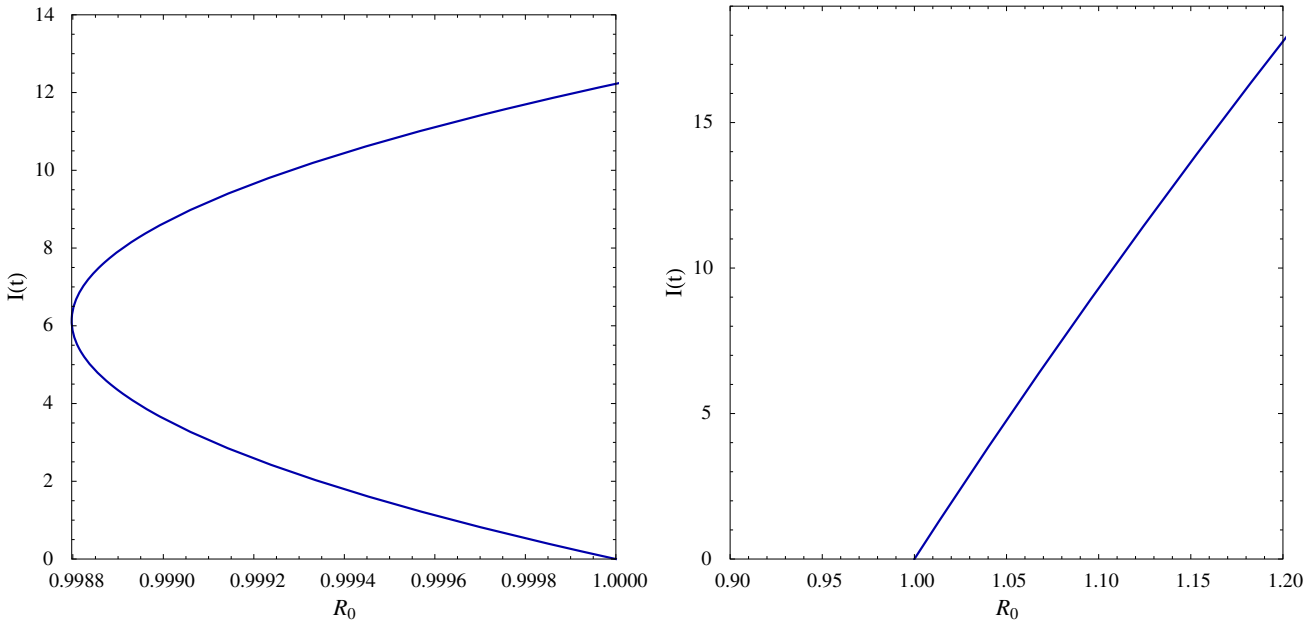
Let a_1^* and a_2^* be two positive roots of the quadratic Eq. (18), given by $a_1^* = \frac{-A_2-\sqrt{D}}{2A_1}$, and $a_2^* = \frac{-A_2+\sqrt{D}}{2A_1}$.

With the help of theorem 4.1 [15], it can be concluded that if $g(a) > 0$, a forward bifurcation occurs, whereas if $g(a) < 0$, the system (16)–(17) exhibits a backward bifurcation. We discuss these two cases separately.

1. Backward bifurcation
The system (16)–(17) exhibits a backward bifurcation when $g(a) < 0$. The existence of a backward bifurcation implies a range which shows the region of coexistence of DFE and two endemic equilibria: a smaller endemic equilibrium (i.e., there is less number of infectives), and a larger endemic equilibrium (i.e., there are a large number of infectives). In this case, the smaller endemic equilibrium is unstable, whereas larger endemic equilibrium is stable. Thus, the occurrence of backward bifurcation shows that R_0 less unity does not confirm the eradication of infection. Therefore, we now find the range of a for which backward bifurcation occurs. Now, $g(a) < 0$ if and only if the discriminant $D > 0$, $A_2 < 0$, and $a_1^* < a < a_2^*$. Thus, the conditions obtained for the occurrence of backward bifurcation are stated below:

$$\begin{cases} A_2 < 0, \\ D > 0, \\ a_1^* < a < a_2^*. \end{cases} \tag{20}$$

2. Forward bifurcation
The presence of forward bifurcation confirms the eradication of the disease when $R_0 < 1$. It shows that the DFE is stable when $R_0 < 1$ and if R_0 cross unity, then the model admits a unique stable endemic equilibrium. From this analysis, we will be able to know that if there is a stable coexistence equilibrium bifurcating from E_0 and E_0 changes its stability from stable to unstable. This behavior is known as forward bifurcation. Therefore, it is very important to find the range for which forward bifurcation occurs. We see that the system (16)–(17) exhibits a forward bifurcation if $g(a) > 0$. $g(a) > 0$ if either $A_2 > 0$ or $D > 0$, $A_2 < 0$, and $a < a_1^*$ or $a > a_2^*$. These cases are described below:



(a) Backward bifurcation. **(b)** Forward bifurcation.

Fig. 1 Bifurcation diagram in the plane (R_0, I). **a** The case when $D > 0$, $A_2 < 0$ and $a_1^* = 0.00644085 < a < 0.847559 = a_2^*$, showing backward bifurcation. **b** The case: $D > 0$, $A_2 < 0$ and $a = 0.005 < a_1^* = 0.00644085$, showing forward bifurcation

$$\begin{cases} A_2 > 0 \text{ or, } b < \dfrac{(2d\vartheta + A\gamma\vartheta + 2\zeta\vartheta + 2\vartheta^2)}{A(A\alpha + \vartheta)}. \end{cases} \tag{21}$$

or

$$\begin{cases} A_2 < 0, \\ D > 0, \\ a < a_1^* \text{ or } a > a_2^*. \end{cases} \tag{22}$$

These bifurcations are illustrated in Fig. 1a, b. We consider the parameter values $A = 5, \alpha = 1, \vartheta = 0.05, d = 0.001, \gamma = 0.01, \zeta = 0.002, b = 0.002$. Most of the parameter values are taken from [33]. At these values of parameters, we obtain that $D = 0.0420559$ and $A_2 = -0.0427$. According to (20), we have a backward bifurcation at $\beta = \beta^*$, for $a_1^* = 0.00644085 < a < 0.847559 = a_2^*$. Setting $a = 0.8$, the system (16)–(17) exhibits a backward bifurcation as shown in Fig. 1a. A forward bifurcation occurs by choosing $a < a_1^*$ or $a > a_2^*$, as stated in (22). Setting $a = 0.005 < a_1^* = 0.00644085$, we see the occurrence of forward bifurcation as shown in Fig. 1b.

Now, in order to satisfy inequality (21), we use the parameter $A = 5, \alpha = 0.004, \gamma = 0.01, \vartheta = 0.05, d = 0.001, \zeta = 0.002, b = 0.002, a = 0.005$.

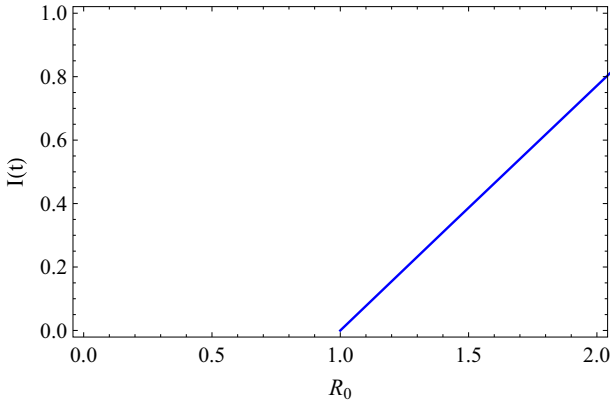


Fig. 2 Forward bifurcation diagram in the plane (R_0, I), when the inequality (21) holds

At these parameter values, we see that the inequality (21), $b = 0.002 < \frac{(2d\vartheta + A\gamma\vartheta + 2\zeta\vartheta + 2\vartheta^2)}{A(A\alpha + \vartheta)} = 0.0222857$ is satisfied and forward bifurcation occurs as shown in Fig. 2.

Theorem 2 *The disease-free equilibrium* $E_0 = \left(\frac{A}{\vartheta}, 0\right)$ *of the system* (3)–(4) *at* $R_0 = 1$ *is linearly neutrally stable for* $\nu > 0$.

Theorem 3 *When* $R_0 = 1$, *then the system* (16)–(17) *exhibits backward (forward) bifurcation at disease-free equilibrium* E_0 *if and only if* $g(a) < 0$ (> 0).

Remark 1 The theorem above states that at $\nu = 0$ and for values of R_0 greater than 1 but close to 1, the system admits a unique infected equilibrium, which is locally asymptotically stable.

4.2 Endemic equilibrium and its stability

To investigate the stability of the system (3)–(4) around the endemic equilibrium $E_e(S^*, I^*)$, we equate the second equation of the system (3)–(4) to zero:

$$\frac{\beta S^* I^*}{1+\alpha S^*+\gamma I^*} - (\vartheta + d + \zeta) I^* - \frac{a I^*}{1+bI^*} = 0. \tag{23}$$

After solving Eq. (23), we get S^* as follows:

$$S^* = \frac{(1+\gamma I^*)(a+(1+bI^*)p)}{(\beta - \alpha a - \alpha p)+bI^*(\beta-\alpha p)}, \; p = \vartheta + d + \zeta. \tag{24}$$

We see that $S^* > 0$ if and only if

$$\beta - \alpha a - \alpha p > 0 \quad \text{i.e.,} \quad \beta - \alpha p > \alpha a > 0. \tag{25}$$

Plug in the value of S^* from Eq. (24), into Eq. (3), we get the cubic equation in I^* as follows

$$H_0 + H_1 I^* + H_2 I^{*2} + H_3 I^{*3} = 0, \tag{26}$$

where

$$\begin{aligned}
H_0 &= (\vartheta + \alpha A)(\vartheta + d + \zeta + a)(1 - R_0),\\
H_1 &= -a^2\alpha - (d+\zeta+\vartheta)(-\beta - 2b\vartheta - \gamma\vartheta\\
&\quad + \alpha(d+\zeta+\vartheta)) + a(Ab\alpha + \beta + b\vartheta + \gamma\vartheta\\
&\quad + Ab(-\beta + \alpha(d+\zeta+\vartheta)),\\
H_3 &= b^2(d+\zeta+\vartheta)(\gamma\vartheta + \beta - \alpha(\vartheta + d + \zeta)).
\end{aligned} \tag{27}$$

Note that $H_3 > 0$ according to the condition given in Eq. (25).

Theorem 4 *If $R_0 > 1$, then there is either a unique or three positive endemic equilibria if all equilibria are simple roots.*

Proof Suppose that $R_0 > 1$. From (27), it follows that H_0 is negative. From equation (26), we have a third-degree polynomial in I^*:

$$H_0 + H_1 I^* + H_2 I^{*2} + H_3 I^{*3} = 0.$$

Since H_3 is positive and H_0 is negative, we have the following possibilities of signs of H_1 and H_2:

$$\begin{aligned}
&\text{(i)}\, H_1 > 0, \text{ and } H_2 < 0,\\
&\text{(ii)}\, H_1 < 0, \text{ and } H_2 > 0,\\
&\text{(iii)}\, H_1 > 0, \text{ and } H_2 > 0,\\
&\text{(iv)}\, H_1 < 0, \text{ and } H_2 < 0.
\end{aligned} \tag{28}$$

Using Descartes' rule of signs [34], it follows that Eq. (26) has either three positive roots or a unique positive root. Hence, the proof. □

To analyze the local stability of endemic equilibrium E_e, we obtain the characteristic equation of the system (3)–(4) at E_e as follows:

$$\lambda^2 + p_0\lambda + q_0 + (p_1\lambda + q_1)e^{-\lambda\nu} = 0, \tag{29}$$

where

$$\begin{aligned}
p_0 &= 2\vartheta + d + \zeta + \frac{a}{(bI^*+1)^2},\\
q_0 &= \vartheta(\vartheta + d + \zeta) + \frac{a\vartheta}{(bI^*+1)^2},\\
p_1 &= \frac{\beta(-S^*(1+S^*\alpha) + I^*(1+I^*\gamma))}{(1+S^*\alpha+I^*\gamma)^2},\\
q_1 &= \frac{\beta(aI^*(1+I^*\gamma) + (1+bI^*)^2(I^*(1+I^*\gamma)(d+\zeta) + (-S^*(1+S^*\alpha) + I^*(1+I^*\gamma))\vartheta))}{(bI^*+1)^2(1+S^*\alpha+y\gamma)^2}.
\end{aligned}$$

Theorem 5 *At $\nu = 0$, E_e is locally asymptotically stable if $p_0 + q_0 > 0$, and $p_1 + q_1 > 0$.*

Proof The characteristic equation of the system (3)–(4) evaluated at E_e and $\nu = 0$ is

$$\lambda^2 + (p_0 + p_1)\lambda + (q_0 + q_1) = 0. \tag{30}$$

Using the Routh–Hurwitz criterion, it can be said that the endemic equilibrium E_e of the system (3)–(4) is locally asymptotically stable when $\nu = 0$ if $p_0 + p_1 > 0$ and $p_1 + q_1 > 0$ are satisfied simultaneuously. □

Theorem 6 *For $\nu > 0$, E_e is locally asymptotically stable if both $p_0^2 - 2q_0 - p_1^2 > 0$ and $q_0^2 - q_1^2 > 0$ are satisfied simultaneuously.*

Proof For $\nu > 0$, if instability occurs for a specific value of the time delay ν, then a characteristic root of (29) must cross the imaginary axis [7].

Assume that $\lambda = \iota\xi, \xi > 0$ is a root of the characteristic Eq. (29). Substituting $\lambda = \iota\xi$ into Eq. (29), we get

$$\begin{aligned}&(-\xi^2 + q_0 + p_1\xi \sin\xi\nu + q_1 \cos\xi\nu)\\ &\quad + \iota(p_0\xi + p_1\xi\cos\xi\nu - q_1\cos\xi\nu) = 0.\end{aligned} \tag{31}$$

On separating real and imaginary part of Eq. (31), we get:

$$p_1\xi \sin\xi\nu + q_1\cos\xi\nu = \xi^2 - q_0, \tag{32}$$

$$p_1\xi\cos\xi\nu - q_1\cos\xi\nu = -p_0\xi. \tag{33}$$

Eliminating ν by squaring and adding Eqs. (32) and (33), we obtain a polynomial in ξ as follows:

$$\xi^4 + (p_0^2 - 2q_0 - p_1^2)\xi^2 + (q_0^2 - q_1^2) = 0. \tag{34}$$

Letting $\xi^2 = x$, $p_0^2 - 2q_0 - p_1^2 = A$, and $q_0^2 - q_1^2 = B$, Eq. (34) becomes

$$x^2 + Ax + B = 0. \tag{35}$$

Using the Routh–Hurwitz Criterion, it can be observed that Eq. (35) has roots with negative real part, if $A > 0$ and $B > 0$ are satisfied simultaneously. It contradicts the assumption that Eq. (29) has a root $\lambda = \iota\xi$. Hence, the endemic equilibrium E_e of the system (3)–(4) is locally asymptotically stable for $\nu > 0$ if $p_0^2 - 2q_0 - p_1^2 > 0$ and $q_0^2 - q_1^2 > 0$ are satisfied simultaneuously. □

4.2.1 Hopf bifurcation

If $q_0^2 - q_1^2 < 0$, then there is a unique positive ω_0 satisfying Eq. (35), i.e., there is a single pair of purely imaginary roots $\pm\iota\omega_0$ to Eq. (29).

Using Eqs. (32) and (33), we obtain ν_k corresponding to ω_0

$$\nu_k = \frac{1}{\omega_0}\arccos\left(\frac{\omega_0^2(q_1 - p_0p_1) - q_0q_1}{p_1^2\omega_0^2 + q_1^2}\right) + \frac{2k\pi}{\omega_0}, \quad k = 0, 1, 2\ldots \tag{36}$$

For $\nu < \nu_0$, the endemic equilibrium E_0 is stable if transversality condition holds, That is, $\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{Re}\,\lambda)\Big|_{\lambda=i\omega_0} \neq 0$.

Differentiating Eq. (29) with respect to ν, we get

$$\begin{aligned}&\left(2\lambda + p_0 + p_1e^{-\lambda\nu} - (p_1\lambda + q_1)\nu e^{-\lambda\nu}\right)\frac{\mathrm{d}\lambda}{\mathrm{d}\nu}\\ &= \lambda(p_1\lambda + q_1)e^{-\lambda\nu}\\ &\left[\frac{\mathrm{d}\lambda}{\mathrm{d}\nu}\right]^{-1} = \frac{2\lambda + p_0 + p_1e^{-\lambda\nu} - (p_1\lambda + q_1)\nu e^{-\lambda\nu}}{\lambda(p_1\lambda + q_1)e^{-\lambda\nu}}\\ &= \frac{2\lambda + p_0}{\lambda(p_1\lambda + q_1)e^{-\lambda\nu}} + \frac{p_1}{\lambda(p_1\lambda + q_1)} - \frac{\nu}{\lambda}\\ &= \frac{2\lambda + p_0}{-\lambda(\lambda^2 + p_0\lambda + q_0)} + \frac{p_1}{\lambda(p_1\lambda + q_1)} - \frac{\nu}{\lambda}.\\ &\frac{\mathrm{d}}{\mathrm{d}\nu}(\mathrm{Re}\,\lambda)\Big|_{\lambda=\iota\omega_0} = \mathrm{Re}\left(\frac{\mathrm{d}\lambda}{\mathrm{d}\nu}\right)^{-1}\Big|_{\lambda=\iota\omega_0}\\ &= \mathrm{Re}\left(\frac{2\iota\omega_0 + p_0}{-\iota\omega_0(-\omega_0^2 + \iota p_0\omega_0 + q_0)}\right.\\ &\quad \left. + \frac{p_1}{-p_1{\omega_0}^2 + \iota q_1\omega_0} + \frac{\iota\nu}{\omega_0}\right)\\ &= \frac{1}{\omega_0}\left(\frac{2\omega_0(\omega_0^2 - q_0) + p_0^2\omega_0}{(p_0\omega_0)^2 + (\omega_0^2 - q_0)^2} - \frac{p_1^2\omega_0}{(p_1\omega_0)^2 + q_1^2}\right)\\ &= \frac{2(\omega_0^2 - q_0) + p_0^2}{(p_0\omega_0)^2 + (\omega_0^2 - q_0)^2} - \frac{p_1^2}{(p_1\omega_0)^2 + q_1^2}\\ &= \frac{2\omega_0^2 + (p_0^2 - 2q_0 - p_1^2)}{p_1^2\omega_0^2 + q_1^2} \text{ (Using Eqs. (32) and (33),}\\ &\quad (p_1\omega_0)^2 + q_1^2 = (p_0\omega_0)^2 + (\omega_0^2 - q_0)^2).\end{aligned} \tag{37}$$

Under the condition $p_0^2 - 2q_0 - p_1^2 > 0$, we have $\frac{\mathrm{d}}{\mathrm{d}\nu}(\mathrm{Re}\,\lambda)\Big|_{\lambda=\iota\omega_0} > 0$. Hence, transversality condition holds and Hopf bifurcation occurs at $\nu = \nu_0$ and ω_0. Thus, we state the following theorem:

Theorem 7 *If the condition $q_0^2 - q_1^2 < 0$ is satisfied, then the endemic equilibrium E_e of the system (3)–(4) is asymptotically stable for $\nu \in [0, \nu_0)$ and it undergoes Hopf bifurcation at $\nu = \nu_0$.*

4.3 Global stability analysis

This subsection is devoted to analyze the global stability behavior of the disease-free and endemic equilibria.

4.3.1 Global stability of disease-free equilibrium

For the global stability of the disease-free equilibrium $E_0\left(\frac{A}{\vartheta}, 0\right)$, we assume the following postulates:

P4. $\Phi_I'(S(t), 0)$ is increasing with respect to $S(t) > 0$.

P5. $\frac{\Phi_I'(S_0,0)}{\Phi_I'(S(t),0)} < 1$ for $S(t) > S_0$, $\frac{\Phi_I'(S_0,0)}{\Phi_I'(S(t),0)} > 1$ for $S(t) \in (0, S_0)$, where $S_0 = \frac{A}{\vartheta}$.

P6. $\Phi(S(t), I(t)) \leq I(t) \left(\left(\frac{\partial \Phi(S(t),I(t))}{\partial I}\right)_{(S_0,0)} - \left(\frac{\partial H(I)}{\partial I}\right)_{I=0}\right) + H(I(t)), I(t) > 0,$
where $H(I(t)) = (\vartheta + d + \zeta)I(t) + h(I(t))$

Under these assumptions, we have the following theorem:

Theorem 8 *Assume that (P1.)–(P6.) are satisfied then, the disease-free equilibrium $E_0(S_0, 0)$ of the system (3)–(4) is globally asymptotically stable for any $\nu \geq 0$ if $R_0 \leq 1$.*

Proof From the conditions (P1.) and (P2.), it follows that the disease-free equilibrium $E_0(S_0, 0)$ is the only equilibrium of the system (3)–(4). We define the following Lyapunov functional:

$$L(t) = L_1(t) + L_2(t)$$

where

$$L_1(t) = S(t) - S_0 - \int_{S_0}^{S(t)} \lim_{I\to 0^+} \frac{\Phi(S_0, I(t))}{\Phi(\sigma, I(t))} \mathrm{d}\sigma + I(t),$$

$$L_2(t) = \int_0^{\nu} \Phi(S(t-\rho), I(t-\rho))\mathrm{d}\rho.$$

The postulates (P1.)–(P3.) imply that $L_1(t)$ is defined and continuously differentiable for all $S(t) > 0$ and $I(t) > 0$, and $L(t) = 0$ at $E_0(S_0, 0)$. We will show that $\frac{\mathrm{d}L(t)}{\mathrm{d}t} \leq 0$ for all $t \geq 0$. For this, first we calculate $\frac{\mathrm{d}L_1(t)}{\mathrm{d}t}$ as follows:

$$\frac{\mathrm{d}L_1(t)}{\mathrm{d}t} = \left(1 - \lim_{I\to 0^+} \frac{\Phi(S_0, I(t))}{\Phi(S, I(t))}\right) S'(t) + I'(t)$$

Since, $A - \vartheta S = -\vartheta\left(S - \frac{A}{\vartheta}\right) = -\vartheta(S - S_0)$, we get

$$\begin{aligned}\frac{\mathrm{d}L_1(t)}{\mathrm{d}t} = &\left(1 - \lim_{I\to 0^+} \frac{\Phi(S_0, I(t))}{\Phi(S, I(t))}\right)(-\vartheta(S - S_0) \\ &- \Phi(S(t-\nu), I(t-\nu))) \\ &+ \Phi(S(t-\nu), I(t-\nu)) - H(I(t)).\end{aligned}$$

We now calculuate $\frac{\mathrm{d}L_2(t)}{\mathrm{d}t}$ as follows:

$$\frac{\mathrm{d}L_2(t)}{\mathrm{d}t} = -\Phi(S(t-\nu), I(t-\nu)) + \Phi(S(t), I(t))$$

Thus, $\frac{\mathrm{d}L}{\mathrm{d}t}$ is given by:

$$\begin{aligned}\frac{\mathrm{d}L(t)}{\mathrm{d}t} &= \frac{\mathrm{d}L_1(t)}{\mathrm{d}t} + \frac{\mathrm{d}L_2(t)}{\mathrm{d}t} \\ &= \left(1 - \lim_{I\to 0^+} \frac{\Phi(S_0, I(t))}{\Phi(S, I(t))}\right)(-\vartheta(S - S_0) \\ &\quad -\Phi(S(t-\nu), I(t-\nu))) \\ &\quad + \Phi(S(t-\nu), I(t-\nu)) \\ &\quad - H(I) - \Phi(S(t-\nu), I(t-\nu)) + \Phi(S(t), I(t)) \\ &= \vartheta\left(1 - \lim_{I\to 0^+} \frac{\Phi(S_0, I(t))}{\Phi(S, I(t))}\right)(S_0 - S(t)) \\ &\quad + \Phi(S(t-\nu), I(t-\nu))\left(\lim_{I\to 0^+} \frac{\Phi(S_0, I(t))}{\Phi(S, I(t))} - 1\right) \\ &\quad + \Phi(S(t), I(t)) - H(I).\end{aligned}$$

Further, (P4.)–(P6.) implies that

$$\begin{aligned}\frac{\mathrm{d}L(t)}{\mathrm{d}t} &\leq \vartheta\left(1 - \frac{\Phi_I'(S_0, 0)}{\Phi_I'(S(t), 0)}\right)(S_0 - S(t)) \\ &\quad + \Phi(S(t-\nu), I(t-\nu))\left(\frac{\Phi_I'(S_0, 0)}{\Phi_I'(S(t), 0)} - 1\right) \\ &\quad + I(t)\left(\left(\frac{\partial \Phi(S(t), I(t))}{\partial I}\right)_{(S_0,0)} \right. \\ &\quad \left. - \left(\frac{\partial H(I)}{\partial I}\right)_{I=0}\right) \\ &= \vartheta\left(1 - \frac{\Phi_I'(S_0, 0)}{\Phi_I'(S(t), 0)}\right)(S_0 - S(t)) \\ &\quad + \Phi(S(t-\nu), I(t-\nu))\left(\frac{\Phi_I'(S_0, 0)}{\Phi_I'(S(t), 0)} - 1\right) \\ &\quad + I(t)\left(\frac{\partial H(I)}{\partial I}\right)_{I=0}(R_0 - 1).\end{aligned}$$

Thus, for $R_0 \leq 1$, $\frac{\mathrm{d}L(t)}{\mathrm{d}t} \leq 0$ for all $t \geq 0$. Also, $\frac{\mathrm{d}L(t)}{\mathrm{d}t} = 0$ if $S(t) = S_0$.

Hence, it follows from the system (3)–(4) that the largest invariant set $\left\{\left(S(t), I(t) \in R_+^2 \middle| \frac{\mathrm{d}L(t)}{\mathrm{d}t} = 0\right)\right\}$ is singleton set $E_0(S_0, 0)$. From the Lyapunov–LaSalle asymptotic stability theorem [35–37], E_0 is the only equilibrium of the system (3)–(4) and globally asymptotically stable. □

4.3.2 Global stability of endemic equilibrium

We now study the global stability of the endemic equilibrium $E_e(S^*, I^*)$ of the system (3)–(4) by Lyapunov direct method. For this, the following hypothesis has been proposed:

P7. $\left(\frac{\Phi(S^*,I^*)}{\Phi(S(t),I^*)} - \frac{I^*}{I(t)}\right) \leq 0$; $\left(\frac{\Phi(S(t),I(t))}{\Phi(S^*,I^*)} - 1\right) \leq 0$; $\left(\frac{\Phi(S(t),I^*)}{\Phi(S(t),I(t))} - \frac{I(t)}{I^*}\right) \leq 0$ for $I \geq I^*$.

P8. $\left(\frac{h(I^*)}{h(I(t))} - \frac{I^*}{I(t)}\right)\left(\frac{I(t)}{I^*} - 1\right) \leq 0$ for $I \geq I^*$.

Under these conditions, we have the following theorem:

Theorem 9 *Suppose that the conditions (P1.)–(P3.) and (P7.)–(P8.) are satisfied then the endemic equilibrium $E_e(S^*, I^*)$ of the system (3)–(4) is globally asymptotically stable for any $\nu \geq 0$ if $R_0 > 1$.*

Proof We assume the following Lyapunov functional

$$M(t) = M_1(t) + M_2(t)$$

where,

$$M_1(t) = S(t) - S^* - \int_{S^*}^{S(t)} \frac{\Phi(S^*, I^*)}{\Phi(\phi, I^*)} \mathrm{d}\phi + I(t) - I^* - I^* \ln \frac{I(t)}{I^*},$$

$$M_2(t) = \Phi(S^*, I^*) \int_0^{\nu} \left(\frac{\Phi(S(t-\theta), I(t-\theta))}{\Phi(S^*, I^*)} - 1 - \ln \frac{\Phi(S(t-\theta), I(t-\theta))}{\Phi(S^*, I^*)}\right) d\theta.$$

By (P1.)–(P3.), $M(t) = M_1(t) + M_2(t)$ is defined and continuously differentiable for all $S(t), I(t) > 0$ and $M(0) = 0$ at $E_e(S^*, I^*)$.

At $E_e(S^*.I^*)$, the system (3)–(4) has

$$A - \vartheta S^* = \Phi(S^*, I^*) = (\vartheta + d + \zeta)I^* + h(I^*) \tag{38}$$

The time derivative of $M_1(t)$ along the solution of system (3)–(4) is given by

$$\begin{aligned}
\frac{\mathrm{d}M_1(t)}{\mathrm{d}t} &= \left(1 - \frac{\Phi(S^*, I^*)}{\Phi(S(t), I^*)}\right) S'(t) \\
&\quad + \left(1 - \frac{I^*}{I(t)}\right) I'(t) \\
&= \left(1 - \frac{\Phi(S^*, I^*)}{\Phi(S(t), I^*)}\right) \big(\vartheta S^* - \vartheta S(t) \\
&\quad + \Phi(S^*, I^*) - \Phi(S(t-\nu), I(t-\nu))\big) \\
&\quad + \left(1 - \frac{I^*}{I(t)}\right) (\Phi(S(t-\nu), I(t-\nu)) \\
&\quad - \frac{(\Phi(S^*, I^*) - h(I^*))}{I^*} I(t) - h(I(t))\Bigg) \\
&= \vartheta(S^* - S(t)) \left(1 - \frac{\Phi(S^*, I^*)}{\Phi(S(t), I^*)}\right) \\
&\quad + \Phi(S^*, I^*) \left(1 - \frac{\Phi(S^*, I^*)}{\Phi(S(t), I^*)} \right. \\
&\quad \left. + \frac{\Phi(S(t-\nu), I(t-\nu))}{\Phi(S(t), I^*)}\right) \\
&\quad - \frac{\Phi(S(t-\nu), I(t-\nu))}{I(t)} I^* \\
&\quad + \Phi(S^*, I^*) \left(1 - \frac{I(t)}{I^*}\right) \\
&\quad + \left(h(I^*) - \frac{h(I(t))}{I(t)} I^*\right) \left(\frac{I(t)}{I^*} - 1\right)
\end{aligned}$$

Further, the time derivative of $M_2(t)$ is given by

$$\begin{aligned}
\frac{\mathrm{d}M_2(t)}{\mathrm{d}t} &= \Phi(S^*, I^*) \frac{\mathrm{d}}{\mathrm{d}t} \int_0^{\nu} \left(\frac{\Phi(S(t-\theta), I(t-\theta))}{\Phi(S^*, I^*)} \right. \\
&\quad \left. - 1 - \ln \frac{\Phi(S(t-\theta), I(t-\theta))}{\Phi(S^*, I^*)}\right) \mathrm{d}\theta. \\
&= \Phi(S(t), I(t)) - \Phi(S(t-\nu), I(t-\nu)) \\
&\quad + \Phi(S^*, I^*) \ln \frac{\Phi(S(t-\nu), I(t-\nu))}{\Phi(S(t), I(t))}.
\end{aligned}$$

Then, we have

$$\begin{aligned}
\frac{\mathrm{d}M(t)}{\mathrm{d}t} &= \frac{\mathrm{d}M_1(t)}{\mathrm{d}t} + \frac{\mathrm{d}M_2(t)}{\mathrm{d}t} \\
&= \vartheta(S^* - S(t)) \left(1 - \frac{\Phi(S^*, I^*)}{\Phi(S(t), I^*)}\right) \\
&\quad + \Phi(S^*, I^*) \left(1 - \frac{\Phi(S^*, I^*)}{\Phi(S(t), I^*)}\right)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{\Phi(S(t-\nu),I(t-\nu))}{\Phi(S(t),I^*)}\Big)\\
&-\frac{\Phi(S(t-\nu),I(t-\nu))}{I(t)}I^*\\
&+\Phi(S^*,I^*)\left(1-\frac{I(t)}{I^*}\right)\\
&+\left(h(I^*)-\frac{h(I(t))}{I(t)}I^*\right)\left(\frac{I(t)}{I^*}-1\right)\\
&+\Phi(S(t),I(t))-\Phi(S(t-\nu),I(t-\nu))\\
&+\Phi(S^*,I^*)\ln\frac{\Phi(S(t-\nu),I(t-\nu))}{\Phi(S(t),I(t))}\\
&=\vartheta(S^*-S(t))\left(1-\frac{\Phi(S^*,I^*)}{\Phi(S(t),I^*)}\right)\\
&+\Phi(S^*,I^*)\left(1-\frac{\Phi(S^*,I^*)}{\Phi(S(t),I^*)}+\ln\frac{\Phi(S^*,I^*)}{\Phi(S(t),I^*)}\right)\\
&+\Phi(S(t-\nu),I(t-\nu))\left(\frac{\Phi(S^*,I^*)}{\Phi(S(t),I^*)}-\frac{I^*}{I(t)}\right)\\
&+\Phi(S^*,I^*)\left(1-\frac{I(t)}{I^*}+\ln\frac{I(t)}{I^*}\right)\\
&+\left(\frac{h(I^*)}{h(I(t))}-\frac{I^*}{I(t)}\right)\left(\frac{I(t)}{I^*}-1\right)h(I(t))\\
&+\Phi(S^*,I^*)\Big(1-\frac{\Phi(S(t-\nu),I(t-\nu))}{\Phi(S(t),I(t))}\\
&\cdot\frac{\Phi(S(t),I^*)}{\Phi(S^*,I^*)}\cdot\frac{I^*}{I(t)}\\
&+\ln\frac{\Phi(S(t-\nu),I(t-\nu))}{\Phi(S(t),I(t))}\cdot\frac{\Phi(S(t),I^*)}{\Phi(S^*,I^*)}\cdot\frac{I^*}{I(t)}\Big)\\
&+\Phi(S^*,I^*)\left(\frac{\Phi(S(t),I(t))}{\Phi(S^*,I^*)}-1\right)\\
&+\Phi(S(t-\nu),I(t-\nu))\\
&\cdot\frac{I^*}{I(t)}\left(\frac{\Phi(S(t),I^*)}{\Phi(S(t),I(t))}-\frac{I(t)}{I^*}\right).
\end{aligned}
$$

The function $\Phi(S(t),I(t))$ is monotonically increasing for all $S(t)>0$. Therefore,

$$\big(S^*-S(t)\big)\left(1-\frac{\Phi(S^*,I^*)}{\Phi(S(t),I^*)}\right)\le 0. \tag{39}$$

The function $f(x)=1-x+\ln x,(x>0)$ has its global maximum $f(1)=0$. Hence, for $x>0$, $f(x)\le 0$ and we have the following inequalities:

$$
\begin{aligned}
&\left(1-\frac{\Phi(S^*,I^*)}{\Phi(S(t),I^*)}+\ln\frac{\Phi(S^*,I^*)}{\Phi(S(t),I^*)}\right)\le 0,\\
&\left(1-\frac{I(t)}{I^*}+\ln\frac{I(t)}{I^*}\right)\le 0,\\
&\left(1-\frac{\Phi(S(t-\nu),I(t-\nu))}{\Phi(S(t),I(t))}\cdot\frac{\Phi(S(t),I^*)}{\Phi(S^*,I^*)}\cdot\frac{I^*}{I(t)}\right)\\
&\quad+\ln\frac{\Phi(S(t-\nu),I(t-\nu))}{\Phi(S(t),I(t))}\cdot\frac{\Phi(S(t),I^*)}{\Phi(S^*,I^*)}\cdot\frac{I^*}{I(t)}\\
&\le 0.
\end{aligned}
\tag{40}
$$

Hence, by (P7.)–(P8.) and inequalities (39)–(40), it follows that $\frac{\mathrm{d}M(t)}{\mathrm{d}t}\le 0$ for all $S(t)\ge 0$, $I(t)\ge 0$. It can be verified that the largest invariant in $\left\{(S(t),I(t))\in R_+^2:\frac{\mathrm{d}M(t)}{\mathrm{d}t}=0\right\}$ is singleton $\{E_e\}$. By the Lyapunov–LaSalle asymptotic stability theorem [35–37], E_e is globally asymptotically stable. □

5 Numerical simulations

In this section, we present numerical simulations of the proposed model. Parameter description and values are given in Table 2, and the total population size is considered as $N=100$.

Initial conditions are given by

$$S(0)=90,\quad I(0)=8,\quad R(0)=2.$$

Figure 3 shows the behavior of susceptible–infected–recovered populations at time delay $\nu=1$. At the parameters values given in Table 2, we obtained that $S^*=23.5845$, $I^*=42.5367$. From this S^*, and I^*, R^* can be evaluated as: $R^*=N-S^*-I^*=100-23.5845-42.5367=33.8788$. This is also verified from Fig. 3. This figure shows that as time increases, all the subpopulations reach to their endemic equilibrium $(S^*,I^*,R^*)=(23.5845,42.5367,33.8788)$.

Figure 4 shows the effect of time delay on the population with respect to time t. In this figure, susceptible, infected population, and recovered populations are being drawn for the time delay $\nu=0$, 3, and 6, respectively. It can be seen that with the increased values of time delay ν, the infected individuals increases, whereas susceptible and recovered individuals decrease, and finally settle down to its endemic equilibrium (23.5845, 42.5367, 33.8788). Thus, the time delay is an influential parameter in the transmission dynamics of infection.

Figure 5 depicts the infected population with and without the effect of the Holling type II treatment rate. This figure shows the major difference between the progression of treated infected individuals and non-treated infected individuals. When Holling type II treatment rate is considered, then the infected individuals first increase, and then stabilize to steady state at a lower

Table 2 List of parameters

Parameter	Description	Value	References
A	Recruitment rate	5	Assumed
α	Inhibitions taken by susceptibles	0.004	[33]
β	Transmission rate	0.0045	Assumed
ϑ	Natural death rate	0.05	[33]
d	Disease-induced death rate	0.001	[33]
γ	Inhibitions taken by infectives	0.002	[33]
ζ	Recovery rate	0.002	[33]
a	Cure rate	0.04	Assumed
b	Rate of limitation in treatment availability	0.002	Assumed

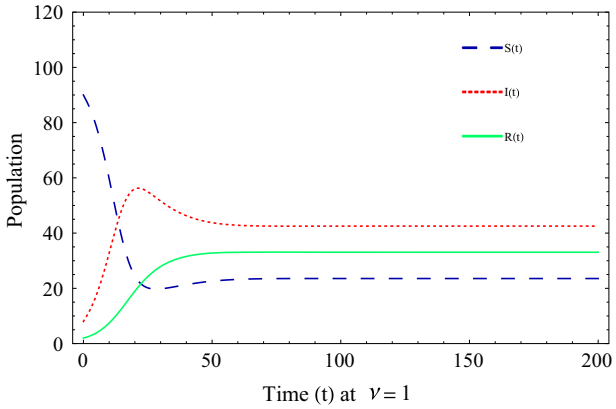


Fig. 3 Behavior of susceptible–infected–recovered populations at time delay $\nu = 1$

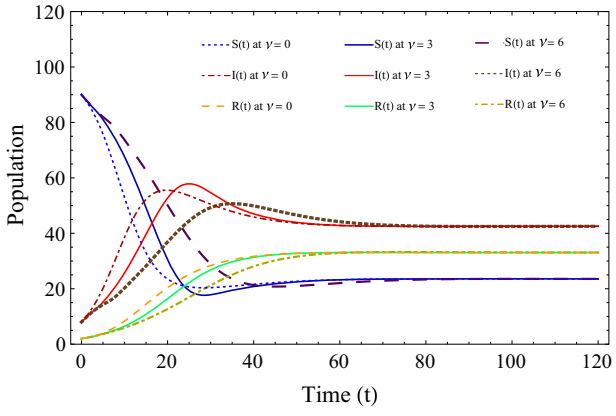


Fig. 4 Impact of time delay ν on susceptible, infected and recovered populations

level, whereas in the absence of treatment rate, infections increase at a higher rate and then settle down to steady-state with a big difference.

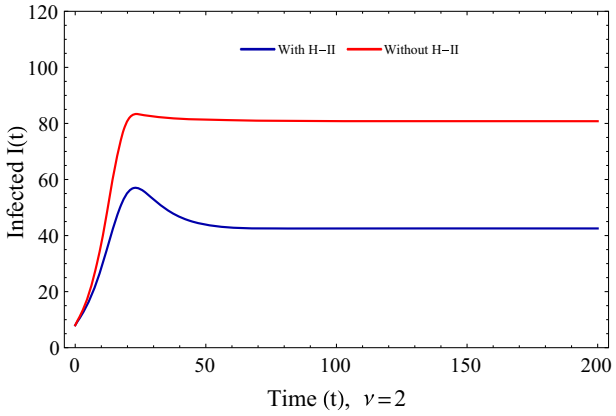
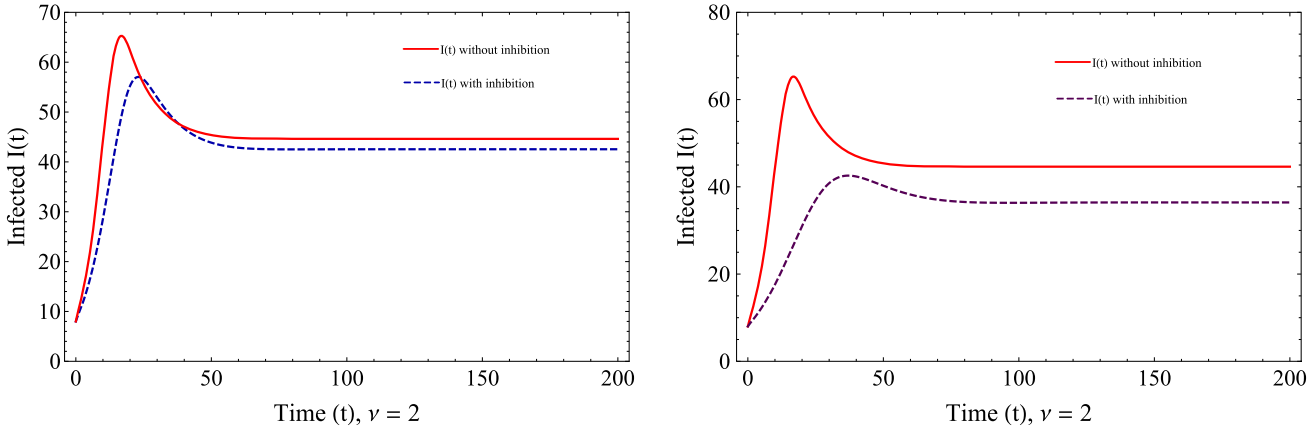


Fig. 5 Infected population with and without Holling type II treatment rate

Figure 6 shows the impact of the measure of inhibitions taken by susceptible and infected individuals. Figure 6a, b shows the difference between the infected population with and without inhibition effects. The solid red line shows the infected population without inhibition effects. When there is no inhibition taken by susceptible and infectives, then infectives are increasing at a higher rate. The dashed blue line shows the infected population when we consider $\alpha = 0.004$ and $\gamma = 0.002$, whereas dashed-purple line shows the infected population when $\alpha = 0.01$ and $\gamma = 0.01$. Thus, the higher the inhibitions rate, the higher the decline in the infected population.

Figure 7 shows the occurrence of Hopf bifurcation at different values of time delay ν in the plane (S, I). We find that $q_0^2 - q_1^2 = -0.0000587362 < 0$, $p_0^2 - 2q_0 - p_1^2 = 0.00573205 > 0$, and there is a unique pair of purely imaginary roots $\pm\iota\omega$ to characteristic equation (29) with $\omega = \pm 0.105111$.



(a) $\alpha = 0.004, \gamma = 0.002.$ **(b)** $\alpha = 0.01,\ \gamma = 0.01.$

Fig. 6 Graphs depicting the effect of inhibitions taken by both susceptible and infected individuals

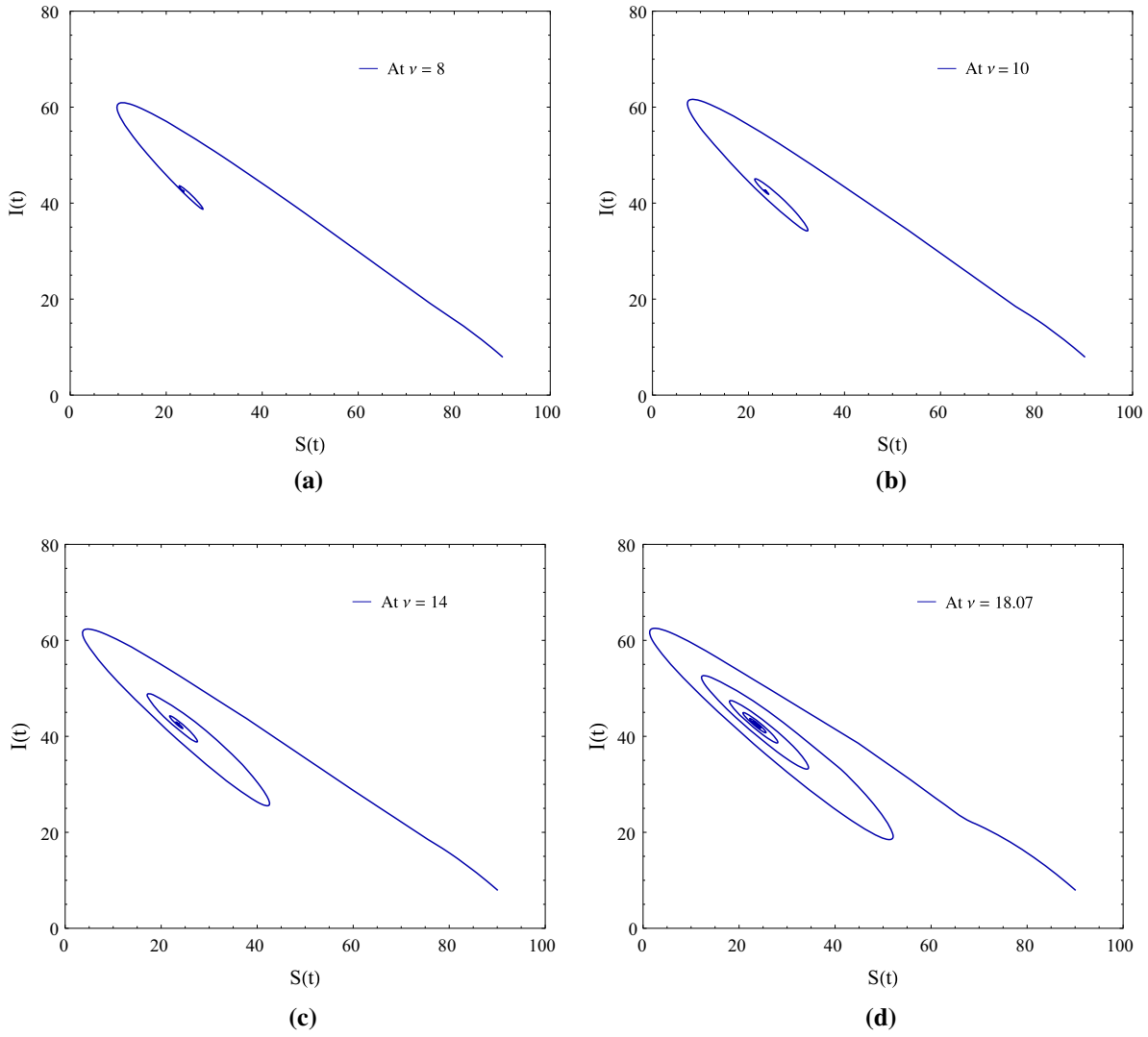


Fig. 7 Hopf bifurcation in the plane (S, I) for various values of time delay ν

Hence, the hypothesis of the occurrence of Hopf bifurcation is satisfied. According to the algorithm given in (36), we obtain the value of ν_0 which is $\nu_0 = 18.07$. Figure 7a–c shows the oscillatory solutions for the time delay $\nu = 8, 10, 14 \in [0, \nu_0]$, respectively. Theorem 7 implies that E_e is asymptotically stable for $[0, \nu_0)$ which follows from Figure 7a–c, and thus, periodic solution appears and endemic equilibrium is locally asympotitically stable which converges to $E_e = (23.5845, 42.5367)$. Also, Hopf bifurcation occurs at $\nu_0 = 18.07$ which confirms Theorem 7.

6 Discussion and conclusion

In this article, we investigate the role of time delay and effects of inhibition measures, and treatment on the SIR model with a nonlinear Beddington–DeAngelis type incidence rate and Holling type II treatment rate. We obtained that the model exhibits two equilibria: disease-free equilibrium E_0 and endemic equilibrium E_e, respectively. Stability analysis is resorted to getting a sense of the behavior of solutions. We investigated the local stability of the model near disease-free equilibrium point E_0 by deriving the basic reproduction number R_0. We observed that the disease-free equilibrium E_0 is locally asymptotically stable when the basic reproduction number R_0 is less than one, unstable when R_0 is greater than one and linearly neutrally stable when R_0 equals to one for the time delay $\nu > 0$. By applying center manifold theory, we investigate the stability behavior of the undelayed system at E_0 when R_0 is one, and showed that the model exhibits a backward (forward) bifurcation under certain conditions. Backward bifurcation plays a relevant role for disease control and eradication. We find the bifurcation range for the occurrence of backward (forward) which depends on the cure rate a. More accurately, if the cure rate a lies in a specific range, $a_1^* < a < a_2^*$, then there is an occurrence of backward bifurcation, and if a is sufficiently small or sufficiently large, i.e., if $a < a_1^*$ or $a > a_2^*$, then forward bifurcation occurs, as shown in inequalities (20) and (22). Also, forward bifurcation may occur if the limitation in the availability of resources is below a certain quantity, as shown in inequality (21). This shows that for eradication of the epidemic when $R_0 < 1$, the resource limitation should be minimized. Thus, the Holling type II treatment rate has a significant role in determining the eradication of disease transmission. Further, we found the conditions for the existence of endemic equilibrium E_e and investigated the local stability behavior of the system near endemic equilibrium E_e and obtained some interesting results stated in Theorems 3 and 4. We also discussed the Hopf bifurcation near endemic equilibrium by considering delay (ν) as a bifurcation parameter, and we obtained that the endemic equilibrium is asymptotically stable for $\nu \in [0, \nu_o)$ and the model undergoes Hopf bifurcation at the critical value ν_0. Moreover, the global stability of disease-free and endemic equilibria has been investigated and obtained that the disease-free equilibrium is globally asymptotically stable when $R_0 \leq 1$, as stated in Theorem 8, whereas, the endemic equilibrium is globally asymptotically stable when $R_0 > 1$, as stated in Theorem 9, epidemiologically this means that the disease will die out or persist in a population with respect to certain restrictions on the parameters.

In addition, we presented the numerical simulations of the proposed model to illustrate our theoretical results. We observed that when the time delay is high, then the infection occur at a higher rate in comparison to the situation when there is no delay. Further, when the Holling type II treatment rate is given to the infected individuals at the appropriate time, then there is a major difference in the occurrence of infection. Hence, to eliminate the disease, there is a need to improve medical technology and invest more medicines, beds, etc. in giving the patients timely treatment. Also, it is depicted that if treatment is not available, then infection is high even if inhibitions are taken into consideration. Therefore, the Holling type II treatment rate led to a diminished level of infection. Impact of inhibitions is also shown in the numerical simulation, which guarantees the control of infection in society if inhibitions are taken properly. From the data, it is observed that critical value ν_0 is 18.07 and for the time lag $\nu \in [0, \nu_0)$, E_e is asymptotically stable as shown in Fig. 7a–c, implies that the disease will persist in society, and at $\nu = \nu_0$, periodic solution appears, which confirms the occurrence of Hopf bifurcation at ν_0.

The epidemic model without delay shows the backward bifurcation which demonstrates that the disease-free equilibrium coexists with the endemic equilibrium when basic reproduction number is less than one. It has important qualitative implications since

reducing the basic reproduction number below one is not sufficient for the eradication of the disease from society. Thus, it is important to explain those components that can control or maintain a strategic distance from such backward situations. Our theoretical results suggest that the existence of backward bifurcation depends on the therapeutic treatment of infected individuals. Under the point of view of the disease control campaigns, public policymakers might try to prevent the dangerous backward scenario by keeping the therapeutic treatments of infected individuals suitably low. In this case, the disease control may straightly follow the classic road of reducing R_0 below 1. If the backward scenario cannot be avoided, then by decreasing contact rates with infected individuals through aimed public sensitization programs which, affecting individual behaviors, can act appropriately to reduce the disease transmission. The nonlinearity in epidemic model due to the latent period (time delay) can give rise to periodicity via a Hopf bifurcation, which shows the oscillatory behavior of the disease at the critical value of time delay.

The findings of the model consisting of Beddington–DeAngelis incidence rate with latent period as time delay and Holling type II treatment rate are capable of demonstration of substantial role of latent period, social awareness among susceptibles, magnitude of interference among infectives, and limitation in available facilities. The results are capable of understanding the transmission dynamics of the outbreak of newly emerging diseases and hence further suggesting the control strategies to prevent the spread of diseases at large scale. For example, largest ebola outbreak in 2014 in the parts of West Africa.

Like all mathematical models, our model also has limitations. For the simulation of our model, we have considered a hypothetical data to investigate the behavior of the system. The estimation of parameters from real data is an important aspect, and we leave it for future study.

In the future, for the memory effect or long-term interactions, fractional-order derivative can be applied in the present epidemic model.

Acknowledgements The authors are grateful to the Delhi Technological University, Delhi, India, for providing financial support to carry this research work. They are also indebted to the anonymous reviewers and the handling editor for their constructive comments and suggestions which led to improvements of our manuscript.

Compliance with ethical standards

Conflict of interest The authors declare that there is no conflict of interests regarding the publication of this article.

Ethical standard The authors state that this research complies with ethical standards. This research does not involve either human participants or animals.

References

1. Kermack, W.O., Mckendrick, A.G.: A contribution to the mathematical theory of epidemic. Proc. R. Soc. A **115**(772), 700–721 (1927)
2. Capasso, V., Serio, G.: A generalization of the Kermack–Mckendrick deterministic epidemic model. Math. Biosci. **42**(1–2), 43–61 (1978)
3. van den Driessche, P., Watmough, J.: A simple SIS epidemic model with a backward bifurcation. J. Math. Biol. **40**(6), 525–540 (2000)
4. Sun, K., Mou, S., Qiu, J., Wang, T., Gao, H.: Adaptive fuzzy control for nontriangular structural stochastic switched nonlinear systems with full state constraints. IEEE Trans. Fuzzy Syst. (2018). https://doi.org/10.1109/TFUZZ.2018.2883374
5. Qiu, J., Sun, K., Wang, T., Gao, H.: Observer-based fuzzy adaptive event-triggered control for pure-feedback nonlinear systems with prescribed performance. IEEE Trans. Fuzzy Syst. (2019). https://doi.org/10.1109/TFUZZ.2019.2895560
6. Kumar, A., Nilam, : Stability of a time delayed SIR epidemic model along with nonlinear incidence rate and Holling type-II treatment rate. Int. J. Comput. Methods **15**(6), 1850055 (2018)
7. Kumar, A., Nilam, : Dynamical model of epidemic along with time delay: Holling type II incidence rate and monod-haldane type treatment rate. Differ. Equ. Dyn. Syst. **27**(1–3), 299–312 (2019)
8. Kumar, A., Nilam, Kishor, R.: A short study of an SIR model with inclusion of an alert class, two explicit nonlinear incidence rates and saturated treatment rate. SeMA J. **76**(3), 505–519 (2019)
9. Dubey, B., Patara, A., Srivastava, P.K., Dubey, U.S.: Modelling and analysis of a SEIR model with different types of nonlinear treatment rates. J. Biol. Syst. **21**(3), 1350023 (2013)
10. Gumel, A.B., Connell Mccluskey, C., Watmough, J.: An SVEIR model for assessing the potential impact of an imperfect anti-SARS vaccine. Math. Biosci. Eng. **3**(3), 485–494 (2006)
11. van den Driessche, P., Watmough, J.: Reproduction numbers and sub-threshold endemic equilibria for compartment models of disease transmission. Math. Biosci. **180**, 29–48 (2002)
12. Okuonghae, D., Aihie, V.U.: Optimal control measures for tuberculosis mathematical models including immigration and isolation of infective. J. Biol. Syst. **18**(1), 17–54 (2010)

13. Gumel, A.B.: Causes of backward bifurcations in some epidemiological models. J. Math. Anal. Appl. **395**(1), 355–365 (2012)
14. Arino, J., McCluskey, C.C., van den Driessche, P.: Global results for an epidemic model with vaccination that exhibits backward bifurcation. SIAM J. Appl. Math. **64**(1), 260–276 (2003)
15. Castillo-Chavez, C., Song, B.: Dynamical models of tuberculosis and their applications. Math. Biosci. Eng. **1**(2), 361–404 (2004)
16. Wang, W.: Backward bifurcation of an epidemic model with treatment. Math. Biosci. **201**(1–2), 58–71 (2006)
17. Xu, R., Ma, Z.: Stability of a delayed SIRS epidemic model with a nonlinear incidence rate. Chaos Solut. Fractals **41**(5), 2319–2325 (2009)
18. Tipsri, S., Chinviriyasit, W.: Stability analysis of SEIR model with saturated incidence and time delay. Int. J. Appl. Phys. Math. **4**(1), 42 (2014). https://doi.org/10.7763/IJAPM.2014.V4.252
19. Goel, K., Nilam, : A mathematical and numerical study of a SIR epidemic model with time delay, nonlinear incidence and treatment rates. Theory Biosci. (2019). https://doi.org/10.1007/s12064-019-00275-5
20. Kumar, A., Nilam, : Mathematical analysis of a delayed epidemic model with nonlinear incidence and treatment rates. J. Eng. Math. **115**(1), 1–20 (2019)
21. Kumar, A., Nilam, : Stability of a delayed SIR epidemic model by introducing two explicit treatment classes along with nonlinear incidence rate and Holling type treatment. Comput. Appl. Math. **38**, 130 (2019). https://doi.org/10.1007/s40314-019-0866-9
22. Beddington, J.R.: Mutual interference between parasites or predators and its effect on searching efficiency. J. Anim. Ecol. **44**(1), 331–340 (1975)
23. DeAngelis, D.L., Goldstein, R.A., O'Neill, R.V.: A model for tropic interaction. Ecology **56**, 881–892 (1975)
24. Kaddar, A.: Stability analysis in a delayed SIR epidemic model with a saturated incidence rate. Nonlinear Anal. Model. Control **15**(3), 299–306 (2010)
25. Elaiw, A.M., Azoz, S.A.: Global properties of a class of HIV infection models with Beddington–DeAngelis functional response. Math. Methods Appl. Sci. **36**(4), 383–394 (2013)
26. Dubey, B., Dubey, P., Dubey, U.S.: Dynamics of an SIR model with nonlinear incidence and treatment rate. Appl. Appl. Math. **10**(2), 718–737 (2015)
27. Wang, W., Ruan, S.: Bifurcation in an epidemic model with constant removal rates of the infectives. J. Math. Anal. Appl. **21**, 775–793 (2004)
28. Zhang, X., Liu, X.: Backward bifurcation of an epidemic model with saturated treatment function. J. Math. Anal. Appl. **348**(1), 433–443 (2008)
29. Upadhyay, R.K., Pal, A.K., Kumari, S., Roy, P.: Dynamics of an SEIR epidemic model with nonlinear incidence and treatment rates. Nonlinear Dyn. **96**(4), 2351–2368 (2019)
30. O'Regana, S.M., Kelly, T.C., Korobeinikov, A., O'Callaghana, M.J.A., Pokrovskii, A.V.: Lyapunov functions for SIR and SIRS epidemic models. Appl. Math. Lett. **23**(4), 446–448 (2010)
31. McCluskey, C.C.: Global stability of an SIR epidemic model with delay and general nonlinear incidence. Math. Biosci. Eng. **7**(4), 837–850 (2010)
32. Sastry, S.: Analysis, Stability and Control. Springer, New York (1999)
33. Dubey, P., Dubey, B., Dubey, U.S.: An SIR model with nonlinear incidence rate and Holling type III treatment rate. In: Cushing, J., Saleem, M., Srivastava, H., Khan, M., Merajuddin, M. (eds.) Applied Analysis in Biological and Physical Sciences. Springer Proceedings in Mathematics and Statistics, vol. 186, pp. 63–81. Springer, New Delhi (2016)
34. Wang, X.: A simple proof of Descartes's rule of signs. Am. Math. Mon. **111**, 525 (2004). https://doi.org/10.2307/4145072
35. Hale, J.K., Lunel, S.M.Verduyn: Introduction to Functional Differential Equations. Springer, New York (1993)
36. Huang, G., Takeuchi, Y., Ma, W.: Lyapunov functionals for delay differential equations model of viral infections. SIAM J. Appl. Math. **70**(7), 2693–2708 (2010)
37. Li, M., Liu, X.: An SIR epidemic model with time delay and general nonlinear incidence rate. Abstr. Appl. Anal. (2014). https://doi.org/10.1155/2014/131257

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

International Journal of Communication
and Media Studies (IJCMS)
ISSN (P): 2250-0014; ISSN (E): Applied
Vol. 9, Issue 2, Apr 2019, 37-50
© TJPRC Pvt. Ltd.

TRANS STELLAR
•Journal Publications •Research Consultancy

THE NEED FOR INNOVATIVE TOOLS IN MULTIMEDIA APPROACH TO REACH THE UNREACHED IN SOCIAL AWARENESS PROGRAMS

R. PRANAVA KUMAR

Assistant Professor, Department of Visual Communication & Electronic Media, Bharathiar University, PSG College of Arts and Science, Coimbatore, Tamil Nadu, India

ABSTRACT

This is essentially a research paper as part of the doctoral thesis work, with a purpose of discerning a myriad social welfare programs and the mode of communicating to the intended target groups. Five case studies have been taken from different secondary source of information to understand the efficacy of communication tools in various social awareness program initiated by Government of India. Case Study 1 covers Water and Sanitation program of DANIDA of Cuddalore district of Tamil Nadu; Case Study 2, discusses the MGNREGA (Mahatma Gandhi National Rural Employment Guarantee Act, 2005) and its performance; Case Study 3 GPDP (Gram Panchayat Development Program) with three criteria in their proposal - hunger free, poverty free and good sanitation; Case Study 4 TSC (Total Sanitation Program) of Melapuram provides information about successful implementation of TSC in rural area and Case Study 5 WASH strategy (water, sanitation and hygiene) to control diarrhea by using potable water for drinking, washing hands and for sanitation. Upshot of the paper is that, no single tool would suffice to communicate all different programs, which in turn necessitate the identification of suitable communication tool and try and use different methods depending upon the program and target group.

KEYWORDS: Disseminate, Sanitation, Communication, Social Awareness & Hygiene

Received: Jan 22, 2019; **Accepted:** Feb 12, 2019; **Published:** Mar 06, 2019; **Paper Id.:** IJCMSAPR20194

Original Article

INTRODUCTION

Any welfare state shall endeavor to promote welfare of its citizens through various measures, and it is their prime duty to do so. Initial step is to provide basic amenities to all people of the State and proceed towards development on all sectors. Development encompasses positive changes in social, economic, cultural, environmental and physical components of humanity. Amartya Sen defines development as a tool enabling people to reach the highest level of their ability, through granting freedom of action i.e. freedom of economic, social and family actions etc. On the basis of development, the World has been divided into developed, developing and under developed nations. India has reached the slot of developing nation, and has been striving to attain developed nation status. Considering the natural resources, manpower, Government's strategies and brain power, India has a huge potential to become a developed nation on par with western countries. The moot question is why India is slow in the progress towards attaining the higher position despite having all resources. Numerous detrimental factors play a role to dampen the process of progress leading to inequitable distribution of wealth, socio cultural taboos, caste and religious intolerance, inadequate health care lack of quality education, corruption on all levels of Government, and lack of awareness among people regarding their rights and so on. The purpose of this paper is to take stock of existing programs at the global, national and state levels with a view to identify the mandates, modalities of

implementing the programs, the degree or response to mass communication modes and the need for improvising the existing methods to achieve the intended targets. It is a preliminary attempt as part of a research work leading to doctoral degree.

Communication is one of the tools to reach a large group of people, spread information, instill ideas, create awareness, change their mindset and create social amity. Mass communication has been used since the inception of newspapers, radio, televisions to reach out to every individual in the nook and corner of the country. Technology has advanced at such a rapid pace that the audio visual medium has attained enormous outreach that instant transmission of videos is possible to every part of the country. This innovate, interactive tool has been used in advertisements, education, politics, entertainment, social issues and creation of social awareness for many years. Innovative formats, new techniques, interactive communication has witnessed advancement from time to time to make communication more effective. Though communication tools spread the message, educate people, entertain and advocate; only a long lasting impression that creates a change in mindset would bring the desired result. Considering the case of creation of social awareness, the desired result would be acceptance of change and sustainable practice of the alternate method. In this context, just information and education would not suffice, though initial results may be encouraging, as people tend to stick to their old methods and customs in the long run. Thus, the search for effective communication tool arises and a strategy to customize the project to suit the stakeholders is mandatory.

Uniqueness of Social Welfare Programs in India

Social welfare, social security or social awareness, vary in their exact meaning but the underlying ideology is human welfare in pursuit of creation of informed society. The social welfare measures adopted vary among developed and developing countries in accordance to the aspirations of the ruling Government and the needs of the countrymen. Each individual is entitled to social security benefits by virtue of just being a citizen of that country, whether he contributes to economy or not. Persistent inequality in the society upholds the need for creation of social equity through advocating social awareness programs. Inequality is an inherent character of any society but the range of difference in the quality of life needs to be minimized. Consequently, the rich poor gap bridged through creation of social awareness and paving the path for all individuals in the society to achieve a better standard of living. Provision of all the basic needs may fulfill *necessary condition* like food, clothing and shelter but *sufficient condition* for development includes parameters like education, health, social status, gender equality etc. Good governance focuses on holistic development on all sectors so as to achieve well-being of its people. The social welfare measures of any Government could be classified as two types; programs which render financial assistance by direct cash or facilities and programs which educate people on various topics- the ultimate goal being improvement in the standard of living.

Comparison of HDI ranking (Human Development Index) among countries rank India at 131 among 188 countries, and it could be construed as the index for development of a country. The components of HDI are three viz. health, education and Gross Domestic Product (GDP), and a combined improvement in all three aspects would only improve the ranking of a country. Central and State Governments of India has been planning and implementing various social welfare schemes since Independence, but the desired effect is lagging and not reaching the whole population. Along with Government, many NGOs work hand in hand to improve the life style of the oppressed groups and create equity, with only partial success. This paper aims to discuss the factors that deter equitable distribution of benefits in India and analyze few factors that could be modified to create a better result.

All nations, be it developed, developing or under developed have charted out a list of welfare programs in accordance to the need of its citizens. Analysis of the various programs of United States of America (Developed) and India (developing) countries (Annexure 1) gave a picture of similarities in the main sectors but marked dissimilarities in many sub sectors. India being a developing country with diversity in socio-cultural parameters in each State, the social welfare measures included schemes for the marginalized through education, health and employment in a special basis so as to create equity. Careful examination of the aforesaid, fairly exhaustive though not complete list, the extra effort and schemes that were incorporated by the Indian Government showed the robust quality of planning and inclusiveness. Women and child health care, old age pension, education for all, housing, disabled care and so on are common to both USA and India. But promotion of sex ratio, assistance to widows, marriage assistance, supplementary nutrition to adolescent girls, scholarship to backward and minority students with supply of free text books and uniforms, free technical education to minorities, SC& STs., free coaching and tuition for weaker economic status students, free hostel for male and female students, vocational training, night shelters, promotion of handloom and handicrafts, public distribution system, Annapurna scheme, mobile dispensary vans, agricultural loans and implements supply, co-operative societies formation, insurance scheme for laborers are few schemes unique to a developing nation like India. Departments concerned with each sector directly appoint staff to carry out various activities or entrust the role to select NGOs by providing them financial assistance. Since the emoluments need to reach every nook and corner of the state at correct time interval, a concerted effort is needed to manage these programs by the Government and NGOs.

Other than Central Government, each State government supplements many programs for its citizens and such programs are specific to the socioeconomic, cultural, and educational and health needs of the people. Each State in addition to Centre's program creates new ones to suit the needs and aspirations of its citizens. Social welfare board of Tamil Nadu has its schemes planned and executed by their ministry Social welfare and the schemes of Tamil Nadu are broadly classified under the following heads, as recorded by the Tamil Nadu Social Welfare Board. Other than the stated programs, every time the government changes, once in five years, new schemes are announced and adequate funds allotted to implement such projects. Success of any scheme depends on the stakeholders, they must be aware of the scheme, its intent, the change it will bring to their lives, acceptance, and sustainability. Thus introduction of the schemes alone would not bring development, only a change in the mindset of the people and their propensity to carry forward the project on a continuous basis would create the real development aspired. Creation of social awareness about the projects, inculcating the idea into the mindset, creating a change in the traditional habits, following the new idea on a sustainable basis would be only possible through communication. Mass communication assumes importance here by serving as a tool to create social awareness and thereby development

Communication Tools

Social awareness programs and its modalities of propagation vary according to the requirements of the particular society. Choice of the communication tool depends on the type of project and the target stakeholders. Consequently, successful implementation of any social welfare or security program depends on creation of awareness through various modes of communication. The tools of communication ranges from lectures, personal talks, posters, banners, pamphlets, newspaper, street songs, puppets, street plays, radio, cinemas, TV, advertisements and so on. Each tool has its own target groups and reach; the selection of tool to be used varies according to the stakeholders. No single medium could be construed as the most effective in all situations as the target population varies in their socioeconomic status, education,

access to the medium, their understanding and capabilities to accept change.

India in general and Tamil Nadu in particular, has a host of good values to be carried on as well as bad outdated customs that need to be changed. 'The Vision of Tamil Nadu 2023' envisages provision of world class facilities to all residents irrespective of caste or class in the form of housing, water, sanitation, energy transport, irrigation, connectivity, transport, health care and education. Open air defecation is one such outdated unhealthy practice that needs to be completely removed from the society. Total Sanitation Program, one of the major components of Vision 2023, the state of Tamil Nadu is far away from achieving it. It is heartening to note that access to connectivity is progressing at a faster pace than access to toilets. *"There are 45 countries including India, where less than half the population has access to improved sanitation facility. Despite increase in sanitation coverage, a vast section of people of India do not have access to improved sanitation facilities"* 2014 update, WHO & UNICEF. "As per 2011 Census, in Tamil Nadu 48 percent of households have access to latrine within their premises and 5 percent use public toilets, remaining 47 percent resort to open air defecation. In rural areas, the proportion is 73.27 percent using open air defecation, in urban areas it is 16.21 percent practicing open air defecation in Tamil Nadu", as per the information provided at the Workshop on Converging Strategies 2015.

"With the declaration of 100 panchayats as ODF (open air defecation Free), nearly one-fourth of the 420 panchayats would become ODF in Madurai district" 2016, B. Rohini Ramdas, continued " the bigger challenge though is to ensure people are consistently using toilets, we had difficulties in convincing people on the need for toilets, but now, that has changed a lot" The strategy would be to motivate people to accept toilets as their fundamental right and practice in door defecation in order to maintain ones dignity and health. In urban areas, the lack of space and absence of toilets in rented buildings lead to open air toilets in crowded slums, whereas in rural areas, irrespective of socio economic status people follow open defecation as this tradition has been inculcated into their minds. Change in the mindset, understanding the health hazards and implementation of indoor defecation among the rural masses need social and behavioral change communication, which is development communication. Thus, communication for development forms the major tool to inculcate changes in the habits of people on a sustainable basis.

Human nature varies from individual to individual that the perception, interpretation and reaction to mass media varies widely. " *It was discovered that audience of mass media were not isolated individuals but remained tied to friends, families, work associates community etc. called the theory of selectivity based on social relationship",* 2011, Rajesh Kumar. *"Reviews of successful BCC- Behaviour Change Communication, campaigns in India indicate that evidence based long term campaigns spanning 4-5 years, an integrated multi-media approach, ongoing monitoring and program modification, if necessary are critical for ensuring sustained behaviour change."* Shalini Prasad and Manisha Singh,2010.

Taking the example of Open air defecation and Swacch Bharath program, it could be understood that communication was not effective to change the habits of the people. Both Central and State Governments have been consistent in building public toilets, giving monetary assistance for construction of household toilets, creation of sanitary marts where you can buy the implements at a nominal rate, educating people about the health hazard of open air defecation, encouraging ODF status by giving awards to Panchayats, door to door talks, posters, advertisements, educational films and creation environmental awareness. These efforts have produced some changes but the desired effect has not been attained. The reason behind such status is lack of mind set change, belief in old customs and traditions, religious causes, lack of maintenance, failure in sustainable use of toilets even when it is available.

Discussion in the preceding pages brings out the fact that there exists a gap between the intended outcome and the actual in all cases of social welfare programs; giving raise to question about the efficacy of the methods of propagation. The solution to one problem may seldom be suitable for another and any methodology could hardly be claimed as the best. How to find which is the best method for spreading the awareness among the people in raising their level of awareness? In this milieu, a study of few success stories published, could throw some insights as to which is best suited for a particular project. Five case studies published in various journals were taken and the methodologies were studied. Each case spelt out a different method depending on their target, budget, stakeholders and time. Some showed success by including women, some using IEC tools, stakeholders' involvement, gram saba meetings and direct door to door canvassing. The main aim, the method and results of the five projects, as culled out from the available literature, are given as case studies below

Case Study 1

Water and Sanitation program of DANIDA carried out in more than hundred villages of Cuddalore district of Tamil Nadu, reiterated people's participation and women empowerment as major focus to achieve the desired result in their projects. Their project motto was *demand driven approach*, wherein they send messages to all villages about their project plan and villagers need to involve voluntarily in the project according to their demand. One such village was Vella Kollai village in Cuddalore district with defunct and poor water supply. Vella Kollai gram panchayat president S. Sundaram, attended the block level meeting and he initiated the project of water supply to his village with the funds from DANIDA. Total project cost of Rs. 3, 09, 485 (year 2000), villagers contributed one –seventh and the rest by DANIDA. Planning, execution and maintenance were done by village groups particularly women. Bore wells, water head tanks, pipes were laid and supply for 4 hours in a day from 20 taps constructed at appropriate sites solved the problem of water supply in that village totally. Similarly sanitation project was carried out in a village called Keerapalayam, Cuddalore district. This was totally a women initiated project, though the panchyat president was a male, K. Paneer Selvam, all households were provided with toilets with the help of DANIDA funding and one – fourth contribution by villagers. The strategy for this project were-strengthening local bodies, peoples' participation, gender equality, women empowerment, maintenance by women user groups. As an example to female involvement, Ms. Suda of Vella Kollai village got trained in pump mechanism and was always ready to rectify any repairs that would arise in the water supply program. Another Ms. Kamalamma trained as mason built most of the toilets in Keerapalayam village.

Case Study 2

MGNREGA (Mahatma Gandhi National Rural Employment Guarantee Act) offers 100 days of work per year to all eligible rural poor in India, and is one of the largest social protection schemes in the world. Since implementation in 2006, it has spreaded to all villages in India covering almost 50 million rural households in 2013-14. Analysis of the project showed marked disparities among and within States of India. ESID, 2012, reports the following, *"Employment outcomes are lower in states with higher poverty levels • MGNREGA implementation depends on the supply of work rather than demand for it • Supply is determined by differences in state capacity and commitment, local power relations, the role of lower-level functionaries, and political competition • Improving employment outcomes requires strengthening the demand-driven aspects of MGNREGA through a focus on local level social audits, funding and tracking of outcomes • States should be provided with support to improve their capacity and allowed a degree of flexibility in MGNREGA implementation "*

Such large projects could not be designed with a common methodology as the needs and aspirations of people vary among states and within it. From the report we incur that, the projects should be demand driven and the funds should be channeled through gram panchayat to villagers rather than district level.

Case Study 3

Gram Panchayat Development Program (GPDP)

The gram panchayat development program plans to formulate their proposals under three major criteria, hunger free, poverty free and good sanitation. On the Tamil Nadu front, there are 12,524 gram panchayats in the 31 districts wherein the GPDP programs has been in vogue. A recent report, puts forth the failures and success of such initiatives in perspective, *There has been a partial success on the implementation of this program and the concluding remarks was, "Only awareness generated on the scope and need of SDGs (Sustainable Development Goals)," would lead to better results in future (2017, Workshop)*

The IEC (Information, Education and Communication) tools play a major role in the successful outcome of these projects. The IEC tools may be in the form of charts, pamphlets, hand books, posters, books, audio, videos and short films. In spite of using various modes of communication, the desired effect was not achieved and the workshop suggested more effective awareness or sensitization of the stakeholders to improve the success rate. The choice of IEC tools must be appropriate to the socio-economic and cultural status of the target leading to customized tools and newer tools of communication to reach the audience more effectively.

Case Study 4

Total Sanitation Program of Melapuram, a success story, was analysed and projected as a model for replication in other parts of the state. Block coordinators and field staff were given adequate inputs and training regarding TSC and entrusted with multi-media tools to sensitize people and help them solve issues that might arise in implementation of the program. Monitoring was carried out through coordinated mechanism from grassroot level through computers by creating a network among all 10 computers of the village panchayat. Adequate training in Computer operation was given to field staff (local residents), which helped to maintain accounts and data. Local NGOs were also involved to make the program robust and sustainable. This model disseminated by creating a web page for this block and spread to all other blocks for replication. In this case, effective use of stakeholders' participation, modern technology and multimedia helped in achieving the desired result.

Case Study 5

The impact of WASH (water, sanitation and hygiene) in controlling diarrhea was studied and found the desired effect lagging far behind the expected outcome. A review of 18 studies revealed that provision of adequate water and sanitation facilities did not effectively bring down the out breaks in cholera. Lack of knowledge that unclean water is responsible for the disease, clean hands form the basis of prevention, proper sanitation can only reduce the incidence were the prime lacunae discovered to impede the success of WASH strategy in prevention of cholera out breaks. Certain cases of improved knowledge had only limited response that those even with awareness did not adhere to clean practices in the long run. Most case studies exhibited success in the first few months followed by back to old practice. As conclusion, it was reiterated that there was a gap in assessing the correct strategy to a particular section of people. More innovative strategies using the limited resources to improve the WASH intervention program, tailored to the target, should be effective in

control of cholera.

DISCUSSIONS

Social welfare programs are stakeholder oriented unlike others, where the project proposer also stands to gain. For example, if a company producing sanitation products like toothpaste or soap advertise their products reiterating the importance of cleanliness and health, through newspapers, TV or posters, their aim is only business. Their target audience is consumers with the capacity to buy their products, others who are in low socio economic strata who are not their customers are out of their concern. They aim at users wooing and maintaining their brand loyalty. Conversely, social welfare program of the Government aims at equity, the program envisages inculcating sanitary practices by changing the mindset of the people. The program aims to cover the whole population. Bringing about a change in their lives and increasing their standard of living. The above case studies are projects implemented by Government or NGOs and their success could not be measured by monetary scale but by the sustainable achievement of the desired result. Thus, the search for different methodology, innovative tools, inclusive participatory modes and new strategies must be delved to fit a project. In this context, *no one size fits all* becomes true as each example project has been designed differently and carried out through various modes and made successful. The common thread is stakeholders' participation, demand driven approach and monitoring by locals. The tools of communication vary with each project, direct talks, *gram sabha* announcements, posters, IEC tools, digital platforms, etc.

The social welfare program chosen for the thesis work is sanitation encompassing construction of toilets, hand washing and safe disposal of waste. Analyzing a few success stories as mentioned above, in search for a suitable tool for creation of awareness towards sanitation, school children was selected as target population, as it would be easier to inculcate sanitary practices and its benefits in young minds rather than old ones who are bound by their traditions and customs. In such a scenario, a clearer understanding of case studies presented in the preceding pages helped and directed the research, as the authors become inquisitive to probe whether animated audio visuals would lead to a better understanding and long standing effect on young minds or not? The animated video is being designed by the authors in line with the principles of development communication to make it more interesting, target oriented, grasp the imagination of the youngsters and create a lasting impression in their minds. The story line would be proactive rather than advisory so that children tend to implement the project at each level with a strong note of do it yourself.

Social welfare forms the major duty of Government that most of its resources are directed in creating an equitable social welfare of its citizens. Comparing the list of social welfare programs of a developed country like USA and a developing country like India, the broad categories like health, education, employment etc. remain the same. But, the exhaustive list documented brings out the variations in the programs between these two countries. More emphasis is given to child health, old age security and rehabilitation of destitute, addicts, juvenile delinquents, immigrants and poor. In India, there is a need to provide special relief to marginalize like minorities, SC, ST and other backward classes. Gender issues need to be taken more effectively to create gender discrimination less society wherein all citizens could enjoy the rights equally irrespective of their gender. Health and sanitation programs require innovative approaches to change the mindset of the people in order to realize MDG (Millennium Development Goals). Other variations are farmer's subsidy, free rations to BPL families, disability pension, old age pension, widow pension and so on. Notwithstanding the disparities among USA and India social welfare programs, the strategy to implement is the hallmark of the project success and communication tools play a major role in translating any social welfare program to reach its target population. Here again "no one size fits

all", each project is unique in its nature, target, resources, and goal. Analysing five of the success stories of Indian social welfare programs, the strategy and methodology differences employed in each program vary reiterating the fact that each program has to be treated unique and tools used customized to the target population and their requirements. The case studies also bring out the difficulties faced and the corrections carried out to make the project more amenable to implement.

The case study of the five projects with all its methodology for successful implementation puts forth the necessity to design the project from conception level. Involvement of stakeholders at the planning level, understanding their views before formulation of a strategy, using local people and language to propagate the idea, identify the target people whose mindset could be influenced, use locals as leaders in bringing a change, use multiple combination of communication tools, upgrade the communication tool to be more interactive and innovative, plans must be personalized for each community or changed in accordance with the necessity of that group.

CONCLUSIONS

Research in media and mass communication is fairly new, when compared to natural, physical and social sciences. Mass communications, although aims to reach the target group as a whole, the success rate has proved to be partial so far. The reasons are manifold; one of the reasons is sticking to anyone particular mode to communicate to the whole lot. This paper infers that new modes need to be tried by taking into account the age, gender, socioeconomic status and cultural background of the target group. Animation as a tool has been tried sporadically with varying degree of success and failure. The authors of this paper, in the backdrop of knowledge gained through the above review of case studies, have been attempting to test the efficacy of animated models to communicate to the younger target group on sanitation practices. In sum, this may be a pointer for researchers to try innovative tools for social awareness communication.

REFERENCES

1. *Alon Meidan, 'What is Development?', Society for International Development, Israel Branch,March 2018*
2. *World development indicators Data@worldbank.org, updated May 3, 2018.*
3. *The millennium development goals report 2015, United Nations New York,2015*
4. *Millennium development goals -India Country report 2015, UNDP in India.*
5. *Social welfare schemes of USA- Wikipedia 2011.*
6. *www.Tamil Nadustat.com*
7. *DNA website, 30 September 2015.*
8. *Seth, P., & Bhatt, V. (2015). Higher Education: Innovative Entrepreneurship Education and Its Promotion.*
9. *Evaluation study on Total sanitation Campaign (TSC), 2013, Planning Commission Government of India*
10. *Dawn L Taylor, Tanya M Kahawita, Sandy Craincross, Jeroen H J Ensink, 2015, 'The Impact of Water, sanitation and Hygiene interventions to control cholera: A systematic review, http://dx.doi.org*
11. *MPLADS (Member of Parliament Local Area Development Scheme),Rural development and Panchayat Raj Department, Government of Tamil Nadu, 2011. www.tnrd.gov.in*

12. *Prudhvi Pavuluri, May 2014, Analysis of Total Sanitation Campaign, Social Sciences. in Case studies of success stories from states of India regarding TSC program, www.mdws.gov.in/sites*
13. *Virginia paul, Priyanka Singh, Sunita B John, August 2013, Role of mass media in social awareness, International Journal of Humanities and social sciences, Voi. 1, August 2013, page 34-38*
14. *B. Rohini Ramdas, Additional Collector, Madurai, in The Hindu dated January 1, 2016,*
15. *The Workshop on Converging Strategies towards open defecation free India, September 15, 2015, Madurai Symposium.*
16. *Gangawane, P., & Sayed, U. (2014). Innovative Use of outdated Doxycycline Hydrochloride to Dye wool. International Journal of Medicine and Pharmaceutical Sciences 2014; 4 (2): 43, 46.*
17. *Progress on drinking water and sanitation- 2014 update- WHO & UNICEF*
18. *Rajesh Kumar, 2011, Mass media can effect socio-political changes: An analysis, Global Media Journal, Indian edition, June 2011.*
19. *Shalini Prasad and Manisha Singh, 2010, Role of advertising agencies in Behaviour Change Communication, The Journal of Family welfare, Vol 56- 2010.*
20. *Nagabhushana Rao. S., 'Water and Sanitation-The DANIDA experiment', Social Welfare, June 2001.*
21. *Rabi Thapa et al 2012, 'Success and failure of MGNREGA implementation in India"2012, Effective States and Inclusive Development Research Centre, School of Environment and Development The University of Manchester*
22. *Ghosh, R., & Goswami, P. (2014). Implementation on adaptive transmission of real time multimedia data. International Journal of Re-search in Engineering & Technology, 2(4), 175-184.*
23. *Workshop on Experience sharing on GPDP By MoPR at NIRD&PR during 1-3, Dec. 2017, RURAL DEVELOPMENT AND PANCHAYAT RAJ DEPARTMENT,Government of Tamil Nadu*
24. *Dr. P.SivaRam*, Dr.P.Durgaprasad "Impact Assessment of Total Sanitation Campaign (TSC) in Nirmal Gram Puraskar (NGP) Awarded Melupuram Block in Kanyakumari District, Tamil Nadu, India*
25. *Dawn L Taylor, Tanya M. Kahawita, Sandy Cairncross, Jeroen H J Ensink, "The impact of Water, Sanitation and Hygiene interventions to control Cholera : A systematic review", PLOS, August 18, 2015, http://doi.org/10.137/journal.pone.0135676*

ANNEXURE

Social Welfare Programs of USA

Social- old age assistance, Temporary assistance to needy, Foster care, Adoption assistance, General assistance in cash, Refugee assistance, and General assistance to Indians, however the list is not exhaustive.

Health- Medicaid, Health insurance, Medical general assistance, Health centers, Maternal & Child health, Medical assistance to refugees, Healthy start

Food- Food stamps, School lunch, Women, Infant & Children Food, School breakfast, Nutrition program for elderly, Commodity supplemental food, Temporary emergency food, Needy families, Farmers market Nutrition program, Special milk program, Summer program

Housing- Section 8 housing, Public housing, Low income housing, Home investment partnership, Assistance to homeless, Rural housing insurance, Housing for elderly, Assisted housing program, Housing for disabled persons

Energy- Low income home energy assistance, Subsidized low income phone service, Weatherization

Education- Pell grants, Grant for local education authorities, 21st century learning centres, Special program for disadvantaged, Supplement educational opportunity, Adult basic education grant, Migrant education, State student incentive grant, Education for homeless children, Aid for graduate & professional study for minorities

Training- Work activities and training, Job corps, Summer youth development, Adult employment & training, Foster grandparents, Youth build, Migrant training, Native American training

Service- block grant services, Social service block grant, Community services grant, Safe & stable families, Aging Americans act, Family planning, Emergency food and shelter, Responsible fatherhood grant, Foster care independence grant, Independent living & training, Maternal, infants and child home visits

Child- Child care & development, Child care entitlement to state, Block grant child care

Community- Block grant, Economic development grant, Appallachian regional development, Empowerment zone, Enterprise communities renewal.

Social Welfare Programs of Central Government of India

- **Social Welfare Schemes** - old age pension, National family benefit, Jan shree bhima yojana- insurance cover to poor, Scholarship to disabled person, Unemployment allowance to disabled person
- **Women and Child Development** - Pension to widows, LADLI - promotion of sex ratio, Financial assistance to lactating and nursing mothers of weaker section, Financial assistance to poor widows for their daughters marriage and marriage assistance to orphan girls, Scheme of Bhagidari- care of Women senior citizens
- **Nutrition Programs** - Supplementary nutritional program, Kishori Sakthi Yojana- improve nutritional status of adolescent girls, National program for adolescent girls.
- **General Education** - Scholarship to students of economically weaker section, Welfare of educationally backward minority students, Lal Bhahadur Sastri Scholarship to meritorious students of economic weaker section, Subsidy for school uniform to students, Free supply of text books, Welfare program for children of MCD school, Welfare scheme for children of NDMC schools, Incentive as scholarship for promoting education among girls/SC/ST/OBC in class I to V
- **Technical Education** - Strengthening facilities to students of SC/ST/OBC/Minorities in polytechnics, Coaching class for SC/ ST/ Weaker sections in Delhi college of engineering, Coaching class for SC/ ST/ Weaker sections in Netaji Subash Institute of technology, Academic development of SC/ST students in College of Art
- **Welfare of SC/ST/OBC/Minorities** - Financial assistance for purchase of stationery, Scholarship, Merit scholarship from class Vi to XII, Merit scholarship foe college and professional courses, Vocational and technical scholarship, Hostel for student boys, Hostel for student girls, Pre- exam coaching, Reimbursement of tuition fees in public schools, Financial assistance/ loan for self-employment, Vocational training, Improvement of living condition
- **Urban Development** - Swaran Jayanthi Sahari Rozqar Yojana- gainful employment to urban unemployed poor, Urban basic services program, Construction of community halls, Environment improvement in urban slum,

Construction of pay and use Jansuvidha complexes

- **Housing** - Houses for weaker sections, Night shelters
- **Industry** - Society for self-employment, Promotion of handlooms, Handloom weavers welfare- insurance, thrift cum security, weavers colony, Promotion of Handicrafts
- **Civil Supplies** - Streamlining of PDS with focus on BPL families- targeted PDS, Antyodaya Anna Yojana- poorest given subsidized wheat and rice, Annapurana Scheme- free grain of 10 Kg per head per month above 65 years.
- **Medical** - Mobile van dispensaries for JJ clusters, Delhi state health mission
- **Science and Technology and Environment** - Involve weaker sections in improvement and upgradation of environment, Labour and Labour welfare- Rashtriya Swasthiya Bhima Yojana- Health cover to BPL families, Technical training for self-employment, Coaching cum guidance facility for ITI students
- **Agriculture and Allied services** - Grant for land improvement and agricultural implements, Training for SC/ST in modern poultry farming, Co-Operation- Credit for SC/ST in cooperative societies
- **Aam Aadmi Bheema Yojana** - Life insurance cover for landless labourers.

Social Welfare Programs of Tamil Nadu

- **20 Point Program**

Agriculture, Children, crime, labour, rural, demography, SC ST, tourism, economy, education, urban, welfare, women, energy, environment, health, housing, industry, infrastructure

- **Backward Class Schemes**

National Backward classes Finance and Development Corporation (NBCFDC), Backward regions grant fund, National Scheduled caste Finance and Development Corporation (NSFDC), National Scheduled Tribes finance and development Corporation (NSTFDC), Aids to NGOs, Financial assistance to upgrade technical and Entrepreneurial skills, Term loan scheme,

- **Child Schemes**

Anganwadi, Balika Samriddh Yojana, ICDS(Integrated child development scheme), ICPS (Integrated child protection scheme), Juvenile justice, Mid day meals, National creche fund, RCH (reproductive and child health, Shishu Greh scheme, others

- **Disabled Schemes**

Educational schemes, vocational training, rehabilitation, artificial limbs, scheme for mental retardation and rehabilitation, special schools for mentally challenged, purchase of aids and appliances, concession certificates.

- **Foreign Contribution**

Regulation of foreign contribution through NGOs., Foreign contribution regulation act, Funds utilized by NGOs.

- **Labour and Employment Schemes**

MGNREGS Mahatma Gandhi National Rural Employment Guarantee Scheme, EAS (Employment Assurance Scheme), Food for work Program, Jawahar Rozgar Yojana (JRY), Million Wells Scheme, National Maternity Benefits Scheme (NMBS), Prime Minister's Rozgar Yojana (PMRY), Rural Employment Generation Program (REGP), Sampoorna Grameen Rozgar Yojana (SGRY), Scheme for working women's hostel, Rehabilitation and bonded labourers, (STEP) Support to Training and Employment Program for Women, Swarnjayanti Gram Swarozgar (SGSY), Training of rural Youth for Self- Employmant (TRYSEM)

- **Minorities Schemes**

Fellowships, National Minorities development and Finance corporation NMDFC, multi sectoral development, Scholarships, Educational loans, micro finance, vocational training, coaching scheme, welfare and development of SC and STs, disbursement of loans, micro credit

- **Rural Schemes**

Rural water supply, Accelerated Rural water supply program(ARWSP), agriculture and rural debt relief, Anti-poverty, Bharath Nirman, Bhoodan Movement, Central schemes, credit cum subsidy, Desert development, District rural development agencies, Drought prone area programs, Employment assurance scheme, Food for work, Indira Awaas Yojana, Integrated rural development program, Integrated wasteland development, Jawahar Geam samridhi, yojana, Jawahar Rozgar Yojana, Million wells, Watershed development, Rural infrastructure development, rural sanitation, land development, self-help groups, updating land records, Toolkits to rural artisans, Rural drinking water, Training for rural youths

- **Social Schemes**

Members of Parliament Local area Development scheme (MPLADS), Prohibition, annapurna, NGO assistance, CAPART sponsor, Freedom fighter's pension, Growth centres, Liberation and rehabilitation, Family benefit, Old age pension, National Social Assistance Program (NSAP), Short stay homes, Social defence schemes

- **Tribal Schemes**

Vocational training, Central sector, centrally sponsored schemes, educational schemes, special component plan, Tribal Co-operative Marketing Federation of India (TRIFED), tribal sub-plan.

- **Urban Schemes**

Accelerated Urban Water Supply Program AUWSP), Jawaharlal Nehru National Urban Renewal Mission (JNNURM), urban infrastructure development for small and medium towns (UIDSSMT), Valmiki Ambedkar Awas Yojana (VAMBAY), Swarna Jayanthi Shahari Rozgar Yojana (SJSRY)

- **Voluntary Schemes**

Schemes and measures taken by Voluntary organisations and Non government organisations NGOs.

- **Women Schemes**

Awareness generation projects for rural and poor women, education for adult women, CSWB scheme, Distance education family counselling, Janani Surakshe Yojana, Kishori Shaki Yojana National Maternity Benefits Scheme (NMBS), NORAD scheme, Rashtriya Mahila KOSH, Working Women's hostel, Science and technology for women, Training and employment program for women, Swa Shakti Project, Swayamsidha Scheme, Vocational rehabilitation for disable women, Vocational training.

Renewable Energy 147 (2020) 2118–2127

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Thermodynamic analysis of Organic Rankine cycle driven by reversed absorber hybrid photovoltaic thermal compound parabolic concentrator system

Deepak Tiwari [a, *], Ahmad Faizan Sherwani [b], Deepali Atheaya [c], Anil Kumar [d, **], Nishant Kumar [e]

[a] Department of Mechanical Engineering, Bhagwan Parshuram Institute of Technology, New Delhi, 110089, India
[b] Department of Mechanical Engineering, Faculty of Engineering and Technology, Jamia Millia Islamia, New Delhi, 110025, India
[c] Mechanical and Aerospace Engineering Department, School of Engineering and Applied Sciences, Bennett University, Tech Zone – II, Greater Noida 201310, UP, India
[d] Department of Mechanical Engineering, Delhi Technological University, Delhi, 110042, India
[e] Department of Embedded System, National Institute of Electronics & Information Technology, Calicut, 673601, India

ARTICLE INFO

Article history:
Received 1 May 2019
Received in revised form
29 September 2019
Accepted 5 October 2019
Available online 7 October 2019

Keywords:
Organic Rankine cycle
Reversed absorber
Concentration ratio
Heptane /R245fa
Zeotropic mixture

ABSTRACT

In current study, an effort is made to improve the overall performance of the solar powered organic Rankine system to meet increasing energy demand and mitigate the resulting menace caused by environmental imbalance. Hourly performance evaluation of Solar powered Organic Rankine Cycle (ORC) system is done for varying concentration ratio in terms of collector output temperature, heat gain in evaporator, combined thermal and exergetic efficiency. The Heptane/R245fa is used as working fluid at fixed (0.1/0.9) fractional mass component. Organic Rankine system is coupled with reversed absorber hybrid photovoltaic thermal compound parabolic concentrator (PVTCPC). The highest collector outlet temperature 125.45 °C, heat gain 3.21×10^5 W, combined thermal efficiency 7.8%, and exergetic efficiency 14.38% is observed at 13:00 h for 6 concentration ratio of the concentrator. This study establishes the utility of low-grade renewable heat source conversion into solar electrification, solar irrigation, solar cooling and solar water purification.

© 2019 Elsevier Ltd. All rights reserved.

1. Introduction

Over-utilization of fossil fuels to meet the world's growing energy demand causes several problems, such as shortage in fossil fuel reserves, global warming, air pollution and depletion in ozone layer. Alternatively, various renewable energy sources such as solar [1], geothermal [2], biomass [3] and ocean-thermal [5,6] have been used in order to curb these concerns. Solar thermal applications leads among these resources due to freely available in nature, easily exploitable, cost-effective and no adverse effect to the atmosphere [7,8]. Therefore, an attempt has been made to utilize solar thermal energy for power generation through organic Rankine cycle (ORC).

Power generation technologies include; organic Rankine cycle [1], Kalina cycle [9], Goswami cycle [10] and Trilateral cycle [11]. Organic Rankine cycle is best among these as simpler structure, flexible, reliable, lower operational cost and improved power generation [12]. The selection of the working fluid is important facet for ORC system [13,14]. The performance of ORC system using the pure fluid as working media is not desirable due to substantial exergy loss in the two-phase region. This is because of thermal mismatch in the evaporator (collector) [15]. Pure working fluids exhibits constant pressure and temperature on phase change process which leads to mismatch of temperature in heat exchangers. The pure working fluids; R245fa, R12, R123, R134a, R1234yf, R290, Ethanol, R141b, R718, water, ammonia, CO_2 were employed in ORC system ([16–19]. Further, flammable pure working fluids such as: R600 (Butane), R600a (Isobutane), R6001 (Benzaldehyde), benzene, isobutene, isopentane, cyclohexane are also used in ORC system [13,17,18,21,22].

The ORC system performance is improved by using zeotropic fluid owing to reduction of exergy destruction in a two-phase

* Corresponding author.
** Corresponding author.
E-mail addresses: deepakmathiya@gmail.com (D. Tiwari), anilkumar76@dtu.ac.in (A. Kumar).

https://doi.org/10.1016/j.renene.2019.10.018
0960-1481/© 2019 Elsevier Ltd. All rights reserved.

region (Xi et al., 2017). Therefore, zeotropic mixture is recommended as working fluid in ORC systems. The zeotropic mixtures exhibits isobaric and non-isothermal phase change processes and lead to improved temperature match in heat exchanger [23,24].

Comparisons of thermo-physical properties for some working fluids are summarized in Table 1. The considered working fluid heptane belongs to a hydrocarbon group. The ozone depletion potential (ODP) of heptane is zero and exhibit negligible global warming potential (GWP) however, it is flammable. The working fluid R245fa is non-inflammable with zero ozone layer depletion potential but exhibits some GWP. Moreover, the GWP of the working fluid heptane/R245fa will be lower than the GWP of pure R245fa. Flammability of heptane is suppressed by blending with more fraction (0.9) of incombustible working fluid mixture component R245fa in whole mixture of (heptanes/R245fa) in comparison to flammable mixture component heptane (0.1) (Xi et al., 2017; [23]. Moreover, heptane and R245fa are dry Working fluids therefore possibility to erode the turbine blade gets reduced [25,26].

Solar collectors are responsible for converting solar radiation into useful heat energy. The selection of solar collector significantly affects the overall Solar powered ORC system performance [13,16]. The collector outlet temperature for flat plate, evacuated and for parabolic collectors are found in the range 80–90 °C, 130–150 °C, 150–165 °C, respectively [22]. Flat plate collector has lowest outlet temperature among reverse two absorber flat plate collector, parabolic collector, evacuated collector owing to substantial radiation and convective heat losses [27]. The flat plate collector is costlier than conventional compound parabolic concentrator because of its higher collector area [28]. Radiation and convectional heat losses of conventional compound parabolic concentrator are substantially reduced which increases collector outlet temperature.

The exit fluid temperature of the concentrator is in the range of 300–550 °C which is higher than the flat plate collector with approximate thermal efficiency of 25–30% [29,30]. Performance of reversed absorber PVT/CPC collector covered with glass is best among unglazed reversed PVT/CPC, inclined PVT/CPC and horizontal PVT/CPC systems in terms of collector outlet water temperature and overall exergetic efficiency [28].

The traditional solar thermal collector converts solar radiation into heat, whereas PV system directly into electricity. The important limitation of PV system is reduction of solar cell efficiency with increases in PV cell surface temperature. The absorbed heat in the PV system is being wasted. Therefore, in order to enhance thermal heat gain and reduction of surface temperature of the PV cell, a traditional solar thermal collector can be interfaced with PV system [31]. Moreover, PVT/CPC exhibits better electrical efficiency than single PV system [32]. Parabolic trough collector exhibits 1500 °C at the concentration ratio 600 to 2000 whereas linear Fresnel reflector gives 500 °C for the concentration ratio 30 to 80 [29,30]. [33] performed thermal analysis for combination of photovoltaic and thermal system by varying concentration ratio. It was observed that at an increased concentration ratio, the electrical, as well as thermal performance, gets improved. According to the best knowledge of authors, some specific studies have been carried out using different zeotropic mixtures and solar collectors for solar powered ORC system which is summarized in Table 2 [1,13,14,16,34]. It is observed that, very few works have been done on energy and exergy analysis of solar collector and ORC system employing zeotropic mixture [1,21,34–36] [37];

The integration of reversed absorber hybrid PVTCPC system is a novel concept (Wang et al., 2011; [18,22,38,39]. The use of reverse absorber (PVTCPC) in low-grade solar heat source power generation for ORC system using zeotropic mixture with low global warming and negligible Ozone depletion potential is being rarely carried out. Therefore, a thermodynamic investigation for Solar powered ORC system system using heptane/R245fa as working fluid coupled with hybrid reversed absorber partially covered photovoltaic thermal compound parabolic concentrator (PVTCPC) system is decided to carry out.

First part of the research describes the introduction, research gaps and objectives. Second part deals with the description of the proposed system, simulation methodology and then third part describes about the mathematical Modelling of the proposed system. Moreover, manuscript describes the analyzed results and finally concludes it.

2. Proposed system description and simulation works

2.1. Basic ORC system

Fig. 1 depicts the basic solar powered ORC system. It consists of a solar collector (evaporator), expander, condenser and pump [14]. The working fluid is pressurized by the pump is shown in (1–2) process and enters to evaporator (solar collector) is shown by state (3). In evaporator, working fluid is heated and changes into the saturated vapor by absorbing incoming radiations ((2–4) process). The saturated working fluid vapor with raised pressure and temperature from evaporator move to expander at state (4) where it gets expand and generate power by reduction in enthalpy ((4–5) process). The working fluid with low pressure and temperature is condensed through condenser and convert into saturated liquid is shown by (5–1) process. The saturated liquid again sent back to the pump and thus cycle is completed. Fig. 2 shows T-s diagram of a basic ORC system when zeotropic working fluid is taken as working fluid. Fig. 3 shows the reversed absorber hybrid photovoltaic thermal compound parabolic concentrator. Its salient parts are; aperture surface, fluid flowing in tubes and metal absorber plate.

The solar beam radiations fall on the aperture area of concentrator and move ahead to the aluminum reflector (Fig. 5) and travels to the bottom part of CPC. The reflected rays further get reflected from the circular glass reflector and received by metal reversed absorber system. The photovoltaic modules are pasted at the base of the metal absorber and the fluid carrying copper tubes are brazed at the top of the system. The upper most portion of the tube in the plate system is glazed. The thermal energy is transferred indirectly from the back of solar cells to the absorber plate and directly from the non-packing area of the PV module which raises thermal energy of the fluid. Moreover, the convection losses are reduced significantly because reversed absorber system faces in

Table 1
Comparisons of thermo-physical properties for working fluids [43]] and [50].

Working fluids	Critical temperature (K)	Critical pressure (M Pa)	Normal boiling temperature (K)	GWP (100 years)	ODP	Types of working fluid	Flammable
Heptane	540.13	2.736	371.53	Low	0	Dry	Yes
R245fa	427.16	3.651	288.29 K	950	0	Dry	No
R134a	374.21	4.0593	247.08	120	0	Wet	No
Cyclohexane	553.64	4.075	353.89	Low	0	Dry	Yes
R152a	386.41	4.5168	249.13	124	0	Wet	No
Cyclopentane	511.69	4.515	322.4	Low	0	Dry	Yes

Table 2
Solar powered ORC system using various zeotropic mixtures and solar collectors.

Reference	Collector outlet temperature (°C)	Types of collector and their analysis	Zeotropic mixture	Remarks
[1]	85	The name of the collector and its thermo-dynamic analysis has not been given	R24fa/R152a	First law efficiency of pure R245fa is better than mixture R245fa/R52a.However, the economy of zeotropic mixture R245fa/R152a in terms of specific volume flow ratio is better than pure R245fa.
[13]	100	Combination of two flat plate collector	R245fa/R152a	At particular mass fraction the thermal performance of zeotropic mixture (R245fa/R152a) is better than pure R245fa.
[16]	The collector outlet temperature has not been given	The name of the collector and its thermodynamic analysis has not been reported.	Isopentane/ R245fa	Use of regenerator, collector outlet temperature, and the mass fraction of the zeotropic mixture are the important factor for solar thermal system.
[34]	150	Individual energy has not been carried out.	Cyclopentane/ Cyclohexane	Evaluation of overall solar power efficiency
[14]	125	Glazed reversed absorber conventional compound parabolic concentrator	Cyclohexane/ R24fa	Total and beam radiation, collector outlet temperature and mass fraction of the zeotropic mixture are salient factors based on energy and exergy analysis.

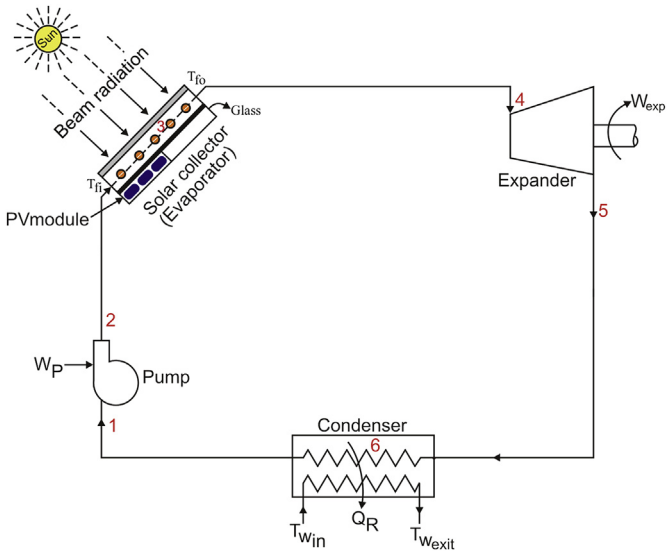


Fig 1. Schematic of basic ORC.

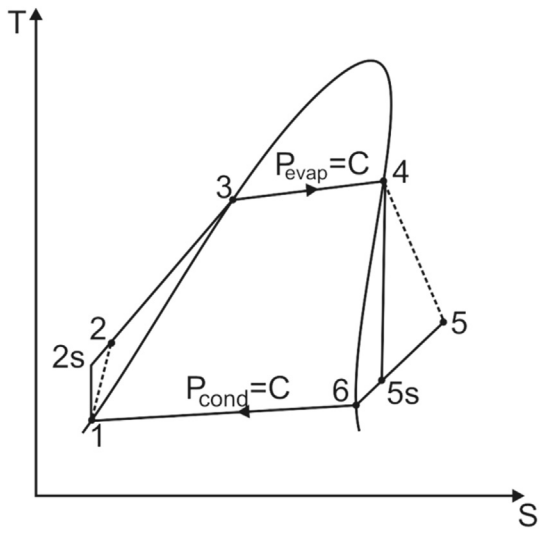


Fig 2. T-s diagram of basic ORC.

downward direction leading improvement in collector system performance.

The design and development of the PVTCPC collector gives the solution to existing energy shortage problems [40]. investigated experimentally performance of hybrid photovoltaic thermal compound parabolic concentrator (PVTCPC) system. It was found that such hybrid systems are cost effective owing to generation of both thermal and electrical energy as well as reduced number of solar cells [41]. studied concentrating photovoltaic thermal system. It was concluded that the system has improved thermal and electrical performance [42]. proposed PVTCPC system used for building integrated purposes. It was observed that CPC-PVT system have increased thermal and electrical efficiencies as compared to flat PVT system [28]. have suggested that PVTCPC collector produces both electrical and thermal energy at reduced cost.

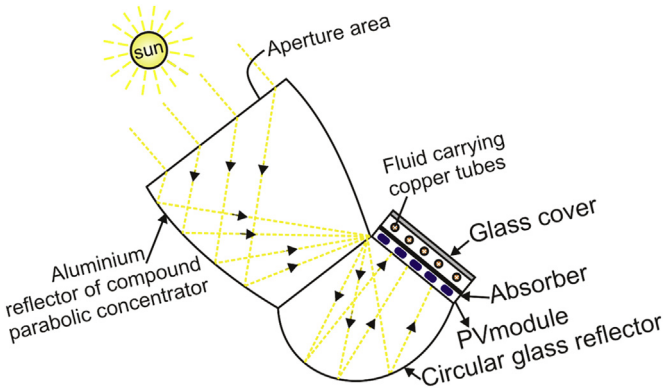


Fig 3. Reverse absorber PVT/CPC.

2.2. Thermal modelling of ORC system

The methodology of this paper has been described Fig. 4. On the basis of detail literature review the methodology of this study includes definition of the research problem, selection appropriate Working fluid, assumption in process parameters, selection of input conditions for ORC-PVTCPC system, determinations of thermodynamic properties at different desired location, application of balance of mass, energy and exergy for solar powered ORC-PVTCPC system and hourly overall evaluation of performance indicators for solar powered ORC system. The thermo-dynamic Modelling of the proposed solar powered system has been completed in MATLAB 2018a computer software. The hourly ambient parameters such as solar radiations and ambient temperature for the month of Jan 2017 of Delhi have been taken from India Meteorological Department (IMD), Pune. The properties of Working fluid can have been evaluated by MATLAB linking with REFPROP 9.0 [43]. The input

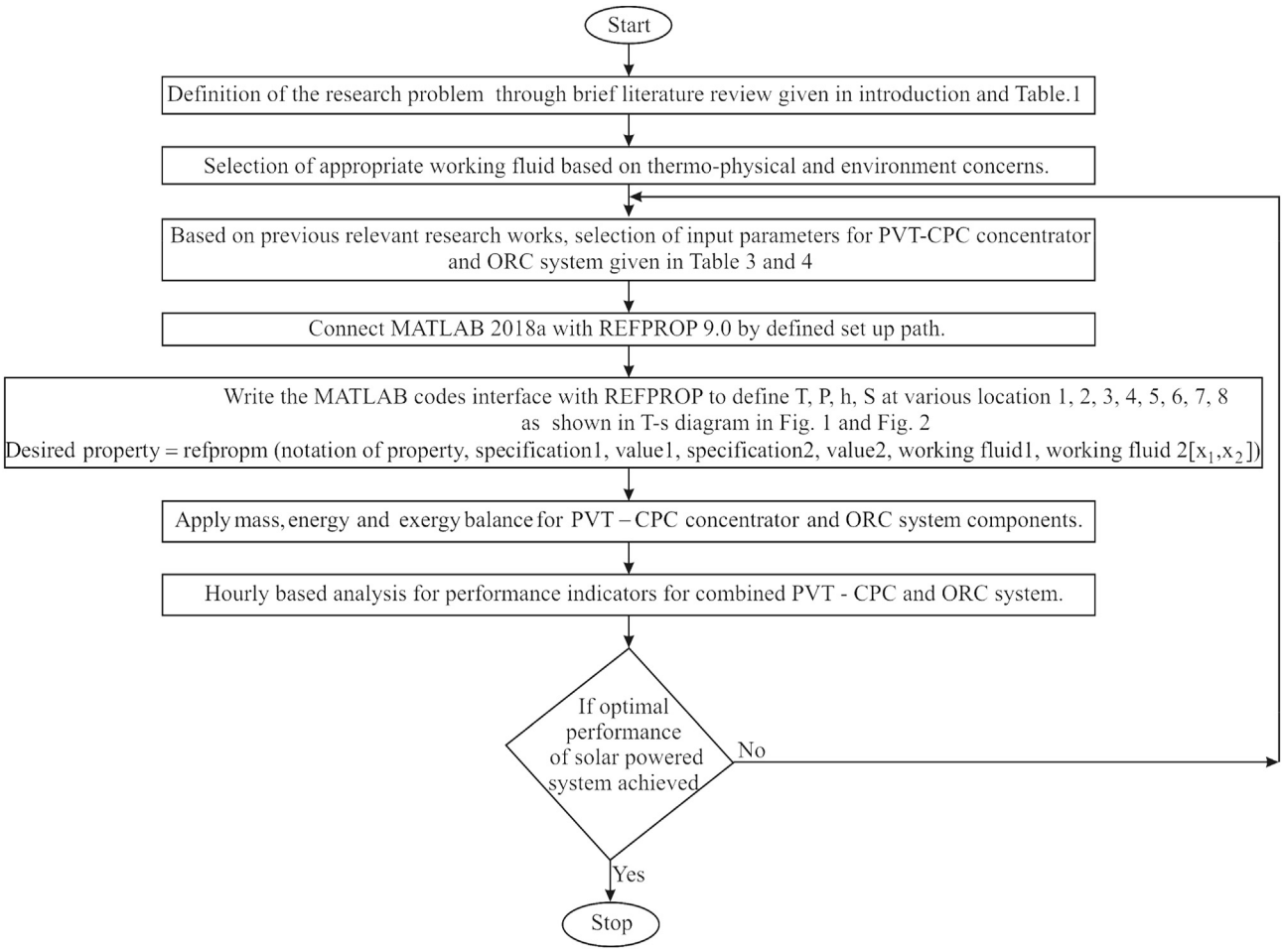


Fig 4. Simulation procedure.

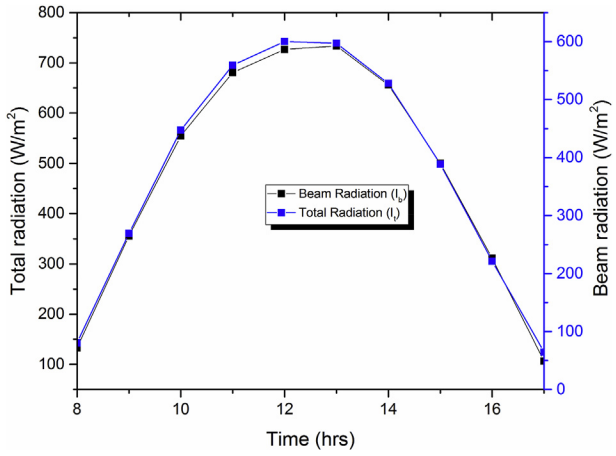


Fig 5. Hourly variation of radiations.

parameters for PVTCPC, and ORC system are given in Tables 3 and 4, respectively.

Assumptions made in order to simplify thermodynamic model of proposed system are given as [1,44].

- No pressure drops and heat losses in heat exchangers.
- The fractionating and shifting in composition for zeotropic mixture are neglected.
- No changes in kinetic and potential energy.

Thermodynamic relations for conservation of mass, energy and exergy balance for control volume system are given as [45]:

2.2.1. Conservation of mass

$$\frac{dm}{dt} = \dot{m}_i - \dot{m}_e \tag{1}$$

where '$\dot{m}$' shows mass flow rate, '*i*' and '*e*' denote the inlet and exit state.

2.2.2. Conservation of energy

$$\frac{dE}{dt} = \dot{Q} - \dot{W} + \sum_i m_i\left(h_i + \frac{V_i^2}{2} + gz_i\right) - \sum_e m_e\left(h_e + \frac{V_e^2}{2} + gz_e\right) \tag{2}$$

where '*E*' denotes energy of the system,' $\dot{Q}$' and '$\dot{W}$ 'denotes heat and work transfer rate, '*h*' denotes enthalpy, '*i*' refers inlet whereas 'e' exit.'*V*' denotes velocity of flowing stream where as '*Z*' height of flowing stream from datum.

Table 3
Input Design parameters and the constants used in modelling of PVTCPC concentrator system [28]] and [49]].

Design Parameters	
Parameters	value
$A_{ac}A_{am},A_a$	1 m^2, 1m^2, 2m^2
A_rA_{rm},A_{rc}	1m^2, 0.5m^2, 0.5m^2
b,b_o	0.5 m, 1 m
L_pL_g,L_i	0.002,0.003,0.100
m_f	0.025 kg/s
Angle of inclination	30^o
L_{rc},L_{rm}	1m,1m
Tube diameter	0.0125 m
Tube spacing	87.5 mm
Number of Tubes	10
Constants used in Modelling	
Constant	Values
C,C_f	2.8J/kgK,4179J/kgK
F'	0.9680
F_{rm},F_{rc}	0.9149,0.9149
F_{rm},F_{rc}	0.9149,0.9149
h_{pf}	100 W/m^2K
h_oh_i,h_{i1}	9.5W/m^2K,5.7 W/m^2K,5.8W/m^2K
K_pK_i,K_g	64 W/mK,0.166 W/mK,0.816 W/mK
PF_cPF_1,PF_2	0.9968,0.7240,0.9818
$U_{l,c}U_{L1},U_{L2}$	3.50 W/m^2K,1.54 W/m^2K,5.05 W/m^2K
$U_{l,m}$	4.96 W/m^2K
τ_g,β_c	0.95,0.89
σ	1000kg/m^3
α_p,α_c	0.8,0.9
ρ, ρ'	0.84, 0.84

Table 4
Input parameters for ORC system.

Input parameters	Values	References
Bubble point evaporation Temperature	Variable	[14]
Mass fraction of mixture component	0.1/0.9	[51]
Expander isentropic efficiency	0.80	[20]
Pump isentropic efficiency	0.70	[14]
Dew point condensation temperature	35 °C	[51]

2.2.3. Exergy balance

Net exergy destruction rate carried by any fluid stream in control volume is expressed as [46]:

$$\frac{dE_{cvD}}{dt}=\sum\left(1-\frac{T_o}{T_k}\right)\dot{Q}_k-\dot{W}+\dot{m}\left[\left(h_i-T_os_i\right)-\left(h_e-T_os_e\right)\right]-\dot{E}_D \tag{3}$$

where 'cv' defines control volume,'E_D' denotes exergy destruction of flow,'$\dot{m}$' denotes mass flow rate of flowing stream,'T_o' denotes ambient temperature, $\dot{Q}_k$ shows the heat transfer rate at temperature ''T_k' whereas 'k' shows location and 's' denotes entropy.

2.2.4. Energy analysis for ORC system

Thermodynamic relations for difference components are given as:

(i) Pump

The isentropic efficiency of the pump is:

$$(h_{2s}-h_2)=\eta_p(h_2-h_1) \tag{4}$$

where, subscript 's' denotes the isentropic.

The pump work is evaluated by the expression:

$$W_p=\dot{m}\,(h_2-h_1) \tag{5}$$

(ii) Expander (Turbine)

The isentropic efficiency is evaluated as:

$$h_4-h_5=\eta_{\exp}(h_4-h_{5s}) \tag{6}$$

The expander work is evaluated as:

$$W_{\exp}=\dot{m}\,(h_4-h_5) \tag{7}$$

(iii) Evaporator (Solar collector)

Heat gain in evaporator (Collector) is calculated as:

$$Q_s=\dot{m}\,(h_4-h_2) \tag{8}$$

(iv) Condenser

Heat rejection from the condenser is calculated as:

$$Q_r=\dot{m}\,(h_5-h_1) \tag{9}$$

(v) Net work done is evaluated as:

$$W_{net}=\left(W_{\exp}-W_p\right) \tag{10}$$

(vi) Thermal efficiency of the ORC is as:

$$\eta_{th}=\frac{(W_{\exp}-W_p)}{Q_S} \tag{11}$$

2.2.5. Exergy analysis of ORC system

(i) The destruction of exergy in pump is as:

$$\dot{E}\,D_P=\dot{m}\,T_o(s_2-s_1) \tag{12}$$

where, 'T_o' denotes ambient temperature in (K).

(ii) The destruction of exergy in evaporator is calculated as:

$$\dot{E}\,D_{ev}=\dot{m}\,T_o\left[(s_4-s_2)-\frac{(h_4-h_2)}{T_{hs}}\right] \tag{13}$$

where 'T_{hs}' is the temperature of the heat source in 'K'.

(iii) The condenser exergy destruction in condenser is calculated as:

$$\dot{E}\,D_c = \dot{m}\,T_o\left[(s_5 - s_1) - \frac{(h_5 - h_1)}{T_o}\right] \tag{14}$$

(iv) The total exergy destruction in ORC system is given as:

$$\dot{E}\,D_{Total} = (\dot{E}\,D_{\exp} + \dot{E}\,D_p + \dot{E}\,D_c + \dot{E}\,D_{ev}) \tag{15}$$

(v) The exergy input to the ORC system is evaluated as:

$$EX_{in} = \left[Q_s\left(1 - \frac{T_o}{T_h}\right) + W_p\right] \tag{16}$$

where 'EX_{in}' is input exergy in 'W'.

(vi) The exergetic efficiency is obtained by the expression as follows:

$$\eta_{ex,ORC} = \frac{W_t}{EX_{in}} \tag{17}$$

'W_t' denotes expander output in 'W'

2.2.6. Energy analysis of (PVTCPC) system

(i) The concentration ratio of the concentrator is evaluated as [28]:

$$CR = \frac{Area\ of\ the\ absorber(A_a)}{Area\ of\ the\ receiver\ (A_r)} \tag{18}$$

$$(\alpha\tau)_{c,eff} = \rho(\alpha_c\tau_g\beta_c - \eta_m)CR \tag{19}$$

(ii) The expression to compute the exit fluid temperature of the (PVTCPC) concentrator is expressed as [47]:

$$T_{fo} = \left[\frac{PF_c(\alpha\tau)_{c,eff}I_b}{U_{l,c}} + T_a\right]\left\{\frac{A_{rc}U_{l,c}\frac{\dot{m}_f c_f}{U_{l,c}A_{rc}}\left[1 - \exp\left\{\frac{-F'U_{l,c}A_{rc}}{\dot{m}_f c_f}\right\}\right]}{\dot{m}_f c_f}\right\} + \left[\left[\frac{PF_2(\alpha\tau)_{m,eff}I_b}{U_{l,m}} + T_a\right]\left\{\frac{A_{rm}U_{l,m}F_{rm}}{\dot{m}_f c_f}\right\} + T_{fi}\exp\left\{\frac{-F'U_{l,m}A_{rm}}{\dot{m}_f c_f}\right\}\right]\exp\left\{\frac{-F'U_{l,c}A_{rc}}{\dot{m}_f c_f}\right\} \tag{20}$$

where 'PF_c' denotes penalty factor of glass cover, 'I_b' beam radiation in (W/m^2), 'A_{rc}', 'A_{rm}' are receiver and PV module area respectively (m^2),'$\dot{m}_f$' water mass flow rate (kg/s),'c_f' shows specific heat of the water (J/kg K) 'PF_c', is penalty factor for the portion covered by the glass, '$U_{l,c}$' is overall loss coefficient for the cover (W/m^{2}K), 'F_{rm}' shows rate flow factor, 'F'' denotes efficiency factor for flat plate collector, 'PF_2' second penalty factor for absorber/receiver.

The estimated exit fluid temperature of the (PVTCPC) concentrator has been validated with the experimental observations given as [47]:

(iii) The expression to calculate the thermal energy obtained from (PVTCPC) system in (W) is as:

$$Q_{th} = \dot{m}_f\,c_f\left(T_{fo} - T_{fi}\right) \tag{21}$$

(iv) Equation (22) to compute the thermal efficiency of the concentrator [48].

$$\eta_{th,cl} = \frac{\dot{m}_f\,c_f\left(T_{fo} - T_{fi}\right)}{(I_b A_a)} \tag{22}$$

2.2.7. Exergy analysis of (PVTCPC) system

(i) The exergy of (PVTCPC) working fluid,'$\dot{E}_{x,cl}$', in W is given as [48]:

$$E_{x,cl} = \dot{m}_f c_f\left(T_{fo} - T_{fi}\right) - \dot{m}_f c_f(T_o + 273)\ln\frac{\left(T_{fo} + 273\right)}{\left(T_{fi} + 273\right)} \tag{23}$$

(ii) The exergy supplied by sun in Watt, is:

$$E_{x,\mathrm{sun}} = [A_a I_b + A_r I_t]\left[1 - \frac{4}{3}\left(\frac{T_o}{T_s}\right) + \frac{1}{3}\left(\frac{T_o}{T_s}\right)^4\right] \tag{24}$$

where, 'T_s' is temperature of the sun (6000 K), and 'I_t' denotes total beam radiation in (W/m^2).

(iii) The exergy efficiency of the compound parabolic concentrator is:

$$\eta_{ex,ccpc} = \frac{E_{x,cl}}{E_{x,sun}} \tag{25}$$

The electrical power developed by the PV module system is evaluated by equation (26) in 'W' [28]:

$$\dot{E}_{e,pv} = \eta_{\mathrm{pv}} I_b \rho\rho' A_{am} \tag{26}$$

where ρ′ denotes the reflectivity and 'A_{am}' shows the area of the module. However, the power consumed in pumping process or by motor (DC) is evaluated by the expression [49].

$$\dot{E}_{e,pv,net} = \eta_{\mathrm{pv}} I_b \rho\rho' A_{am} - W_p \tag{27}$$

(iv) The electrical exergy of (PVT CPC) system from the photovoltaic module in 'W' is:

$$\dot{E}_{pv} = \eta_{ex,\mathrm{pv}} I_b \rho A_{am} \tag{28}$$

'$\eta_{ex,pv}$' exhibit the exergy efficiency of the PV module.

(v) The overall exergy of (PVT CPC) system is calculated by combining thermal exergy and electrical exergy:

$$\eta_{ov,ex,pvtcpc} = \eta_{ex,ccpc} \times \eta_{ex,pv} \tag{29}$$

The combined thermal efficiency efficiency of the solar collector as well as organic Rankine cycle system is calculated as [13]:

Table 5
Validation of present basic ORC with model of [26]].

Working fluids	Mass fractions	Thermal efficiency (η_{th}) Present Work	Thermal efficiency (η_{th}) Wu et al.(2016)	Deviation (%)
R227ea/R245fa	0.1/0.9	8.95	9.07	0.45
Butane/R245fa	0.1/0.9	8.94	8.99	0.55
R227ea/R245fa	0.1/0.9	8.95	9.07	0.88

$$\eta_{ov,th,SORC} = \eta_{th,cl} \times \eta_{ORC} \quad (30)$$

Similarly overall exergy efficiency of solar powered ORC is as:

$$\eta_{ov,ex,\text{sorc}} = \eta_{ov,ex,\text{pvtcpc}} \times \eta_{ex,orc} \quad (31)$$

3. Results and discussion

Thermal efficiency from First Law of Thermodynamics is evaluated in present research which shows the fair agreement with results obtained by Ref. [26] in terms of percentage deviation (less than 1%) as shown in Table 5. Further, Table 6 shows the comparison of the outlet temperature for the PVTCPC concentrator of the present study with [28] at different concentration ratio with same given input parameters (Table 3). It clearly explains the better match of collector outlet temperature of the reversed absorber hybrid PVTCPC at various concentration ratio 1, 2 and 3 of the present study with [28]. The data for fluid inlet temperature, total radiation and beam radiation of the collector has been collected from (Indian metrological department (IMD), Pune) for January 17, 2017 for Delhi, India.

3.1. Hourly variation in beam radiations, total radiations and collector outlet temperatures

Fig. 5 exhibits hourly (08:00 h to 16: 00 h), beam radiations, total radiations for January 17, 2017, Delhi. The maximum value of total radiation (733 W/m^2) occurs at 12:00 noon whereas beam radiation (600 W/m^2) at 013:00 h (Fig. 5). Maximum collector temperature 125.45 °C occurs at 13:00 h for concentration ratio 6 (Fig. 6). The collector outlet temperature increases when concentration ratio increases. It can explain this, when concentration ratio the aperture area gets increased which facilitate to significant absorption of beam radiation over the aperture surface. The collector outlet temperature increases by increasing concentration ratio up-to 13:00 h and then decreases because after 13:00 h beam radiation decreases.

3.2. Effect of mass fraction of zeotropic mixture on glides in temperature for evaporator and condenser

Fig. 7 exhibits the variation in condensation and evaporation temperature glides when mass fraction of considered zeotropic mixture heptane/R245fa changes from (0.1–0.9). The maximum temperature glide (51.64 K) in evaporator occurs at 0.58/0.42 mass fraction whereas condensation temperature exhibits its maximum value (49.26 K) at (0.63/.37) mass fraction.

In this zeotropic mixture the first working fluid is heptane and second R245fa. The temperature glide is isobaric temperature difference between heat transferring and heat receiving fluid during evaporation and condensation process. When mass fraction of zeotropic mixture changes the boiling point of the mixture component also changes. The difference in boiling point of the mixture component creates temperature glide [23]. The temperature glide becomes largest when highest difference between the boiling point temperatures of the mixture component appear [50]. Thus, the temperature glide in condensation and evaporation process is dependent on variation in mass fraction of the zotropic mixture component. The temperature glide brings a good thermal match during phase change which improves the thermal and exergetic efficiency of the system [51,52].

3.3. Heat gain in collector

Fig. 8 exhibits the hourly heat gain in collector at mass fraction (0.1/0.9). The highest heat gain (3.21×10^5 W) occurs at 13:00 h for concentration ratio (6). It occurs because when concentration ratio increases the aperture area of the concentrator also increases whereas receiver area of the photovoltaic-cell decreases. Therefore, more beam radiations are entrapped and get absorbed over the increased surface area of the aperture leading to enhanced heat gain in the collector.

3.4. Overall thermal efficiency of solar ORC system

Fig. 9 shows variation of combined thermal efficiency of solar ORC system at a various concentration ratio of the concentrator for the mass fraction 0.1/0.9. Maximum value of overall thermal efficiency (7.8%) occurs at 13:00 h at collector concentration ratio 6. It occurs because at the same time and with same concentration ratio, when collector exhibits a maximum value of heat gain (3.21×10^5 W). Moreover, owing to maximum heat gain in collector working fluid temperature inlet to the expander increases leading

Table 6
Comparison of collector outlet temperature of present study with [28]] for PVTVCPC system at different concentration ratio.

Time (hrs)	Fluid inlet Temperature of the collector T_{fi} (°C)	Collector outlet temperature (°C) at concentration ratio (1)		Collector outlet temperature (°C) at concentration ratio (2)		Collector outlet temperature (°C) at concentration ratio (2)	
		Present study	[28]	Present study	[28]	Present study	[28]
08:00	9.90	11.64	12.01	14.45	14.23	17.31	17.27
09:00	9.90	16.51	16.41	24.47	24.51	32.77	33.01
10:00	9.90	20.72	21.12	33.35	33.20	46.83	47.01
11:00	8.60	21.96	22.23	37.50	37.9	54.34	55.32
12:00	8.40	22.70	23.01	39.33	38.23	57.48	57.11
13:00	9.70	24.09	27.09	40.83	41.11	59.10	58.89
14:00	12.60	25.66	26.01	40.84	41.21	57.27	57.10
15:00	15.00	25.15	24.10	37.02	36.01	49.64	50.01
16:00	18.00	24.22	23.59	32.32	31.87	40.77	41.02

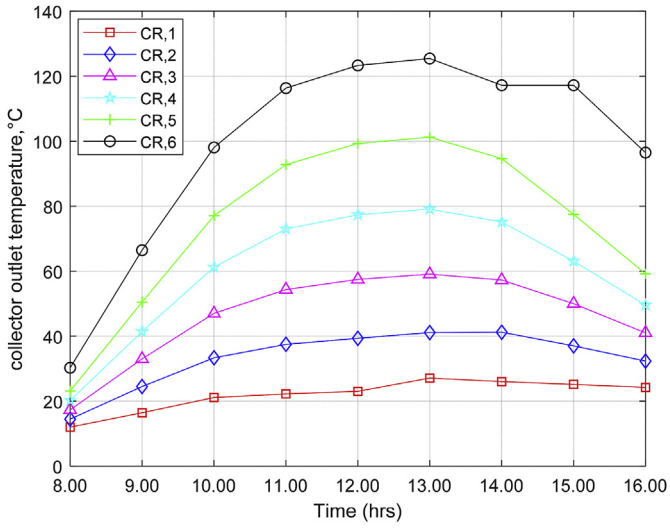


Fig 6. Hourly collector outlet temperatures at various concentration ratios.

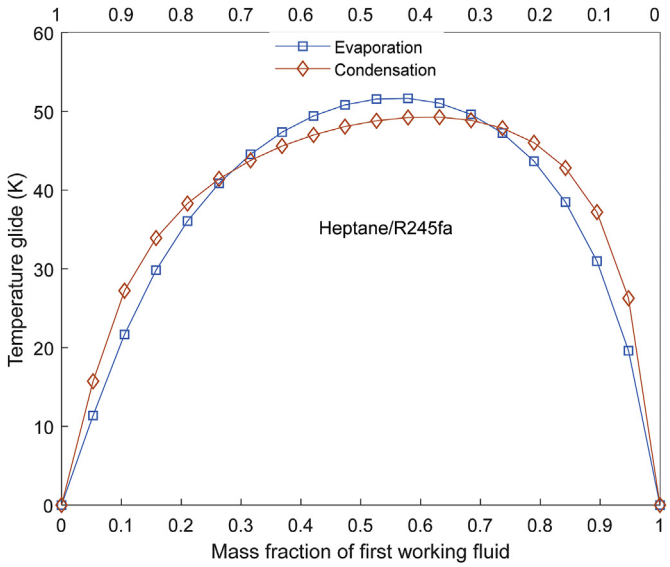


Fig 7. Condensation and Evaporation temperature glide.

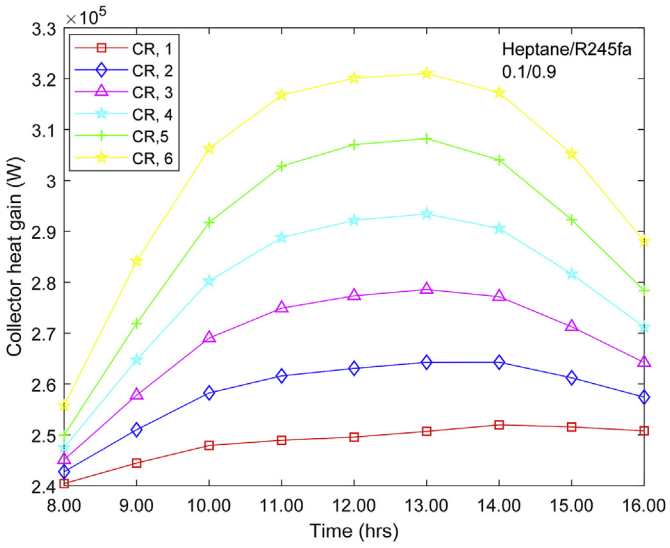


Fig 8. Hourly heat gain at various concentration ratio.

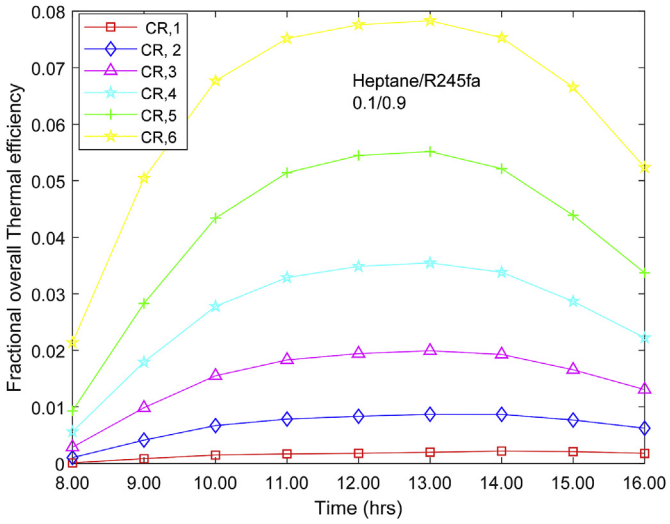


Fig 9. Hourly overall thermal efficiency at various concentration ratio.

to improvement in thermal efficiency.

3.5. Overall exergetic efficiency of solar ORC system

Fig. 10 shows variation of overall exergetic efficiency on hourly basis at 0.1/0.9 fixed mass fraction of the mixture. The overall exergetic efficiency (14.38%) is maximum at 13:00 h for mass fraction 0.1/0.9. It occurs, because with increasing concentration ratio (1–6) heat accumulation in collector gets increased. Therefore the temperature of the working fluid inside the evaporator (collector) increases. Due to the increased working fluid temperature in evaporator, the mean temperature of heat addition gets increased. Therefore the temperature difference between heat source and heat absorbing fluid gets reduced. The reduction in temperature difference during heat transfer improves the exergetic efficiency and reduces the exergy destruction.

3.6. Loss of exergy in components of ORC system

Fig. 11 shows the extent of destruction in exergy occurred in ORC system components at the mass fraction (0.1/0.9). It is observed that highest exergy destruction (40%) in evaporator followed by the condenser (38%), expander (21%) and pump (<1%). The major part of the exergy destruction has been developed in condenser and evaporator due to larger temperature difference in heat addition and heat rejection processes.

4. Conclusions

In this present work, hourly (8–16, hrs) thermodynamic investigation for Solar powered ORC system has been carried out coupled with hybrid reversed absorber partially covered PVTCPC system using heptane/R245fa as working fluid in ORC system. Further, the effect of variation in concentration ratio has also been observed on overall system performance of Solar powered ORC. The following conclusions are drawn from aforementioned results and discussion.

- The collector gets maximum total radiation (733 W/m^2) at 12:00 h whereas maximum beam radiation is available (600 W/m^2) at 13.00 h.

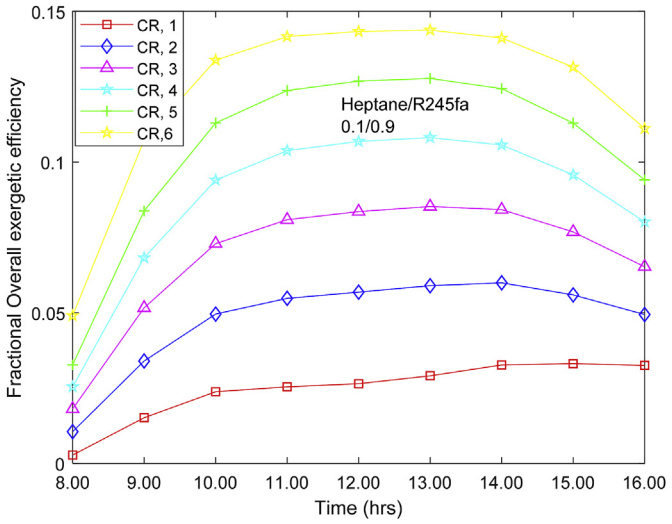


Fig 10. Hourly overall exergetic efficiency at various concentration ratio.

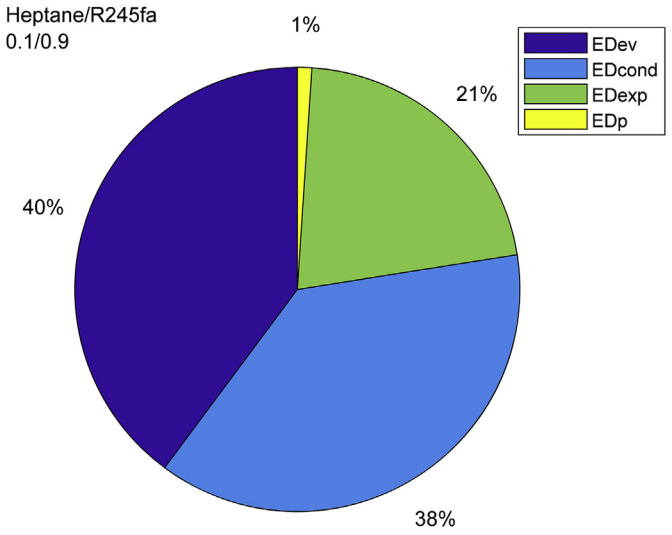


Fig 11. Exergy destruction in ORC components.

- At the mass fraction 0.58/0.42 the zeotropic mixture Hepatne/R245fa exhibits highest temperature glide (51.64 K) in evaporator whereas in condenser (49.26 K) at 0.63/0.37 mass fraction.
- The outlet fluid temperature and heat gain of the collector are improved by increasing concentration ratio. At concentration ratio 6 the collector exhibit highest outlet temperature (125.45 °C) and maximum heat gain (3.21×10^5 W) for 13:00 h.
- The combined thermal efficiency of Solar powered ORC system is 7.8% and exergetic efficiency of 14.38% are achieved at 13:00 h for concentration of ratio of 6.
- The evaporator exhibits (40%) of total exergy destruction and (38%) contributed by condenser, (21%) appeared in expander and approximately (1%) in pump at (0.1/0.9) mass fraction of zeotropic mixture.

This study deals with the combined application of photovoltaic thermal and compound parabolic concentrator using low-grade heat source for power generation using safe reliable zeotropic mixture. The findings of the study are fruitful in identifying salient aspects while making mathematical Modelling and also for installation of experimental setup.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

Authors are highly thankful to Bhagwan Parshuram Institute of Technology New Delhi (India), Jamia Millia Islamia, New Delhi (India), Bennett University, Greater Noida - 201310 (India), Department of Mechanical Engineering, Delhi Technological University, Delhi (India) and National Institute of Electronics & Information Technology, Calicut (India) and for providing computational resources in carrying out present research work.

Nomenclature

A	area (m^2)
B	receiver breadth (m)
b_o	breadth of the receiver area
L_{rm}	length for the receiver covered by PV module
L_{rc}	length covered by the glass for the receiver
h	specific enthalpy (kJ/kg)
h	coefficient of heat transfer (W/m^2K)
$\dot{Q}$	rate of heat transfer (W)
s	specific entropy (kJ/kg K)
T	temperature (K)
$\dot{W}$	work transfer rate (W)

Greek letters

$(\alpha\tau)_{cl,eff}$	product of effective absorptivity and transmissivity of the collector
ρ	reflectivity
τ	transmittivity
α	absorptivity
η	efficiency
β	packing factor

Subscripts

a	aperture
b	beam
c	condenser
c	solar cell
ccpc	conventional compound parabolic concentrator
cl	collectror
c	cell
ex, x	exergetic
ev	evaporator
g	glass
in	inlet
f	fluid
f_i	fluid inlet
f_o	fluid outlet
m	module
o	ambient, overall
ov	overall
p	pump
p	palte
pv	photovoltaic
r	receiver
s	supply
$\dot{E}\,D$	Exergy destruction

th thermal

Abbreviations

CR	Concentration Ratio
GWP	global warming potential
ORC	organic Rankine cycle
ODP	ozone depletion potential
(PVTCPC)	photovoltaic thermal compound parabolic concentrator
SORC	solar ORC

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.renene.2019.10.018.

References

[1] X.D. Wang, L. Zhao, Analysis of zeotropic mixtures used in low-temperature solar Rankine cycle for power generation, Sol. Energy 83 (2009) 605–613.
[2] F. Heberle, M. Preissinger, D. Brueggemann, Zeotropic mixtures as working fluids in Organic Rankine Cycles for low-enthalpy geothermal resources, Renew. Energy 37 (2012) 364–370.
[3] A. Demribas, Potential applications of renewable energy sources, biomass combustion problems in boiler power system and combustion related environment issues, Prog. Energy Combust. Sci. 31 (2) (2005) 171–192.
[5] M. Yang, R. Yeh, Analysis of optimization in an OTEC plant using organic Rankine cycle, Renew. Energy 68 (2014) 25–34.
[6] M. Wang, R. Jing, H. Zhang, C. Meng, N. Li, Y. Zhao, An innovative organic rankine cycle (ORC) based ocean thermal energy conversion (OTEC) system with performance simulation and multi-objective optimization, Appl. Therm. Eng. 145 (2018) 743–754.
[7] H. Zhai, Q. An, L. Sh, V. Lemort, S. Quolin, Categorization and analysis of heat sources for organic Rankine cycle, Renew. Sustain. Energy Rev. 64 (2016) 790–805.
[8] A. Ramos, M.A. Chatzopoulou, J. Freeman, C.N. Markides, Optimisation of a high-efficiency solar-driven organic Rankine cycle for applications in the built environment, Appl. Energy 228 (2018) 755–765.
[9] C.P. Jawahar, R. Saravanan, Joan Carles Bruno, Alberto Coronas, Simulation studies on gax based Kalina cycle for both power and cooling applications, Appl. Therm. Eng. 50 (2) (2013) 1522–1529.
[10] Y.D. Goswami, Solar thermal power technology: present status and ideas for the future, Energy Sources 20 (2) (1998) 137–145.
[11] C. Zamfierescu, I. Dincer, Thermodynamic analysis of novel ammonia-water trilateral Rankine cycle, Termochim Acta 477 (1) (2008) 7–15.
[12] T.C. Hung, T.Y. Shai, S.K. Wang, A review of organic Rankine cycle (ORCs) for the recovery of low-grade waste heat, Energy 22 (1997) 661–667.
[13] X.D. Wang, L. Zhao, J.L. Wang, W.Z. Zhang, X.Z. Zhao, W. Wu, Performance evaluation of low-temperature solar Rankine cycle system utilizing R245fa, Sol. Energy 84 (2010) 353–364.
[14] D. Tiwari, A.F. Sherwani, D. Atheaya, A. Arora, Energy and exergy analysis of solar driven recuperated organic Rankine cycle using glazed reverse absorber conventional compound parabolic concentrator (GRACCPC)system, Sol. Energy 155 (2017) 1431–1442.
[15] G. Venkatarathnam, G. Mokashi, S. Srinivasa Muthy, Occurrence of pinch point in condenser and evaporation for zeotropic mixtures, Int. J. Refrig. 19 (6) (1996) 361–368.
[16] J.J. Bao, L. Zhao, W.Z. Zhang, A novel auto-cascade low-temperature solar Rankine cycle system for power generation, Sol. Energy 85 (2011) 2710–2719.
[17] N. Yamada, M.N.A. Mohmad, T.T. Kien, Study on thermal efficiency of low-temperature organic Rankine cycle using HFO-R1234yf, Renew. Energy 41 (2012) 368–375.
[18] M. Wang, J. Wang, Y. Zhao, Y. Zhao, P. Zhao, Y. Dai, Thermodynamic analysis and optimization of solar-driven regenerative organic Rankine cycle (ORC) based on flat-plate solar collector, Appl. Therm. Eng. 50 (2013) 816–825.
[19] X.B. Bu, H.S. Li, L.B. Wang, Performance analysis and working fluids selection of solar powered organic Rankine-vapor compression ice maker, Sol. Energy 95 (2013) 271–278.
[20] D. Geng, Y. Du, R. Yang, Performance analysis of an organic Rankine cycle for reverse osmosis desalination system using zeotropic mixtures, Desalination 381 (2016) 38–46.
[21] B.F. Tchanche, G. Papdakis, G. Lambrinos, A. Frangoudakis, Fluid selection for low temperature solar organic Rankine cycle, Appl. Therm. Eng. 29 (2009) 2468–2476.
[22] R. Rayegan, Y.X. Tao, A procedure to select working fluids for solar organic Rankine cycles (ORCs), Renew. Energy 36 (2011) 659–670.
[23] G.B. Abadi, K.C. Kim, Investigation of organic Rankine cycles with zeotropic mixtures as working fluid: advantages and issues, Renew. Sustain. Energy Rev. 73 (2017) 100–1013.
[24] Z. Miao, K. Zhang, M. Wang, J. Xu, Thermodynamic selection criteria of zeotropic mixtures for subcritical organic Rankine cycle, Energy 167 (2018) 484–497.
[25] H. Chen, D. Yogi Goswami, E. K Stefanakos, A review of thermodynamic cycle and working fluids for conversion of low-grade heat, Renew. Sustain. Energy Rev. 14 (2010) 3059–3067.
[26] Y. Wu, Y. Zhu, L. Yu, Thermal and economic performance analysis of zeotropic mixture for organic Rankine cycles, Appl. Therm. Eng. 96 (2016) 57–63.
[27] V.K. Goel, R. Chandra, R. Chaudhari, A study on the performance of a Two-absorber Reverse flat-plate collector, Energy Convers. Manag. 27 (1987) 335–341.
[28] D. Atheaya, A. Tiwari, G.N. Tiwari, I.M. Al-Helal, Performance evaluation of inverted absorber photovoltaic thermal compound parabolic concentrator (PVT-CPC): constant flow rate mode, Appl. Energy 167 (2016) 70–79.
[29] A. Kumar, O. Prakash, A. Dube, A review on progress of concentrated solar power in India, Renew. Sustain. Energy Rev. 79 (2017a) 304–307.
[30] A. Kumar, O. Prakash, A. Dube, A review on technology and promotional initiatives for concentrated solar power in world, Int. J. Ambient Energy 39 (2017b) 297–316.
[31] E.C. Kern Jr., M.C. Russel, Combined photovoltaic and thermal hybrid collector system, in: Proceeding of the 13th IEEE Photovoltaic Specialistic, 1978, pp. 1153–1157. Washington D.C. USA.
[32] S. Jakhar, M.S. Soni, Experimental and theoretical analysis of glazed tube-and-sheet Photovoltaic/Thermal system with earth water heat exchanger cooling, Energy Convers. Manag. 153 (2017) 576–588.
[33] H. Chen, J. Ji, Y. Wang, W. Sun, G. Pei, Z. Yu, Thermal analysis of high concentration photovoltaic/thermal system, Sol. Energy 107 (2014) 372–379.
[34] E. Baldasso, J.G. Andreasen, A. Modi, F. Haglind, A. Stoppato, Performance Analysis a of Solar Driven Organic Rankine Cycle Using Multi-Component Working Fluids, 2015.
[35] J. Freeman, K. Hellgarddt, C.N. Markides, An assessment of solar-powered organic Rankine cycle systems for combined heating and power in UK domestic applications, Appl. Energy 138 (2015) 605–620.
[36] M. Ashouri, M.H. Ahmadi, S.M. Pourkiaei, F.R. Astaraei, R. Ghasempour, T. Ming, J.H. Hemanti, Exergy and exergo-economic analysis and optimization of solar double pressure organic Rankine cycle, Therm. Sci. Eng. Prog. 6 (2017) 72–86.
[37] M. Villarini, R. Tascioni, A. Arteconi, L. Cioccolanti, Influence of the incident radiation on the energy performance of two small-scale solar Organic Rankine Cycle trigenerative systems: a simulation analysis, Appl. Energy 242 (2019) 1176–1188.
[38] H.U. Helvaci, Z.A. Khan, Thermodynamic Modelling and analysis of solar organic Rankine cycle employing thermo-fluids, Energy Convers. Manag. 138 (2017) 493–510.
[39] L. Cioccolanti, S.R. Hamedani, M. Vilarani, Environmental and energy assessment of a small-scale solar Organic Rankine Cycle trigeneration system based on Compound Parabolic Collectors, Energy Convers. Manag. 198 (2019) 111829.
[40] M. Brogen, P. Nostell, B. Karlsson, Optical efficiency of a PV thermal hybrid CPC module for high latitudes, Sol. Energy 69 (2000) 173–185.
[41] J.S. Coventry, Performance of a concentrating photovoltaic/thermal solar collector, Sol. Energy 78 (2005) 211–222.
[42] L. Guiqiang, P. Gang, Y. Su, Z. Xi, J. Jie, Preliminary study based on building integrated compound parabolic concentrators (CPC) PV/thermal technology, Energy Procedia 14 (2012) 343–350.
[43] E.W. Lemon, M.L. Huber, Mo Mclinden, NIST Standard Reference Database 23: Reference Fluid Thermodynamic and Transport properties-REFPROP.9.0, 2010.
[44] W. Su, Y. Hwang, S. Deng, L. Zhao, D. Zhao, Thermodynamic Performance Comparison of Organic Rankine Cycle between Zeotropic Mixtures and Pure Fluids under Open Heat Source, vol. 165, 2018, pp. 720–7373.
[45] M.J. Moran, H.N. Shapiro, Fundamental of Engineering Thermodynamics, sixth ed., John Willey & Sons Ltd, UK, 2010.
[46] Y.A. Cengel, M.A. Boles, Thermodynamics: an Engineering Approach, seventh ed., Tata McGraw Hill, New York, 2011.
[47] D. Atheaya, A. Tiwari, G.N. Tiwari, I.M. Al-Helal, Analytical characteristics equation for partially covered photovoltaic (PVT) compound parabolic concentrator (CPC), Sol. Energy 111 (2015) 176–185.
[48] G.N. Tiwari, Solar Energy: Fundamentals, Design, Modelling and Applications, Alpha Science International, UK, 2015.
[49] R. Tripathi, G.N. Tiwari, Annual performance evaluation (energy and exergy) of fully covered concentrated photovoltaic thermal (PVT) water collector: an experimental validation, Sol. Energy 146 (2017) 180–190.
[50] L. Zhao, J. Bao, Thermodynamic analysis of organic Rankine cycle using zeotropic mixtures, Appl. Energy 130 (2014) 748–756.
[51] Y.R. Li, M.T. Du, S.Y. Wu, C. Liu, Potential of organic Rankine cycle using zeotropic mixture as working fluids for waste heat recovery, Energy 77 (2014) 509–519.
[52] Q. Liu, Y. Duan, Z. Yang, Effect of condensation temperature glide on performance of organic Rankine cycle with zeotropic mixture fluids, Appl. Energy 115 (2014) 394–404.

Suhash Chandra Dutta Roy

Topics in Signal Processing

Analog and Digital

Springer

Topics in Signal Processing

Suhash Chandra Dutta Roy

Topics in Signal Processing

Analog and Digital

Springer

Suhash Chandra Dutta Roy
Indian Institute of Technology Delhi
New Delhi, Delhi, India

ISBN 978-981-13-9531-4 ISBN 978-981-13-9532-1 (eBook)
https://doi.org/10.1007/978-981-13-9532-1

© Springer Nature Singapore Pte Ltd. 2020
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Dedicated
to my family, comprising
Sudipta, my wife,
Sumantra, my elder son,
Sroddha, my elder daughter-in-law,
Shoubhik, my younger son,
Parul, my younger daughter-in-law,
Soham, my elder grandson,
and
Malhar (Tutu), my younger grandson.

Acknowledgements

I owe so much to so many that it is not possible to acknowledge all of them in this limited space. To the ones whose credit I have not been able to mention, I ask for forgiveness.

To the President of the IETE, Prof. (Dr.) K. T. V. Reddy, I am indebted for his enthusiasm and support.

To the successive Publication Committee Chairpersons and Members of the Governing Council, I would like to say a big thank you for approving the proposal for writing this book.

To Shrimati Sandeep Kaur Mangat, I would like to express my deep gratitude for successfully leading this project to a conclusion, with her immense patience and love and passion for work, much beyond her duty and for carrying out all the necessary hard work.

To all the workers of the Publications Section and the Secretariat in general, I have no words to convey my heartfelt appreciation for their diligent work.

To my students—particularly, the research students, I would like to express my love and best wishes for making me learn many things. In fact, I learnt more from them than they from me.

Suhash Chandra Dutta Roy

About This Book

This book, in common with my first book, is for you—students, teachers and researchers. I hope you would enjoy reading it.

Starting from 1962, I have written and published a large number of articles in IETE Journals. At various points of time, starting from 1982, I have been requested by students,as well as teachers and researchers, to publish a book of collected reprints, appropriately edited and sequenced. Of late, this request has intensified, and the demand for reprints of some of the tutorial papers I wrote has increased considerably, not only from India, but also from abroad, because of my five video courses related to Circuits, Systems and Signal Processing (CSSP) successfully uploaded by NPTEL on YouTube. I thought it would be a good idea to venture into such a project at this time. This is the reason for my first book, *Circuits, Systems and Signal Processing*, published by Springer in 2018. Its performance encouraged me to venture into working on a second book, entitled *Topics in Signal Processing*, and I am glad that Springer has agreed to publish this book also. As in the first book, this also is a book for you, students, teachers and researchers in the subject Signal Processing. It is a collection of research papers published by me in collaboration with some of my colleagues and students in IETE Journals, duly edited and sequenced by me.

As you would notice, I have written this book also in a conversational style to make you feel at ease while reading it.

The contents of the book are divided into five parts—Parts I–IV dealing with Analog Signal Processing and Part V dealing with Digital Signal Processing. There are a total of 28 chapters. For details, see Contents.

I hope students and teachers will benefit from the book. It is only then that I shall feel adequately rewarded. For any mistakes/confusions/clarifications, please feel free to contact me via email at s.c.dutta.roy@gmail.com. I take such mails on top priority, I assure you.

Happy learning!

New Delhi, India Suhash Chandra Dutta Roy

Contents

Part I Analog Signal Processing: Signal Processing by Passive Networks

1 Finite Zero Butterworth and Chebyshev Filters 3
References 5

2 Bandwidth Enhancement with the Asymmetrical Bridged T-coil Network 7
2.1 Introduction 7
2.2 Analysis of the General BTC 9
2.3 Symmetrical BTC 12
2.4 BWER of the Asymmetrical BTC 13
2.5 Technology Simulation Results 16
2.6 Concluding Discussion 16
References 17

3 Multiple Frequency Impedance Matching by Frequency Transformation 19
3.1 Introduction 19
3.2 The Basic Network 20
3.3 Impedance Matching at a Single Frequency 21
3.4 Application of Frequency Transformation 21
3.5 Dual-Frequency Impedance Matching 23
3.6 Triple-Frequency Impedance Matching 24
3.7 Impedance Matching at Four Frequencies 25
3.8 Impedance Matching at Five Frequencies 26
3.9 Summary of the Procedure 27
3.10 Concluding Discussion 27
References 28

4 Triple Frequency Impedance Matching by Frequency Transformation 29
4.1 Introduction 29
4.2 General Theory 30
4.2.1 Basic L-Network for Impedance Matching 30
4.2.2 Single-Frequency Matching 31
4.2.3 Application of Frequency Transformation 31
4.3 Triple Frequency Matching Network 32
4.4 Alternative Circuits for Triple Frequency Matching 34
4.5 Conclusions 35
References 35

Part II Analog Signal Processing: Signal Processing by Distributed Networks

5 Application of Topological Formulas to Distributed Parameter Networks 39
5.1 Introduction 39
5.2 Topological Formulas 40
5.3 Formulas for Ladder Networks and Extension to Non-uniform Transmission Lines 40
References 44

6 On Equivalent Representation of Lossy Transmission Lines 45
6.1 Introduction 45
6.2 Derivation 46
6.2.1 Two-Port Equivalence 47
References 48

7 Efficient Approximation of Variable Delay Using Tapped Delay Line 49
7.1 Introduction 49
7.2 Derivation of Formulas for the Weighting Coefficients 50
7.2.1 Recursive Formula 50
7.2.2 Explicit Formula 51
7.3 Implementation 52
7.4 Performance 52
7.5 Conclusions 53
References 53

Part III Analog Signal Processing: Signal Processing by Active Networks

8 Second-Order Active RC Filters Using a Single Operational Amplifier 57
8.1 Introduction 57
8.2 Basic Circuit 58
8.3 Low-Pass Realization 59
8.4 Bandpass Filter Realizations 61
8.4.1 First Bandpass Realization 62
8.4.2 Second Bandpass Realization 62
8.5 High-Pass Realization 63
8.6 Band Stop Realization 65
8.6.1 Second Band Stop Realization 68
8.6.2 Third Band Stop Realization 69
8.7 Economic Aspect of Filter Realizations 70
8.8 Conclusion 71
References 72

9 Tolerance Minded Design of Active RC Bandpass Filters 73
9.1 Introduction 73
9.2 Butterworth Approximation 74
9.3 Conclusion 77
References 78

10 Active RC Filters Using a Single Differential Input Operational Amplifier 79
10.1 Introduction 79
10.2 The New Method 80
10.3 Two Examples 82
10.4 A Comparison of the Two Methods 83
10.5 An Illustration of the Flexibility of the New Method 84
References 86

11 Second-Order Active RC Bandpass Filters Using Operational Amplifier in Low-Gain Connection 87
11.1 Introduction 87
11.2 The Circuit Configuration 88
11.3 Sensitivity Considerations 89
11.4 Design Equations 90
11.4.1 Case A: Unity Gain Configuration 90
11.4.2 Case B: General K Configuration 91

11.5 Frequency Limitation 92
11.5.1 Case A: OA in Unity Gain Configuration 92
11.5.2 Case B: OA in Gain = 2 Configuration 93
11.6 Conclusion 95
References 96

12 Some Recent Contributions to Inductor Simulation and Applications to Low-Sensitivity Biquad Design 97
12.1 Introduction 97
12.2 Gyrator Based Methods 98
12.3 General Characteristics of Lossy, Grounded *L* Simulators 100
12.4 Simple Circuits for Lossy Grounded Inductance 101
12.4.1 Four-Component Circuits 101
12.4.2 Five-Component Circuit 103
12.4.3 Circuits Using Six Components 103
12.4.4 Comments 105
12.5 Lossy Floating Inductance Simulation 106
12.6 Applications of Simulated *L* in the Design of Insensitive Biquadratic Sections 108
12.7 Conclusion 110
References 111

13 Evaluating the Frequency Limitations of Operational Amplifier RC Networks by the Dominant-Pole Technique 113
13.1 Introduction 113
13.2 Dominant-Pole Technique 114
13.3 Frequency Limitations of Lossy Grounded Inductance Simulators 116
13.4 Experimental Investigations 123
13.5 Conclusion 124
References 124

14 'Shadow' Filters—A New Family of Electronically Tunable Filters 125
14.1 Introduction 125
14.2 Analysis 126
14.3 Special Cases 127
14.4 Range of Variation 128
14.5 An Open Problem and a Suggestion 128
14.6 An Example 129
14.7 Conclusion 129
References 130

Part IV Analog Signal Processing: Signal Processing by Discrete Time Networks

15 Some Aspects of Digital and CTD Signal Processing 133
15.1 Introduction 133
15.2 Our Early Work in Digital Signal Processing 133
15.3 Linear-Phase Recursive Filters 134
15.4 Discrete Hilbert Transform and Its Applications in Digital Filtering 135
15.5 Roundoff Noise in Digital Filters 136
15.6 Quantized Coefficent Design of Digital Filters 137
15.7 Variable IIR Digital Filters 138
15.8 Variable FIR Digital Filters 139
15.9 Nested Structures for FIR Filters 140
15.10 Fourier Transform at Arbitrary Frequencies 142
15.11 Inefficiency Compensation in CTD Delays and Signal Processors 142
15.12 CTD Linear Phase Versus Minimum Phase Design 145
15.13 Coefficient Sensitivity in CTD Filters 145
15.14 Spectrum Analysis by CTD 146
15.15 Looking Ahead 146
References 147

16 Signal Processing Applications of Charge-Coupled Devices 151
16.1 Introduction 151
16.2 Physics of Charge-Coupled Devices 153
16.3 The Input and Output Schemes 155
16.3.1 Input Circuit 155
16.3.2 Output Circuit 157
16.4 Characterization, Measurement and Compensation of Charge-Transfer Inefficiency 158
16.5 CCD Transversal Filters 162
16.5.1 Limitations of This Technique 165
16.5.2 CCD Low-Pass Filters 167
16.5.3 Bandpass Filters 167
16.5.4 Matched Filtering 168
16.5.5 Hilbert Transformer 169
16.5.6 Adaptive Filtering 169
16.5.7 Complex Coding 169
16.5.8 Interference Rejection Filter Using CCDs 170
16.5.9 Linear Regression Analysis Using CCDs 170
16.5.10 Radar Display System Using CCDs for Time Expansion 170

16.5.11 Waveform Generation Using CCDs 170
16.5.12 Cost and Performance Trade-Offs in Various Approaches to Signal Processing 171
16.6 Spectral Analysis Using CCDs 173
16.7 MTI Radar and Television Ghost Suppression Applications . . . 174
16.7.1 Simple Cancelling Filter 174
16.7.2 Application to Practical MTI Filters 175
16.7.3 Processor for Pulse-Doppler Radar 177
16.7.4 Charge-Transfer Memory for ECM 177
16.7.5 Television Ghost Suppression Using CCD 178
16.8 Concluding Remarks 179
References 180

17 Linear-Phase Versus Minimum-Phase CCD Transversal Filters 185
17.1 Introduction 185
17.2 Analysis 185
17.3 Results of Numerical Computations 188
17.4 Conclusion 190
Reference 190

Part V Digital Signal Processing

18 A State-of-the-Art Survey on Linear Phase Digital Filter Design 193
18.1 Introduction 193
18.2 The Problem of Linear Phase and FIR Design 195
18.3 Linear Phase IIR Design 197
18.4 Conclusion 199
References 199

19 Generalized Linear Phase Conditions for Causal FIR Systems Revisited 201
19.1 Introduction 201
19.1.1 Basic Concepts 201
19.1.2 Conditions for GLP and Proof of Sufficiency 202
19.1.3 Necessary and Sufficient Conditions 203
19.2 Method 1 204
19.3 Method 2 205
19.4 Method 3 206
19.5 Method 4 208
19.6 Method 5 210

19.7 Method 6 . . . 211
19.8 Conclusions . . . 212
References . . . 213

20 Review of Pseudo-Boolean Methods with Applications to Digital Filter Design . . . 215
20.1 Introduction . . . 215
20.2 Pseudo-Boolean Functions . . . 216
20.3 Linear PB Equations and Inequalities . . . 216
20.4 Nonlinear PB Equations and Inequalities . . . 219
20.5 Minimization of PB Functions . . . 221
20.6 Applications to Digital Filters . . . 224
20.7 Discussion . . . 227
References . . . 228

21 Maximally Linear FIR Digital Differentiators: A Review . . . 229
21.1 Introduction . . . 229
21.2 DDs for Low Frequencies: $H_1(\omega)$ . . . 231
21.2.1 Design for Approximation to $H_1(\omega)$ . . . 231
21.2.2 Performance of $H_1(\omega)$ . . . 234
21.3 DDs for Mid-band Frequencies with Zero Phase: $H_2(\omega)$ . . . 234
21.3.1 Design for Approximation to $H_2(\omega)$ . . . 235
21.3.2 Performance of $H_2(\omega)$ . . . 236
21.4 DDs for Mid-band Frequencies with $\pi/2$ Phase: $H_3(\omega)$ . . . 236
21.4.1 Design for Approximation to $H_3(\omega)$ . . . 239
21.4.2 Performance of $H_3(\omega)$ . . . 241
21.5 DDs for High Frequencies: $H_4(\omega)$ . . . 243
21.5.1 Design for Approximation to $H_4(\omega)$ . . . 243
21.5.2 Performance of $H_4(\omega)$ . . . 244
21.6 Variable Frequency DDs: Universal Differentiators . . . 244
21.6.1 Design of Universal Digital Differentiators . . . 246
21.6.2 Performance of Universal Differentiators . . . 248
21.7 Conclusions . . . 250
References . . . 250

22 Design of Maximally Flat and Monotonic FIR Filters Using the Bernstein Polynomial . . . 253
22.1 Introduction . . . 253
22.2 Review . . . 254
22.3 The Bernstein Polynomial and Design of Maxflat Filters . . . 259
22.4 Transformation Matrix . . . 260
22.4.1 Coefficients a_i's . . . 261
22.4.2 Coefficients b_i's . . . 261

22.5 Maxflat FIR Filters . . . 263
22.6 Monotonic FIR Filters . . . 265
22.6.1 Monotonic Filters with Transition Width 2/N . . . 266
22.6.2 Monotonic Filters with Transition Width 3/N . . . 268
22.6.3 The New Method . . . 269
22.6.4 Comparative Study . . . 271
22.7 Quadrature Mirror Filter . . . 273
22.7.1 Design of MAXFLAT QMF . . . 276
22.7.2 Monotonic QMF with Zero Reconstruction Error at the End and Mid Points . . . 278
22.7.3 Minimization of the Reconstruction Error at x_m . . . 281
22.8 Conclusion . . . 283
References . . . 284

23 On the Design of FIR Notch Filters . . . 287
23.1 Introduction . . . 287
23.2 Design . . . 287
23.3 Design Examples and Performance . . . 291
23.3.1 Equiripple Design . . . 292
23.4 Conclusion . . . 293
References . . . 293

24 A Simple Solution for the Analytic Inversion of Van der Monde and Confluent Van der Monde Matrices . . . 295
24.1 Introduction . . . 295
24.2 Inversion Formula for Van der Monde Matrix . . . 296
24.3 Inversion Formula for the Confluent Van der Monde Matrix . . . 298
References . . . 299

25 A New Canonic Lattice Realization of Arbitrary FIR Transfer Functions . . . 301
25.1 Introduction . . . 301
25.2 Derivation of the New Structure . . . 303
25.3 Cases Requiring Modification of the Procedure . . . 305
25.4 Realization of a Single Transfer Function . . . 306
25.5 Limitations of the New Procedures . . . 308
25.6 Concluding Comments . . . 309
References . . . 309

26 A Note on Canonic Lattice Realization of Arbitrary FIR Transfer Functions . . . 311
26.1 Introduction . . . 311
26.2 The New Procedure . . . 311
Reference . . . 313

27 A New Canonic Lattice Realization of Arbitrary IIR Transfer Functions 315
27.1 Introduction 315
27.2 Derivation of the New Structure 316
27.3 Examples 318
27.4 Conclusions 321
References 321

28 A New Feedback Configuration for Canonical Realization of IIR Filters and Its Application to Lattice Realizations 323
28.1 Introduction 323
28.2 The Method 325
28.3 Applications to Canonical IIR Lattice Realizations 328
28.4 Conclusions 330
References 330

Author and Contributors

About the Author

Suhash Chandra Dutta Roy was fortunate enough to be educated at the Calcutta University, where great professors and researchers like C V Raman, S N Bose, M N Saha and S Radhakrishnan taught and made a name, globally, not only for them, but also for the university and the country. Immediately after my master's, I had to take up a job in a Government Research Institute. I soon discovered that many things are done at the Institute, but not research. I therefore quit and shifted to the newly established University of Kalyani as a lecturer. There I spent a few years, and after my Ph.D., I left the country to take up an Assistant Professorship at the University of Minnesota. A few years passed by, but I felt increasingly guilty that I was not doing anything for my motherland. I therefore returned to the country and joined IIT Delhi, as an Associate Professor where I served for more than four decades as Professor, Head of the Department and as a Dean. The last two positions were imposed upon me, and I suffered because they cut down on the time available for my research and interactions with my dear students. I formally retired at the age of 60, but continued to teach and carry out research, as an Emeritus Fellow, followed by INSA Senior Scientist and INSA Honorary Scientist. Later, when my term as an Emeritus Fellow was over, I did not take any money from IIT Delhi and served voluntarily, simply for the love of teaching and research. I finally left IIT Delhi

permanently, 12 years after the retirement. I am now settled at Hauz Khas, New Delhi, where I live happily with my wife, but still continue to do research. We enjoy good health because of Yoga and strict diet.

I have been extraordinarily lucky to have had gems of Ph.D. students, 30 of them, who have done exemplary work in research and innovation. Some of them have figured in 'Contributors'. In fact, I feel myself shining with reflected glory.

I have also been lucky to have received high recognition through Fellowship of IEEE, Distinguished Fellowship of IETE and Fellowship of all the relevant national academies. I have been awarded some prestigious national awards, including the Shanti Swarup Bhatnagar Prize. Over and above all the awards and recognitions, however, what I value most is the love and affection of my students.

I now spend time giving professional lectures on innovations in teaching and research, in general, and also on how to improve their standards in the country. I have the hobby of listening to Hindustani classical music and researching on its maestros. I also love spending quality time with my grandson, Soham. Reading political history and detective stories are my other hobbies. I also read poetry and compose some poems and short stories, for my own pleasure.

–In short, I have lived a complete life, with nothing to complain about.

Contributors[1]

Prof. Amalendu Bhushan Bhattacharyya did his M.Sc. (Physics) and Ph.D. from BHU, and was the Head of the Centre for Applied Research in Electronics (CARE) at IIT Delhi. I collaborated with him in the Analog Signal Processing Systems (ASPS) project of the Department of Electronics, Government of India. In this project, fabrication and applications of the charge-transfer device (CTD)—both bucket-brigade device (BBD) and charge-coupled device (CCD)—formed a major

[1]I shall now introduce my co-authors in some of the papers on which the book is based. The first four were my colleagues in the Institute. The next eleven were my Ph.D. students, listed in alphabetical order. The last person listed here was my M.Tech. student.

part. After retirement, he joined ITM Gurgaon, as Director. He then transferred to Goa University as Professor and Director, VLSI Design program, and finally to the JP University, Noida, as Emeritus Professor. He lives in Noida.

Dr. Alladi Prabhakar was a colleague of mine in the Electrical Engineering Department (EED) of IIT Delhi. An IISc Ph.D. holder specializing in Graph Theory, he later transferred to ITI Bangalore, became its Managing Director, and is now settled at Bangalore after retirement.

Prof. Surendra Prasad did his B.Tech. at IIT Kharagpur and M.Tech. and Ph.D. from IIT Delhi. After M.Tech., he joined CARE as a lecturer, and after Ph.D., he became an assistant professor in the EE Department, where he later became a professor. He also became a dean, Deputy Director (Faculty) and finally retired as Director. He is now Emeritus Professor at IITD.

Prof. Prem Saran Satsangi was a dear colleague of mine at IITD. He did his bachelor's degree from BHU, master's from Michigan State University, East Lansing, USA, and Ph.D. from Waterloo at Canada. On return from Waterloo, he became an associate professor, and later a full professor. He occupied several important positions at IIT Delhi, including Head of the Department, Dean and Managing Director, FITT. Later, he took voluntary retirement to take up Directorship of the Dayalbagh Educational Institute (DEI), Agra, a deemed university. After retirement, he was elected to head the Satsang Community and is now the most respected Guruji Maharaj there.

Prof. Juzer Vasi is a B.Tech. graduate from IIT Bombay and a Ph.D. holder from Harvard. He joined CARE at IIT Delhi as an assistant professor and transferred to IIT Bombay for family reasons. He established a state-of-the-art microelectronics facility at IITB. Now retired, he is settled at Mumbai.

Monica Agrawal was jointly registered with me and Surendra Prasad for her Ph.D. After graduation, she worked with Hughes Software for a few years and then returned to IITD as an assistant professor in CARE. She is now a full professor there, doing important work in Underwater Signal Processing.

Muslim Taj Ahmed was a B.Sc. (Engg) graduate from AMU, who also served as a lecturer there. Later, he came to IIT Delhi for his M.Tech., and seeing his extraordinary performance in course work, he was transferred to Ph.D. Program with me as the guide. After Ph.D., he went back to AMU, became an assistant professor and later a professor there. After a few years, he went to Dubai and Saudi Arabia with teaching assignments. After returning to AMU, he served the Zakir Hussain College of Engineering, as its Principal. Unfortunately, he died prematurely.

Vuppuluri Gurusharan Das a B.Sc. (Engg) graduate from Dayalbagh Engineering College (DEI, then under Agra University) and an M.Tech. graduate from IIT Kanpur, came as a QIP candidate to IITD to work for his Ph.D. under my supervision. He was exceptionally brilliant and completed his research in record time. He worked on the CCD as part of the ASPS project. After completion, he went back to DEI, became a full professor there and finally DEI's Director. After retirement, he rejoined DEI as Professor Emeritus and remained active in teaching and project supervision there. He passed away recently in May 2018.

Shail Bala Jain did B.Sc. (Engg) from the Delhi College of Engineering (DCE) and M.Tech. from IITD. She was a faculty in DCE for many years. During this, she registered for Ph.D. at Delhi University with me as her guide and Prof. A. K. Sinha (DCE) and Balbir Kumar (DIT, now NSIT) as co-guides. She became a full professor at IGIT, GGSIP University, and later joined the Indira Gandhi Delhi Technical University for Women as an academic advisor. She is now leading a retired life.

Navin Kapur did his B.Tech. at IIT Delhi and then took up a job in CDIL. After a few years, he came to me for Ph.D. I got him a research associateship in the ASPS project to work on CCDs. He did good work, got his Ph.D. and went back to CDIL. After retirement from there, he worked in a private institution as Professor and later became its Director. On superannuation from there, he settled at Delhi. He also passed away recently.

Balbir Kumar did his B.E. from Punjab Engineering College (now a deemed university) and was commissioned in the Indian Navy. He came as a sponsored candidate from the Navy to do his M.Tech. at IITD. He worked for his M.Tech. thesis with me and then rejoined the Navy. After a few years, he took voluntary retirement and joined IITD for his Ph.D. He did excellent work on digital differentiators. After Ph.D., he joined the newly established Delhi Institute of Technology (now Netaji Subhas Institute of Technology [NSIT]) as an assistant professor and was promoted to full professorship after a few years. At NSIT, he also worked as an acting director for some time. After retirement, he worked in a private institution for some years. Unfortunately, like Ahmed and Das, he also died prematurely.

Shailey Minocha registered for her Ph.D. with me when she was working on a project in CARE. She only had a bachelor's degree from DCE, but with some course work, she did excellent research. After her Ph.D., she took up a job with TCS, which was followed by a postdoctoral fellowship at Technical University, Braunschweig, Germany. On completion of her fellowship, she moved to the UK and is now a full professor at UK's Open University.

Rakesh Kumar Patney came to do his M.Tech. at IIT Delhi, after a B.Tech. from IIT Kanpur. He did his M.Tech. thesis with me and continued for his Ph.D. After graduation, he was inducted to the faculty. He recently retired as full Professor and is continuing at IIT Delhi as Emeritus Fellow.

L. R. Rajagopal is a graduate of the National Defence Academy, Khadakwasla, Pune, and the Naval College of Engineering at Lonavala, Pune, and was sponsored by the Indian Navy to do his M.Tech. at IITD. He did his M.Tech. thesis with me and Rakesh. Then, he continued at IIT Delhi for his Ph.D. under my supervision. The use of Bernstein polynomials in DSP was his innovation, and the work is widely cited in the literature, even today. After Ph.D., he went back to the Navy and did commendable work in the WESEE project. After superannuation, he started his own industry at Chennai, named 'Signals and Systems', which is now well established and reputed for their high-quality power instruments.

L. Shankaranarayanan came to work in the ASPS project after his M.Sc. (Physics) from Anna University. Shankar registered for his Ph.D. with Prof. Bhattacharyya and me. He later went to work with Philips and rose to the position of Research Director. I do not know his whereabouts at the present time.

Mahesh Chandra Srivastava came to do his Ph.D. under the supervision of Prof. Satsangi and me. He was sponsored by the MMM Engineering College, Gorakhpur. After his Ph.D., he joined BIT Mesra, Ranchi. After a few years, he changed to KNIT Sultanpur and then served several institutions till he came to Jaypee Institute of Technology, Noida. That is what I last knew about him. Recently, I came to know that Mahesh is no more.

Phani Bhushan Tholetti was an M.Tech. student. The paper in which he collaborated and contributed to the work in this book was done by him while taking my course on DSP. He is now working in industry at Bangalore.

Part I
Analog Signal Processing: Signal Processing by Passive Networks

This part contains 4 chapters. Chapter 1 deals with finite zero filters of the Butterworth and Chebyshev types. Chapter 2 contains one of my best works. I show that the bandwidth enhancement capability of the asymmetrical bridged T-coil, used as the inter-stage network, is virtually unlimited. This is something new and may open up new investigations, because it offers a technological challenge. Chapters 3 and 4 deal with multiple frequency impedance matching by using the frequency transformation techniques.

Chapter 1
Finite Zero Butterworth and Chebyshev Filters

Abstract An accurate expression is derived for the minimum order of a finite-zero Butterworth filter (Dutta Roy in IEEE Trans AU-19(1):58–63, 1971) [1], which has a steeper cut-off slope than the corresponding Chebyshev filter (Agarwal and Sedra in IEEE Trans AU-20(2):138–141, 1972) [2]. The inaccuracy occurring in the latter is illustrated by a numerical example.

Keywords Finite-zero filters · Butterworth filters · Chebyshev filters · Cutoff slope

This chapter is concerned with a comparison of the finite-zero, low-pass, Butterworth filter, defined by the magnitude squared function [1]:

$$|F_{mn}(j\omega)|^2 = \frac{1}{1+\omega^{2n}[(\omega_0^2-1)/(\omega^2-\omega_0^2)]^{2m}} \tag{1.1}$$

with the corresponding Chebyshev filter, defined similarly by [2]:

$$|T_{mn}(j\omega)|^2 = \frac{1}{1+\varepsilon^2 C_n^2(\omega)[(\omega_0^2-1)/(\omega^2-\omega_0^2)]^{2m}} \tag{1.2}$$

where, in either case, n denotes the order of the filter, m is the order of zeros at $\pm j\omega_0$, $\omega_0 > 1$ and $n \geq 2\,m+1$. In Eq. (1.2), $C_n(\omega)$ denotes the Chebyshev polynomial of the first kind, of order n, and ε^2 is a ripple controlling factor. We define the cut-off slope in either case as the negative differential coefficient of the magnitude with respect to ω, evaluated at $\omega = 1$, and denote it by S_{bmn} in the Butterworth and S_{tmn} in the Chebyshev case. Further, for the purpose of comparing the two cut-off slopes, we set $\varepsilon^2 = 1$; then it can be shown that [1, 2].

$$S_{bmn} = [n/(2\sqrt{2})]\{1 + 2m/[n(\omega_0^2-1)]\} \tag{1.3}$$

and

Source: Suhash C. Dutta Roy, "Comparison of Finite Zero Butterworth and Chebyshev Filters," *JIETE*, vol 21, p 350, June 1975.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_1

$$S_{tmn} = [n^2/(2\sqrt{2})]\{1 + 2m/[n^2(\omega_0^2 - 1)]\} \tag{1.4}$$

Let n_B denote the *minimum* order of the modified Butterworth function in Eq. (1.1) with m_B pairs of zeros at $\pm j\omega_0$ for which the cut off slope is steeper than that given by Eq. (1.4). Then from Eqs. (1.3) and (1.4), it follows that

$$n_B + \{2m_B/(\omega_0^2 - 1)\} \geq n^2 + \{2m/(\omega_0^2 - 1)\} \tag{1.5}$$

Note that if $m_B = m$, then n_B is simply equal to n^2. To allow the Butterworth function to have the highest possible slope, we set

$$m_B = (n_B - r)/2 \tag{1.6}$$

where, since m_B is to be an integer,

$$r = \begin{cases} 1, & \text{for odd } n_B \\ 2, & \text{for even } n_B \end{cases} \tag{1.7a, b}$$

Substituting Eq. (1.6) into Eq. (1.5), we get

$$n_B \geq (1/\omega_0^2)[2m + r + n^2(\omega_0^2 - 1)] \tag{1.8}$$

which, combined with Eq. (1.7a, b), gives

$$n_B = \text{Lower of} \begin{cases} \text{odd} & < (1/\omega_0^2)[2m + 1 + n^2(\omega_0^2 - 1)] > \\ \text{even} & < (1/\omega_0^2)[2m + 2 + n^2(\omega_0^2 - 1)] > \end{cases} \tag{1.9}$$

where the notation <a> means the lowest integer greater than or equal to a.

Agarwal et al. [2] have assumed Eq. (1.7a), without regard to the possibility of an even n_B, and the resulting equation can be written in our notations as

$$n_B = \ <(1/\omega_0^2)[2m + 1 + n^2(\omega_0^2 - 1)] > \tag{1.10}$$

As an example of the inaccuracy of Eq. (1.10), let $n = 5$, $m = 2$ and $\omega_0 = 2.0$; then Eq. (1.10) gives $n_B = 20$ and according to Eqs. (1.6) and (1.7a), $m_B = 9.5$. Since m_B has to be an integer and $m_B = 10$ is not allowed, the highest possible value of m_B is 9. Under these conditions, the left-hand side of inequality (1.5) becomes 26 while the right-hand side becomes 26.33.

In the above situation, Eq. (1.9), on the other hand, gives n_B = lower of {odd <20>; even <20.25>} = 21 and from Eqs. (1.6) and (1.7a), one obtains, $m_B = 10$. With these values, the left-hand side of inequality (1.5) is 27.67 which is greater than 26.33, the value of the right-hand side.

As another example of the application of Eq. (1.9), let $n = 5$ and $m = 2$, as in the previous case, but $\omega_0 = \sqrt{6}$. Then Eq. (1.9) gives n_B= lower of {odd <21.67>; even <21.83>} = 22 and hence $m_B = 10$. It may now be verified that these satisfy the inequality (1.5).

References

1. S.C. Dutta Roy, On maximally flat sharp cut-off low-pass filters. IEEE Trans. AU-**19**(1), 58–63 (1971)
2. M.C. Agarwal, A.S. Sedra, On designing sharp cutoff low-pass filters. IEEE Trans. AU-**20**(2), 138–141 (1972)

Chapter 2
Bandwidth Enhancement with the Asymmetrical Bridged T-coil Network

Abstract This chapter presents the results of a theoretical investigation of the asymmetrical bridged T-coil (BTC) bandwidth enhancement network under the condition of maximally flat magnitude (MFM) response. Conventionally, the BTC is used in the symmetrical configuration and it is well known that, as compared to the commonly used RC load, the maximum bandwidth enhancement ratio (BWER) under the MFM condition is $2\sqrt{2}$. It is shown here that by introducing asymmetry, even under the constraint of MFM response, the maximum achievable BWER is theoretically unlimited, the limit being set only by practical considerations. Unlimited BWER has never been considered possible, even theoretically, with BTC or any other network, and the result reported here is believed to be a breakthrough in the design of wide-band and ultra wide-band amplifiers. Simulations for a few designs indicate that practical realization of the full potential of the asymmetrical BTC may have to wait for improved technologies which facilitate tight coupling between the two parts of the coil and also minimize parasitic effects.

Keywords Bandwidth enhancement · Bridged T-coil · Maximally flat magnitude · Wide-band amplifiers

2.1 Introduction

Considerable interest has been shown in the analysis, design, and applications of the bridged T-coil (BTC) network during the past three decades [1–30]. While major application of the BTC is in wide-band amplifiers [1–25], several recent papers relate to its application in replacing transmission line sections required in ultra wide-band power dividers and bandpass filters, thereby significantly reducing the circuit size without affecting the performance [26–28]. This equivalence was first proposed in [29] for modelling low-temperature co-fired ceramic embedded inductors. Yet another very recent application of the BTC has been reported in [30] in the design of all-pass networks and phase-shifters.

Source: Suhash C. Dutta Roy, "A Theoretical Investigation into the limits of Bandwidth Enhancement with the Asymmetrical Bridged T-Coil Network," *IETE Journal of Research*, vol 62, pp 379–386, May–June 2016.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_2

This chapter will focus on the application of the BTC in wide-band amplifier design only, a typical stage of which is shown in Fig. 2.1, where the active device is an NMOS transistor. As mentioned in [13] and [16], the BTC has been used for about four decades by Tektronix Inc. in oscilloscope circuitry, but its design remained a trade secret for a long time. A US patent on a bridged T-termination network, which can be considered as a lossy BTC was granted as early as 1964 [31], but the first public disclosure of analysis and design of the BTC appeared in 1982 [1] (not 2004 [13], as mentioned in [16]). Most of the contributions reported so far used the symmetrical form of the BTC, and it has been well established that under the maximally flat magnitude (MFM) condition, the maximum bandwidth enhancement ratio (BWER), as compared to the commonly used RC load, is $2\sqrt{2}$. Asymmetrical BTC was considered in [6, 7, 9] and [14]. In [6] as well as [9], the maximum BWER achieved was the same as that of the symmetrical BTC under the MFM condition. It was mentioned in [7] that the BWER considerably exceeded the conventional limit but no theoretical analysis or the condition (MFM or otherwise) under which the same was achieved was given. There are indications in [14] of achieving, with an asymmetrical BTC, a BWER greater than 4 with zero-dB ripple; however, zero-dB ripple does not necessarily mean the MFM condition, and no theoretical analysis was provided. If ripples are allowed, [14] shows that the BWER can even exceed 5. It is worth mentioning here that in [32], published as early as 1954, the circuits labelled D are asymmetrical BTCs; again, there was no theoretical analysis, but the plots of gain frequency characteristics given in Fig. 2.2 of that paper indicated a BWER much in excess of $2\sqrt{2}$.

A comprehensive theoretical investigation of the general asymmetrical BTC is carried out in this chapter, and it is shown that even under the MFM constraint, the BWER achievable by this network is theoretically unlimited, the realizable limit being set only by practical considerations. This result is believed to be a breakthrough in the design of wide-band and ultra wide-band amplifiers.

Some of the theoretical designs were simulated using models of 130 nm technology, which show that parasitic effects severely limit the bandwidth enhancement, and lead to the conclusion that improved technologies, which allow tight coupling

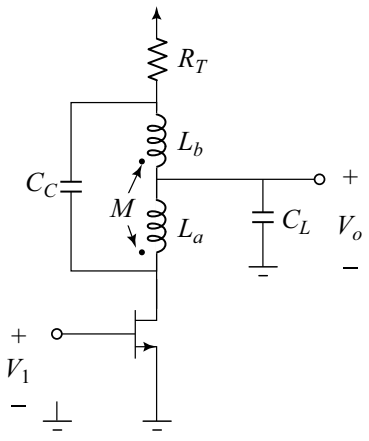


Fig. 2.1 A common source NMOS amplifier stage with a bridged T-coil network in the drain circuit

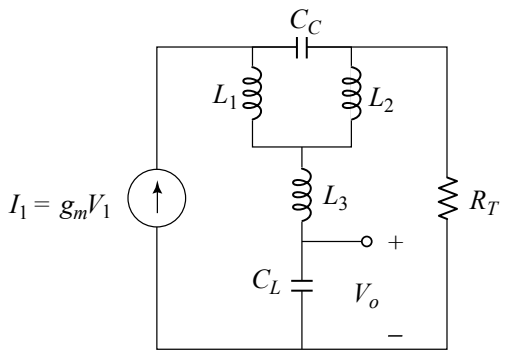


Fig. 2.2 Small-signal equivalent circuit of Fig. 2.1, where $L_1 = L_a + M$, $L_2 = L_b + M$, and $L_3 = -M$

between the two parts of the asymmetrical BTC, and also minimize the parasitic effects, will be needed to realize the full potentials of the BTC.

The chapter is organized as follows. In the next section, we give the equivalent circuit of Fig. 2.1 and the expression for its transfer impedance, which is of order 4. This transfer function is then reduced to one of order 2 by following [16] and the conditions under which this can be done are stated. Section 2.3 deals with the application of these results to a symmetrical BTC, where a misconception in the existing literature [13] is cleared. Section 2.4 contains the details of derivation needed to find the BWER of the asymmetrical BTC under the MFM constraint, and establishes the fact that in theory, the BWER can be infinitely large. A plot of BWER versus the coefficient of coupling is given and some specific designs are given in the form of a table. Section 2.5 comments on the results of simulation of a few designs while Sect. 2.6 contains the concluding discussions.

2.2 Analysis of the General BTC

The small-signal equivalent circuit of Fig. 2.1 is shown in Fig. 2.2, where g_m is the transconductance of the NMOS transistor. In [16], this circuit was analysed by using the extra-element theorem [33] whereas in [17], a much simpler analysis was carried out by using classical techniques.

In either case, the result obtained is the following:

$$\begin{aligned} Z_T &= V_0/I_i = R_T N(s)/D(s) \\ N(s) &= s^2 C_c(L_a + L_b + 2M) + s((L_b + M)/R_T) + 1 \\ D(s) &= s^4 C_c C_L(L_a L_b - M^2) + s^3 C_c C_L R_T(L_a + L_b + 2M) \\ &\quad + s^2(C_c(L_a + L_b + 2M) + C_L L_b) + s C_L R_T + 1. \end{aligned} \tag{2.1}$$

The transfer impedance Z_T is of the fourth order, and it is difficult to proceed further analytically. Following [16], we, therefore, convert it to a second-order one by forcing pole-zero cancellation, i.e. we write

$$D(s) = (ps^2 + qs + 1)N(s) \tag{2.2}$$

and determine the constraints on the element values by equating the coefficients of powers of s on both sides. Carrying out these steps, we get [16]

$$p = C_L(L_aL_b - M^2)/(L_a + L_b + 2M), \tag{2.3}$$

$$q = C_L R_T - \frac{C_L(L_aL_b - M^2)(L_b + M)}{C_c R_T(L_a + L_b + 2M)^2}, \tag{2.4}$$

$$\frac{q(L_b + M)}{R_T} + p = C_L L_b, \tag{2.5}$$

and

$$q = C_L R_T - \frac{L_b + M}{R_T}. \tag{2.6}$$

Comparing Eqs. (2.4) and (2.6) gives

$$\frac{C_c}{C_L} = \frac{L_aL_b - M^2}{(L_a + L_b + 2M)^2}. \tag{2.7}$$

Also, by substituting Eqs. (2.3) and (2.6) in Eq. (2.5) and carrying out some algebraic simplifications, we obtain

$$C_L R_T^2 = \frac{(L_b + M)(L_a + L_b + 2M)}{L_a + M}. \tag{2.8}$$

Let $L_a = L$, $L_b = bL$, and $M = mL$; then the coefficient of coupling between the two inductors in Fig. 2.1 is given by

$$k = \frac{M}{(L_aL_b)^{1/2}} = \frac{m}{\sqrt{b}} \tag{2.9}$$

Also, Eq. (2.8) gives

$$L = \frac{C_L R_T^2(1 + m)}{(b + m)(1 + b + 2m)} \tag{2.10}$$

and Eq. (2.7) gives

$$\frac{C_c}{C_L} = \frac{b - m^2}{(1 + b + 2m)^2}. \tag{2.11}$$

When L and C_c obey Eqs. (2.10) and (2.11), respectively, Z_T/R_T becomes a second-order function of the form

$$\frac{Z_T}{R_T} = \frac{1}{ps^2 + qs + 1}, \tag{2.12}$$

where

$$p = \frac{LC_L(b - m^2)}{1 + b + 2m} \text{and } q = C_L R_T - \frac{(b + m)L}{R_T}. \tag{2.13}$$

Equation (2.12) can be recast in the standard form

$$Z_T/R_T = \omega_0^2/(s^2 + 2\varsigma\omega_0 s + \omega_0^2). \tag{2.14}$$

Comparing Eqs. (2.12) and (2.14) and using Eqs. (2.10), (2.11), and (2.13), we get, after simplification,

$$\omega_0 = \frac{1}{\sqrt{p}} = \frac{1 + b + 2m}{C_L R_T}\sqrt{\frac{b + m}{(1 + m)(b - m^2)}} \tag{2.15}$$

and

$$2\varsigma = \frac{q}{p\omega_0} = \sqrt{\frac{(b + m)^3}{(1 + m)(b - m^2)}}. \tag{2.16}$$

As is well known [34], the transfer function given by Eq. (2.14) has MFM response when $\varsigma = 1/\sqrt{2}$ (critically damped case). For $0 < \varsigma < 1/\sqrt{2}$ (underdamped case), the magnitude response exhibits a peak of value $1/[2\varsigma(1 - \varsigma^2)^{1/2}]$ at the frequency $\omega_0(1 - 2\varsigma^2)^{1/2}$. When $\varsigma = 1/\sqrt{2}$ (overdamped case), the magnitude response decays monotonically. Typical plots of the magnitude of Eq. (2.14) for underdamped, critically damped and overdamped cases are shown in Fig. 2.3.

The 3-dB bandwidth of Eq. (2.14) is given by

$$\omega_3 = \omega_0\sqrt{(1 - 2\varsigma^2) + \sqrt{(1 - 2\varsigma^2)^2 + 1}} \tag{2.17}$$

The 3-dB bandwidth of the amplifier with the commonly employed $R_T C_L$ load is given by $1/(R_T C_L)$. Hence, the BWER with the BTC load is given by

$$\eta = \omega_3 R_T C_L \tag{2.18}$$

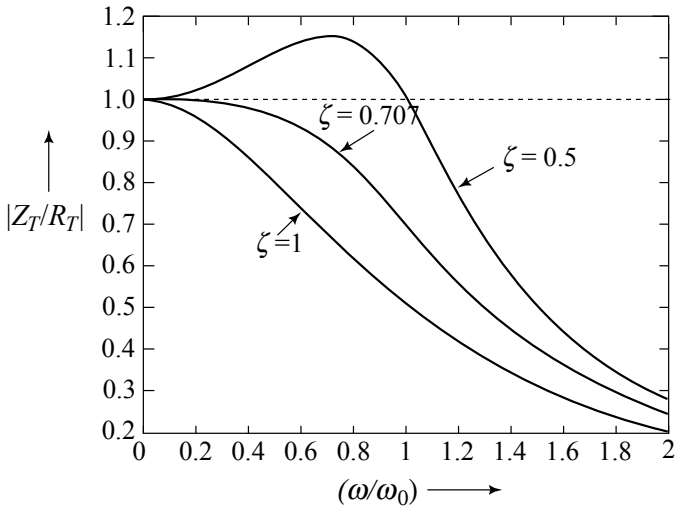


Fig. 2.3 Frequency response of (2.14) for $\varsigma = 0.5$ (underdamped), 0.707 (critically damped), and 1 (overdamped)

2.3 Symmetrical BTC

For the symmetrical BTC, $b = 1$, and Eqs. (2.15) and (2.16) simplify to the following:

$$\omega_3 = \frac{2}{C_L R_T}\sqrt{\frac{1+m}{1-m}} \tag{2.19}$$

and

$$2\varsigma = \sqrt{\frac{1+m}{1-m}} \tag{2.20}$$

With $b = 1$, Eq. (2.9) gives $k = m$, and from Eq. (2.20), we get

$$k = \frac{4\varsigma^2 - 1}{4\varsigma^2 + 1} \tag{2.21}$$

Putting $b = 1$ in Eqs. (2.10) and (2.11), and using Eq. (2.21), we get

$$L = \frac{C_L R_T^2}{4}\left(1 + \frac{1}{4\varsigma^2}\right) \text{ and } C_c = C_L/(16\varsigma^2). \tag{2.22}$$

These results agree with those derived by many authors earlier.

Now combine Eqs. (2.18) with (2.17) and then use Eqs. (2.19)–(2.21), along with the fact that $k = m$. The result is

$$\eta = 4\varsigma\sqrt{(1-2\varsigma^2) + \sqrt{(1-2\varsigma^2)^2 + 1}} \tag{2.23}$$

Equation (2.17) shows that as ς increases from 0 to 1, ω_3 decreases monotonically from $(\sqrt{2}+1)^{1/2}\omega_0$ to $(\sqrt{2}-1)^{1/2}\omega_0$, passing through the value ω_0 at $\varsigma = 1/\sqrt{2}$.

In Eq. (2.23), however, because of the factor ς, the BWER will show a maximum at some value of ς, and it is not difficult to show that $\varsigma = 1/\sqrt{2}$ is this value. Thus, the condition for MFM response is also the condition for maximum BWER. Except for [25], this fact has not been explicitly mentioned in the literature. The maximum η is calculated from Eq. (2.23) as $2\sqrt{2}$.

Lee [13] mentions that a bandwidth greater than ω_0 and a value of η almost equal to 3 are possible to be achieved by using $\varsigma = 0.7507$. However, the analysis given here does not corroborate this assertion. In fact, for $\varsigma = 0.7507$, Eq. (2.23) gives $\eta = 2.8184$, less than $2\sqrt{2}$, as expected, because η is a maximum when $\varsigma = 1/\sqrt{2}$.

2.4 BWER of the Asymmetrical BTC

For a general transfer function of the form of Eq. (2.14), combining Eqs. (2.17) and (2.18) with Eq. (2.15), we get the following expression for the BWER of the asymmetrical BTC:

$$\eta = (1 + b + 2m)\sqrt{\frac{b+m}{(1+m)(b-m^2)}} \times \sqrt{(1-2\varsigma^2) + \sqrt{(1-2\varsigma^2)^2 + 1}} \tag{2.24}$$

For MFM, $\varsigma = 1/\sqrt{2}$ and Eq. (2.24) becomes

$$\eta = (1 + b + 2m)\left(\frac{b+m}{(1+m)(b-m^2)}\right)^{\frac{1}{2}}. \tag{2.25}$$

Further, putting $\varsigma = 1/\sqrt{2}$, we get

$$2 = \frac{(b+m)^3}{(1+m)(b+m^2)}. \tag{2.26}$$

Thus for MFM response, b and m have to satisfy Eq. (2.26) and for any such set, η is given by Eq. (2.25). Combining Eqs. (2.25) with (2.26), the expression for η gets further simplified to the following:

$$\eta = \sqrt{2}\left(1 + \frac{1+m}{b+m}\right) \tag{2.27}$$

It is difficult to find, analytically, the variation of b with m from Eq. (2.26) or that of η with m from Eq. (2.27). One way is to simplify Eq. (2.26) to get a cubic equation in b and to solve it; however, as is well known [35, 36], an explicit expression for b in terms of m cannot be written and one has to take recourse to numerical computations. An alternative, and a simpler way, is to introduce the variable $\beta = b/m$; which converts

Eq. (2.26) to the following quadratic equation in m:

$$((\beta+1)^3+2)m^2+2(1-\beta)m-2\beta=0 \tag{2.28}$$

The solution of Eq. (2.28) is

$$m=\frac{(\beta-1)\pm\sqrt{(\beta-1)^2+2\beta((\beta+1)^3+2)}}{(\beta+1)^3+2} \tag{2.29}$$

The sign in Eq. (2.29) is dictated by the fact that b has to be positive. Thus if β is taken as positive, then m has to be positive; hence only the positive sign in Eq. (2.29) is acceptable.

A discussion on the possible range of values of m and b is in order at this point. Figures 2.1 and 2.2 show that m has been assumed to be positive, i.e. the coupling between L_a and L_b is negative. Can m be negative, i.e. is it possible to achieve a BWER > 1 with positive coupling? This question has been answered in the negative in the past, see e.g. [14] and [21].

Note that $m = 0$ means that there is no coupling between the two coils and the advantages of using the BTC are lost. Also, from Eq. (2.26), m is calculated as 1/3 when $b = 1$, which corresponds to the symmetrical case. Hence the range of m to be considered is $0 < m < 1/3$. The range of b is of course $0 < b < 1$. What about $b > 1$? Computation shows that $b > 1$ leads to BWER $< 2\sqrt{2}$. These points need to be kept in mind in interpreting Eqs. (2.28) and (2.29).

We computed the values of m for about two decades of values of β, and also the corresponding values of b, the coupling coefficient $k = m/\sqrt{b}$, and the BWER η. The results are given in Table 2.1 along with the design values of the normalized inductance $L/(C_L R_T^2)$ calculated from Eq. (2.10) and the capacitance ratio C_C/C_L calculated from Eq. (2.11). The variation of η with k is shown in Fig. 2.4.

Table 2.1 Some specific designs for the asymmetrical BTC

β	m	β	$k = m/\sqrt{b}$	η	$L/(C_L R_T^2)$	C_c/C_L
0.02	0.198×10^{-1}	0.396×10^{-3}	0.995	72.825	48.544	0.366×10^{-5}
0.04	0.392×10^{-1}	0.157×10^{-2}	0.990	37.463	23.585	0.269×10^{-4}
0.06	0.581×10^{-1}	0.349×10^{-2}	0.984	25.710	15.367	0.880×10^{-4}
0.10	0.946×10^{-1}	0.946×10^{-2}	0.973	16.296	8.776	0.355×10^{-3}
0.20	0.177	0.354×10^{-1}	0.941	9.251	3.988	0.211×10^{-2}
0.40	0.303	0.121	0.871	5.758	1.778	0.985×10^{-2}
0.60	0.383	0.230	0.799	4.606	1.131	0.208×10^{-1}
1.00	0.447	0.447	0.669	3.702	0.691	0.451×10^{-1}
2.00	0.408	0.815	0.451	3.042	0.438	0.937×10^{-1}
3.00	0.333	1.000	0.333	2.828	0.375	0.125

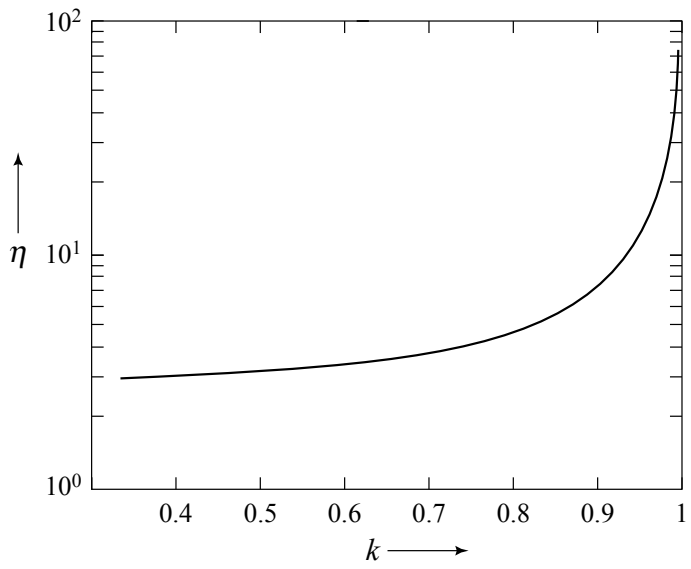


Fig. 2.4 Plot of BWER (η) versus coefficient of coupling (k)

The results for η in Table 2.1 and Fig. 2.4 are truly amazing, because they show that the BWER achievable by the asymmetrical BTC is theoretically unlimited, the realizable limit being set only by practical limitations of tight coupling and large inductance and capacitance spreads needed for large η. Parasitics will also play a major negative role, as discussed under Sect. 2.5.

The capability of unlimited BWER can also be justified from the expression (2.27), which can be rewritten in terms of β and m as follows:

$$\eta = \sqrt{2}\left(1 + \frac{1 + (1/m)}{1 + \beta}\right) \tag{2.30}$$

This clearly shows that $\eta \to \infty$ when $m \to 0$.

The values of the intermediate variable β are retained in Table 2.1 so as to facilitate the design for any value of η which lies in between two successive entries. Note that the last row in the table corresponds to the symmetrical BTC.

In order to display the possibility of achieving large BWERs with the asymmetrical BTC, we have calculated, for typical values of R_T and C_L, the element values corresponding to the fifth row in Table 2.1, along with those for the commonly used $R_T C_L$ load and the case of symmetrical BTC corresponding to the last row in Table 2.1. These are given below.

Design A with simple $R_T C_L$ load:
$R_T = 50\ \Omega$ and $C_L = 5$ pF.
Design B with symmetrical BTC ($b = 1$):
$R_T = 50\ \Omega$, $C_L = 5$ pF, $L_a = L_b = 4.688$ nH, and $C_C = 0.625$ pF.
Design C with asymmetrical BTC ($b = 0.0354$):
$R_T = 50\ \Omega$, $C_L = 5$ pF, $L_a = 49.85$ nH, $L_b = 1.765$ nH, and $C_c = 0.01055$ pf.

The corresponding expected values of the -3 dB frequencies are 0.6366, 1.8002, and 5.889 GHz, respectively. These designs were simulated with P Spice and the

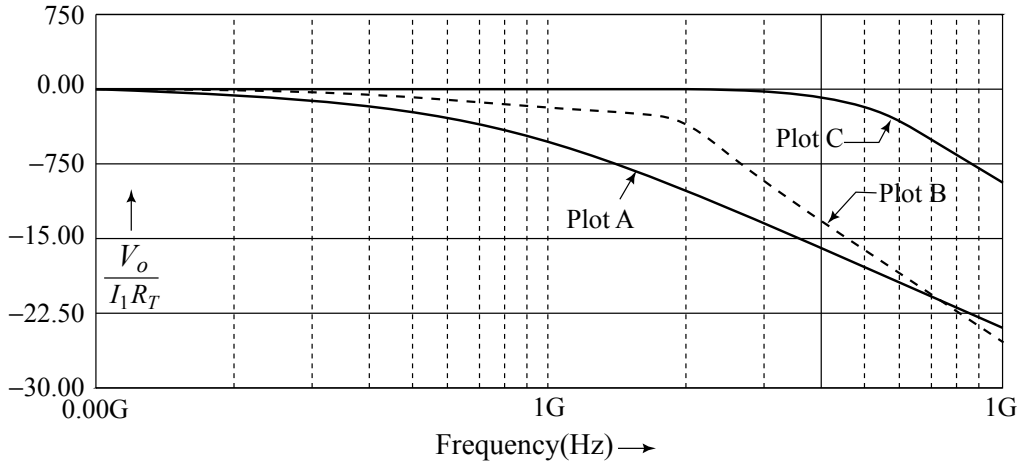


Fig. 2.5 Frequency response for designs A, B, and C obtained by P Spice simulation

resulting magnitude responses are shown in Fig. 2.5. These results agree with the expected ones, thus validating the theoretical investigations reported in this chapter.

2.5 Technology Simulation Results

Using Agilent's Advanced Design System tools with models of United Microelectronics 130 nm (1P8M) technology, simulation studies were carried out for a few specific designs of Table 2.1 with $R_T = 50\ \Omega$ and $C_L = 5$ pF. The results of simulation are rather disappointing and indicate that parasitics cause appreciable deviations from the ideal characteristics and that the distortions increase with increasing coefficient of coupling. Simulations with only the device parasitics indicate that the BTC parasitics have dominant effect on the deviations from maximal flatness and reduction of the BWER. Realizing the full potential of the theoretical results derived here will therefore have to wait for improved technologies for inductor fabrication in future.

2.6 Concluding Discussion

A comprehensive theoretical analysis has been carried out in this chapter of the general asymmetrical BTC network used as the load of a wide-band amplifier, and it has been shown that the BWER achievable is unlimited, the limit being set only by practical considerations of tight coupling, large spreads in inductance and capacitance values, and of course, parasitics. Unlimited bandwidth has never been reported earlier, by either the BTC or any other network, and it is believed that the results of this chapter will set a new trend in the design of wide-band and ultra wide-band amplifiers.

Simulation results indicate that the realization of the full potential of the asymmetrical BTC in wide-band amplifiers will have to wait for substantial improvements in the technology of inductance fabrication.

Acknowledgements This work was supported by the Indian National Science Academy through the Honorary Scientist scheme. The author is thankful to Y. V. Joshi, S. K. Koul, R. Bhattacharya, and G. Chowdary for their help in preparing the manuscript.

References

1. P. Staric, An analysis of the tapped coil peaking circuit for wideband/-pulse amplifiers. Elektrotehniski Vestnik. **49**, 66–79 (1982)
2. P. Staric, Three and four pole tapped coil circuits for wideband/-pulse amplifiers. Elektrotehniski Vestnik. **50**, 129–137 (1983)
3. E.M. Chase, W. Kennan, A power distributed amplifier with constant-R network, in *Proceedings of the IEEE GaAs IC Symposium*, Genelefe, FL (1986), pp. 13–17
4. C. Hutchinson, W. Kennan, A low noise distributed amplifier with gain control, in *Proceedings of the IEEE GaAs IC Symposium*, Anaheim, CA (1987), pp. 119–122
5. D.L. Feucht, *Handbook of Analog Circuit Design* (Academic Press, San Diego, CA, 1990)
6. L. Selmi, D.B. Estreich, B. Ricco, Small-signal MMIC amplifier with bridged T-coil matching networks. IEEE J. Solid State Circuits **7**, 1093–1096 (1992)
7. J. Kim, J.-K. Kim, B.-J. Lee, M. S. Huang, H.-R. Lee, S.-H. Lee, N. Kim, D.-K. Jeong, W. Kim, Circuit techniques for a 40 Gb/s transmitter in 0.13 μm CMOS, in *Proceedings of the IEEE International Solid-State Circuits Conference*, San Francisco, CA (2005), pp. 150–151
8. S. Galal, B. Razavi, Broadband ESD protection circuits in CMOS technology, in *Proceedings of the IEEE International Solid-State Circuits Conference*, San Francisco, CA (2003), pp. 182–183
9. T. Toifl, M. Kossel, C. Menolfi, T. Morf, M. Schmatz, A 23 GHz differential amplifier with monolithically integrated T-coil in 0.09 μm CMOS technology, in *IEEE Microwave Theory and Techniques Society Symposium*, Philadelphia, PA (2003), pp. 239–242
10. S. Galal, B. Razavi, 10 Gb/s limiting amplifier and laser modulation driver in 0.18 μm CMOS technology. IEEE J. Solid State Circuits **38**, 2334–2339 (2003)
11. S. Galal, B. Razavi, Broadband ESD protection circuits in 0.18 mm CMOS technology. IEEE J. Solid State Circuits **38**, 2138–2146 (2003)
12. S. Galal, B. Razavi, 40 Gb/s amplifier and ESD protection circuit in 0.18 mm CMOS technology. IEEE J. Solid State Circuits **39**, 2389–2396 (2004)
13. T.H. Lee, *Planar Microwave Engineering: A Practical Guide to Theory, Measurement and Circuits* (Cambridge University Press, Cambridge, 2004)
14. S. Shekhar, J.S. Walling, D.J. Allstot, Bandwidth extension techniques for CMOS amplifiers. IEEE J. Solid State Circuits **41**, 2414–2439 (2005)
15. H. Zheng, S.Z. Lou, H.P. Leung, RF system architectures and circuit techniques for wideband transceiver, in *Proceedings of the IEEE International Workshop on Radio Frequency Integration Technology*, Singapore (2005), pp. 30–37
16. J. Paramesh, D.J. Allstot, Analysis of the bridged T-coil circuit using the extra-element theorem. IEEE Trans. Circuits Syst. II Express Briefs **53**, 1408–1412 (2006)
17. S.C. Dutta Roy, Comments on 'Analysis of the bridged T-coil circuit using the extra element theorem'. IEEE Trans. Circuits Syst. II Express Briefs **54**, 673–674 (2007)
18. M.T. Reiha, J.R. Long, Symmetrical monolithic T-coils for broadband IC design, in *IEEE Microwave Theory and Techniques Society Symposium*, Honolulu, HI (2007), pp. 1909–1911

19. E. Pillai, J. Wiess, Novel T-coil structures and implementation in a 6.4-Gb/s CMOS receiver to meet return loss specifications, in *Proceedings of the IEEE Electronic Components and Technology Conference*, Reno, NV (2007), pp. 147–153
20. D. Linton, S. Thijs, J. Borremans, M. Dehan, D. Tremouilles, M. Scholz, M. I. Natarajan, P. Wambacq, G. Groeseneken, T-diodes—a novel plug-and-play wideband RF circuit for ESD protection methodology, in *Proceedings of the EOS/ESD Symposium*, Anaheim, CA (2007), pp. 4A.1.1–4A.1.8
21. M. Kossel, C. Menolfi, J. Weiss, P. Buchmann, G.V. Bueren, L. Rodoni, T. Morf, T. Toifl, M. Schmatz, A T-coil enhanced 8.5 Gb/s high swing SST transmitter in 65 nm bulk CMOS with <-16 dB return loss over 10 GHz bandwidth. IEEE J. Solid State Circuits **43**, 2905–2920 (2008)
22. B.-J. Huang, C.-H. Wang, C.-C. Chen, M.-F. Leei, P.-C. Huang, K.-Y. Lin, H. Wang, Design of amplifier for a 60-GHz low-noise amplifier with RF ESD protection. IEEE Trans. Microw. Theory Tech. **57**, 298–305 (2009)
23. J. Kim, J.K. Kim, B.J. Lee, D.K. Jeong, Design optimization of on-chip inductive peaking structures for 0.13-μm CMOS 40-Gb/s transmitter circuits. IEEE Trans. Circuits Syst. I Regul. Papers **56**, 2544–2555 (2009)
24. S. Kim, S. Kim, G. Jung, K.-W. Kwm, J.-H. Chun, Design of a reliable broadband I/O employing T-coil. J. Semicond. Technol. Sci. **9**, 198–204 (2009)
25. S. Cao, J.H. Chun, S.G. Beebe, R.W. Dutton, ESD design strategies for high-speed digital and RF circuits in deeply scaled silicon technologies. IEEE Trans. Circuits Syst. I Regul. Papers **57**, 2301–2310 (2010)
26. Y.-S. Lim, J.-H. Lee, Miniature ultra-wideband power divider using bridged T-coils. IEEE Microw. Wirel. Compon. Lett. **22**, 391–393 (2012)
27. Y.S. Lim, J.H. Lee, Miniature Butler matrix design using glass based passive device technology for 2.5-GHz application. IEEE Trans. Microw. Theory Tech. **61**, 2590–2602 (2013)
28. W.-T. Fang, T.-H. Tseng, Y.-S. Lin, Miniature ultra-wideband bandpass filter using bridged T-coil. IEEE Microw. Wirel. Compon. Lett. **24**, 367–369 (2014)
29. T.-S. Hong, J.-M. Wu, L.-Q. Yang, S.-T. Fang, A novel modified T-equivalent circuit for modeling LTCC embedded inductors with a large bandwidth, in *Proceedings of the IEEE Microwave Theory and Techniques Society Symposium*, Philadelphia, PA (2003), pp. 1015–1018
30. H.-Y. Lin, J.-S. Fu, Analysis of magnetically coupled all-pass networks for phase-shifter design. IEEE Trans. Microw. Theory Tech. **62**, 2625–2637 (2014)
31. T. True, Bridged-T termination network. U.S. Patent 3, 155, 927 (1964)
32. F.A. Muller, High frequency compensation of RC amplifiers. Proc. IRE **43**, 1271–1276 (1954)
33. R.D. Middlebrook, Null double injection and the extra-element theorem. IEEE Trans. Educ. **32**, 167–189 (1989)
34. M.E. Van Valkenburg, *Introduction to Modern Network Synthesis* (Wiley, New York, 1960)
35. M. Abramowitz, I. Stegun, *Handbook of Mathematical Functions* (Dover, New York, 1970)
36. S.C. Dutta Roy, On the solution of quartic and cubic equations. IETE J. Educ. **47**, 91–95 (2006)

Chapter 3
Multiple Frequency Impedance Matching by Frequency Transformation

Abstract It is shown that the application of frequency transformation to a single-frequency impedance matching network simplifies the design of networks for impedance matching at any number N of frequencies. The required design procedure has been presented in formal network theoretic terms. A systematic procedure is given for generating all the equivalent networks for this purpose, from which the best one can be chosen according to a predefined criterion. Illustrative designs and some of the relevant circuits are given for some low values of N.

Keywords Frequency transformation technique · Multiple frequency impedance matching · Network design

3.1 Introduction

Frequency transformation is a well-known technique in filter design [1, 2], by which a normalized low-pass filter with a cutoff at 1 rad/s can be transformed to any other kind of filter, be it high-pass, bandpass, bandstop or a multi-bandpass [3] one. That this technique can be applied to design networks for impedance matching at multiple frequencies has recently been explored in [4]. The basic philosophy is that the frequency transformation technique simply transforms a frequency of 1 rad/s to a number of frequencies ω_i, $i = 1$ to N (henceforth assumed to be in ascending order), thereby reproducing the performance at 1 rad/s at the transformed frequencies. In filter design, performance refers to cutoff at a certain level of magnitude. However, there is no reason why performance cannot refer to some other feature like impedance matching. Basically, if one can design a network for matching the load impedance to the source impedance at a single frequency ω_0, which is normalized to 1 rad/s, then frequency transformation applied to this network should produce another network which exhibits the same phenomenon of impedance matching at multiple frequencies.

Source: Suhash C. Dutta Roy, "Network Design for Multiple Frequency Impedance matching by the Frequency Transformation Technique," *IETE Journal of Research*, vol 59, pp 698–703, November–December 2013.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_3

The purpose of this chapter is to supplement the work of [4] in several ways. First, we formulate the procedure in formal network theoretic terms. Second, we show that this allows us to generate, systematically, all possible networks for impedance matching at the given set of frequencies. From these, one can select the best one according to a predefined criterion. Next, we present the design formulas and some relevant circuits for matching at 2, 3, 4 and 5 frequencies. Finally, some general observations are made and scope of further work is indicated in the concluding section.

We shall assume in this chapter that the load is resistive, of value R_L, and likewise, the source has a resistive impedance, of value R_S.

3.2 The Basic Network

Consider the simplest configuration for impedance matching, namely, the L-network shown in Fig. 3.1. Z_s is an LC impedance and Y_p is an LC admittance such that when the network is terminated in R_L, the input impedance is R_S at the frequencies of interest. Let, at $s = j\omega$, $Z_s = jX_s$ and $Y_p = jB_p$, where X denotes reactance and B denotes susceptance. Then, at the frequencies of interest, we should have

$$R_S = jX_s + 1/\left(jB_p + G_L\right) \tag{3.1}$$

where $G_L = 1/R_L$. Cross multiplying, simplifying, and equating the real and imaginary parts on both sides give the two equations

$$X_S B_p = 1 - R_S G_L \text{ and } R_S B_p = X_S G_L \tag{3.2}$$

The second equation shows that X_s and B_p must be of the same sign, both positive or both negative. Combining this fact with the first Equation in (3.2), *we note that* $R_S G_L$ *must be less than unity,* i.e. R_S *must be less than* R_L. However, *this is no restriction because the other situation*, i.e. $R_S > R_L$, *can be taken care of* by simply interchanging the positions of R_S and R_L in Fig. 3.1. Eliminating B_p from the two Equations in (3.2), we get

$$X_s^2 = R_S(R_L - R_S) = R_1^2, \text{ say} \tag{3.3}$$

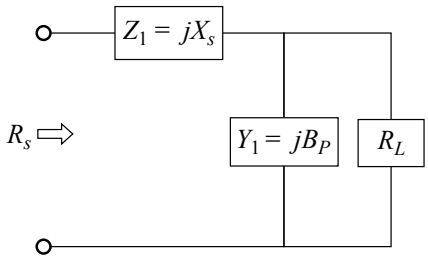


Fig. 3.1 The basic L-network

Or,

$$X_s = \pm R_1 \tag{3.4}$$

Combining this with the second Equation in (3.2) gives

$$B_p = \pm R_1/(R_L R_S) = \pm G_2, \text{ say} \tag{3.5}$$

As already stated, the signs in Eqs. (3.4) and (3.5) should be either both positive or both negative.

3.3 Impedance Matching at a Single Frequency

For matching R_L to R_S at a single frequency ω_0, we can choose either an inductance L_s for X_s and a capacitance C_p for B_p, or a capacitance C_s for X_s and an inductance L_p for B_p. In the first case, to be referred to as Design 1 (D1)

$$L_s = R_1/\omega_0 \text{ and } C_p = G_2/\omega_0 \tag{3.6}$$

while for the alternative design, to be called Design 2 (D2),

$$C_s = 1/(R_1\omega_0) \text{ and } L_p = 1/(G_2\omega_0) \tag{3.7}$$

We shall mostly use D1 in our further discussions, it is being implied that D2 is equally applicable, giving another set of solutions.

3.4 Application of Frequency Transformation

Let ω_0 be normalized to 1 rad/s and let S denote the complex frequency variable in the single-frequency matching network. Then the problem of multiple frequency impedance matching reduces to that of transforming $S = \pm j1$ to another set of frequencies $s = \pm j\omega_i$, $i = 1$ to N, where s denotes the complex frequency variable in the transformed network. It is well known [1, 2] that the functional relationship between S and s is an LC impedance or admittance function of the general form

$$S = (k_0/s) + k_\infty s + \sum\nolimits_i k_i s/\left(s^2 + q_i\right) \tag{3.8}$$

Using Eq. (3.8) on D2, for example, $C_s = 1/R_1$ (recall that ω_0 has been normalized to unity) transforms to a parallel connection of a capacitance of value k_∞/R_1, an inductance of value R_1/k_0, and a number of series LC circuits, the number being the

same as the upper limit in the summation of Eq. (3.8). The ith series LC circuit will have an inductance of value R_1/k_i and a capacitance of value $k_i/(q_i R_1)$. Similarly, the inductance $L_p = 1/G_2$ transforms to a series connection of an inductance of value k_∞/G_2, a capacitance of value G_2/k_0 and a number of parallel LC circuits, the number being the same as in the case of C_s. The ith parallel LC circuit will have an inductance of value $k_i/(q_i G_2)$ and a capacitance of value G_2/k_i. Similarly, Eq. (3.8) can be used on D1 to derive another circuit which can be used for the same purpose.

In addition, note that applied to an inductance, Eq. (3.8) results in a Foster 1 (F1) form of realization, while applied to capacitance, the result is a Foster 2 (F2) form of realization. Each of them can be converted to the other Foster form or Cauer 1 (C1) or Cauer 2 (C2) form. By different combinations of these forms for the two LC networks, a large number of competing circuits can be generated. The set can be further expanded by synthesizing each LC impedance by combining two or more of F1, F2, C1 and C2 forms to generate more circuits having the same impedance-matching properties.

A further dimension is added to the generation of competing circuits by the flexibility in choosing the transformation function. In general, for odd N, one can choose either of the first two terms in Eq. (3.8) and the summation with $i = 1$ to $(N - 1)/2$. For even N, one can choose both of the first two terms in Eq. (3.8) and the summation with $i = 1$ to $(N - 2)/2$; alternatively, one can use only the summation with $i = 1$ to $(N/2)$. The number of circuits that can be generated is, therefore, virtually indefinite. These can be compared, and the best one can be selected according to some appropriate criterion.

However, for small values of N, the number of alternative circuits is limited and the designer need not feel uncomfortable in leaving out some a from consideration. For example, for $N = 2$, we can choose either the first two terms only or a single term in the summation in Eq. (3.8). However, the latter is of the form of reciprocal of the former, and it is easily verified by choosing either of the two options, and using D1 as well as D2 suffice. Also, a single series or parallel LC circuit cannot be converted to any other form, thus giving only two alternative circuits for impedance matching at two frequencies. We shall investigate this case in more details in the next Section.

For $N = 3$, either of the first two terms and a single summation in Eq. (3.8) can be selected and applied to D1 as well as D2, giving four alternative circuits. In addition, each Foster form can be converted to the other, and four combinations, namely, (F1, F1), (F1, F2), (F2, F1) and (F2, F2), are possible for the series and shunt arms of Fig. 3.1. Converting a Foster form of three elements to either of the Cauer forms is not useful here because it gives the other Foster form. It is also not possible to make partial synthesis with one of the Foster or Cauer forms and switching to another form. Thus, a total of eight competing circuits can be found. We shall investigate this case in Sect. 3.6, but shall work out only one design.

It can be easily appreciated that the number of alternative circuits grows with increasing N. We shall not discuss this issue further.

For finding the element values in any circuit, we require the values of $k_i's$ and $q_i's$ in terms of $\omega_i's$. These can be obtained by the following procedure. Put $S = +j1$ and $s = j\omega$ in the chosen transformation equation and simplify. This will give an algebraic

equation in ω of order N, containing the constants $k_i's$ and $q_i's$ in the coefficients. To find how these constants are related to $\omega_i's$, one way is to make a reactance plot of the transformation, i.e. a plot of S/j versus s/j to find out at which values of ω, S/j becomes $+1$. These frequencies must correspond to $\omega_i's$, but not all with positive signs. A simpler way is to use the facts that the reactance (or susceptance) is an odd function of ω and its slope is strictly positive. From these, it is easily shown that if S has a pole at $s = 0$, i.e. if the first term in Eq. (3.8) is included in the chosen transformation, then the required frequencies are $-\omega_1$, ω_2, $-\omega_3$. On the other hand, if S has a zero at $s = 0$, then the required frequencies are ω_1, $-\omega_2$, ω_3. Note that the signs alternate, and it is adequate to find the sign of ω_1 only. Now equate the previously obtained equation to $(\omega + \omega_1)\ (\omega - \omega_2)\ (\omega + \omega_3)\ldots. = 0$, or $(\omega - \omega_1)\ (\omega + \omega_2)\ (\omega - \omega_3)\ldots. = 0$, as the case may be. By comparing the coefficients, one can easily find the required $k_i's$ and $q_i's$. The procedure will be amply clear from the examples worked out in the following Sections.

3.5 Dual-Frequency Impedance Matching

Let the frequencies at which matching is needed be ω_1 and ω_2. We use the transformation

$$S = (k_0/s) + k_\infty s \tag{3.9}$$

As mentioned in the previous section, using this on D1 and D2 suffices to find the only possible competing networks. Note, in passing, that Eq. (3.9) is the well-known low-pass to bandpass transformation. By putting $S = +j1$ and $s = j\omega$ in Eq. (3.9) and simplifying, we get the following quadratic equation in ω:

$$\omega^2 - (1/k_\infty)\omega - (k_0/k_\infty) = 0 \tag{3.10}$$

Since S has a pole at $s = 0$, the roots of this equation are $-\omega_1$ and ω_2, so that Eq. (3.10) should be identical to the following:

$$\omega^2 - (\omega_2 - \omega_1)\omega - \omega_1\omega_2 = 0 \tag{3.11}$$

Comparing coefficients, we get

$$k_\infty = 1/(\omega_2 - \omega_1) \text{ and } k_0 = \omega_1\omega_2/(\omega_2 - \omega_1) \tag{3.12}$$

The networks obtained by using Eq. (3.9) on D1 and D2 are shown in Fig. 3.2a and b, respectively.

We shall now compare the two designs with the help of a numerical example. Let $\omega_1 = 0.5$ rad/s and $\omega_2 = 0.7$ rad/s; these are normalized frequencies and the results can be applied to any situation where $\omega_1{:}\omega_2 = 5{:}7$ by appropriate frequency

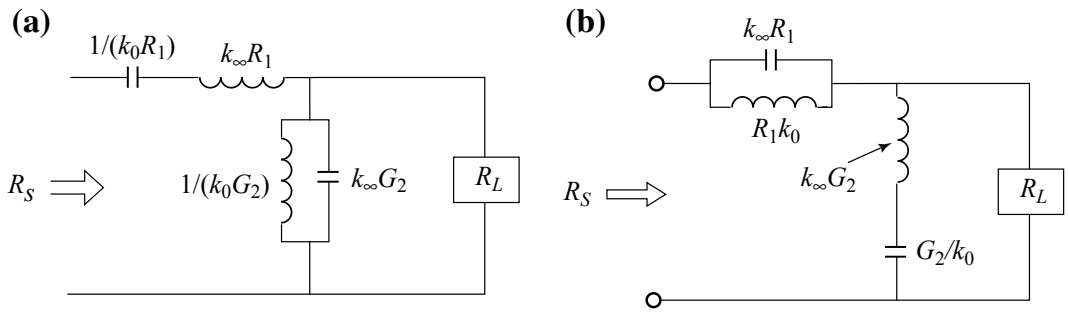


Fig. 3.2 Circuits obtained by application of Eq. (3.9) to **a** D1 and **b** D2

Table 3.1 Comparison of the two circuits of Fig. 3.2

Parameter	(a)	(b)
Total inductance	7.5 H	10.5714 H
Total capacitance	1.7142 F	5.2857 F
Highest inductance	5 H	10 H
Highest capacitance	1.1428 F	5 F
Inductance spread	2	17.5
Capacitance spread	2	17.5

scaling. Also, let $R_L = 2$ Ohms and $R_S = 1$ Ohms; again, the results can be applied to any situation where R_L: $R_S = 2$:1 by impedance scaling. Calculation gives $k_\infty = 5$, $k_0 = 1.75$, $R_1 = 1$ Ohms, and $G_2 = 0.5$ mho. The element values, from left to right, are calculated as 5H, 0.5714F, 2.5H and 1.1428F for Fig. 3.2a and 0.5714H, 5F, 10H and 0.2857F for Fig. 3.2b. Table 3.1 shows a comparison of the two circuits for some relevant implementation parameters. Clearly, the circuit of Fig. 3.2a would be a better choice from all considerations. Note that the inductance and capacitance spreads are the same in each circuit; this is expected from the general values given in Fig. 3.2.

3.6 Triple-Frequency Impedance Matching

To bring variety into experience, let us use the transformation

$$S = k_\infty s + k_1 s/\left(s^2 + q_1\right) \tag{3.13}$$

Applied to D1, this gives the circuit shown in Fig. 3.3. By following the procedure given in Sect. 4 and illustrated in the previous Section, the two cubic equations to be compared are obtained as follows:

$$\omega^3 - (1/k_\infty)\omega^2 - [q_1 + (k_1/k_\infty)]\omega + (q_1/k_\infty) = 0 \tag{3.14}$$

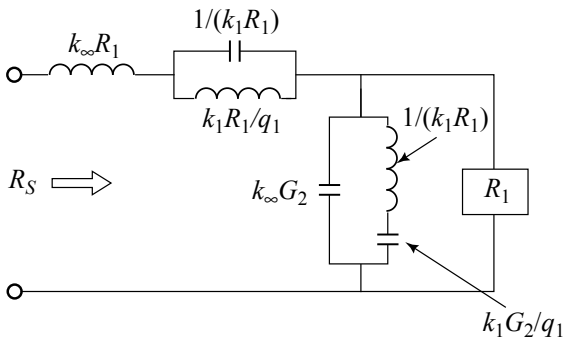


Fig. 3.3 Circuit obtained by applying Eq. (3.19) to D1

And

$$(\omega - \omega_1)(\omega + \omega_2)(\omega - \omega_3) = \omega^3 - (\omega_1 - \omega_2 + \omega_3)$$
$$\text{or,} \quad \omega^2 - (\omega_1\omega_2 + \omega_2\omega_3 - \omega_1\omega_3)\omega + \omega_1\omega_2\omega_3 = 0 \tag{3.15}$$

Comparison gives

$$k_\infty = \omega_1 - \omega_2 + \omega_3, q_1 = \omega_1\omega_2\omega_3 k_\infty, \text{ and}$$
$$k_1 = k_\infty(\omega_1\omega_2 + \omega_2\omega_3 - \omega_1\omega_3 - q_1) \tag{3.16}$$

The element values of the circuit in Fig. 3.3 can now be calculated.

3.7 Impedance Matching at Four Frequencies

In this as well as the next Section, for brevity, we shall skip the routine steps and the circuit diagrams, and shall only give the design equations. We choose the transformation

$$S = \left[k_1 s/\left(s^2 + q_1\right)\right] + \left[k_2 s/\left(s_m^2 + q_2\right)\right] \tag{3.17}$$

The roots of the quartic equation obtained from Eq. (3.17) should be ω_1, $-\omega_2$ ω_3 and $-\omega_4$. The equations for finding $k_i's$ and $q_i's$ are

$$k_1 + k_2 = -\omega_1 + \omega_2 - \omega_3 + \omega_4 = A, \text{ say} \tag{3.18}$$

$$q_1 + q_2 = \omega_1\omega_2 - \omega_1\omega_3 + \omega_1\omega_4 + \omega_2\omega_3 - \omega_2\omega_4 + \omega_3\omega_4 = B, \text{say} \tag{3.19}$$

$$k_1q_2 + k_2q_1 = \omega_2\omega_3\omega_4 - \omega_1\omega_3\omega_4 + \omega_1\omega_2\omega_4 - \omega_1\omega_2\omega_3 = C, \text{say} \tag{3.20}$$

$$q_1q_2 = \omega_1\omega_2\omega_3\omega_4 = D, \text{say} \tag{3.21}$$

q_1 and q_2 can be found from Eqs. (3.19) and (3.21) by solving a quadratic equation. This gives two solutions, one being the reciprocal of the other. Either of them can be chosen for q_1; then the other is q_2. Substituting these values in Eq. (3.20) and combining with Eq. (3.18), one can find k_1 and k_2. The final results are

$$q_1 = (1/2)[B + \sqrt{B^2 - 4D})], q_2 = D/q_1, k_1 = (Aq_1 - C)/(q_1 - q_2), \text{ and } k_2 = A - k_1 \tag{3.22}$$

It is easily shown that interchanging the values of q_1 and q_2 has the effect of interchanging the values of k_1 and k_2, so that the circuit does not change.

3.8 Impedance Matching at Five Frequencies

Let the transformation function be

$$S = (k_0/s) + \left[k_1 s/\left(s^2 + q_1\right)\right] + \left[k_2 s/\left(s^2 + q_2\right)\right] \tag{3.23}$$

By applying Eq. (3.23) to D1 or D2, the resulting circuit configuration can be easily drawn. For the constants, the equations to be solved are obtained as

$$k_0 + k_1 + k_2 = \omega_1 - \omega_2 + \omega_3 - \omega_4 + \omega_5 = A, \text{ say} \tag{3.24}$$

$$q_1 + q_2 = \omega_1\omega_2 - \omega_1\omega_3 + \omega_1\omega_4 - \omega_1\omega_5 + \omega_2\omega_3 - \omega_2\omega_4 + \omega_2\omega_5 + \omega_3\omega_4 - \omega_3\omega_5 + \omega_4\omega_5 = B, \text{ say} \tag{3.25}$$

$$\begin{aligned} k_1q_2 + k_2q_1 + k_0(q_1 + q_2) = {} & \omega_1\omega_2\omega_3 - \omega_1\omega_2\omega_4 + \omega_1\omega_2\omega_5 + \omega_1\omega_3\omega_4 \\ & - \omega_1\omega_3\omega_5 + \omega_1\omega_4\omega_5 - \omega_2\omega_3\omega_4 + \omega_2\omega_3\omega_5 \\ & - \omega_2\omega_4\omega_5 + \omega_3\omega_4\omega_5 = C, \text{ say} \end{aligned} \tag{3.26}$$

$$q_1q_2 = \omega_1\omega_2\omega_3\omega_4 + \omega_1\omega_2\omega_4\omega_5 - \omega_1\omega_2\omega_3\omega_5 - \omega_1\omega_3\omega_4\omega_5 + \omega_2\omega_3\omega_4\omega_5 = \mathrm{D}, \text{ say} \tag{3.27}$$

$$k_0q_1q_2 = \omega_1\omega_2\omega_3\omega_4\omega_5 = E, \text{ say} \tag{3.28}$$

As in the case of the previous Section, q_1 and q_2 can be solved from Eqs. (3.25) and (3.27) by solving a quadratic equation, whose solutions are reciprocals of each other. Also, as in the previous case, interchanging q_1 and q_2 has the effect of interchanging the values of k_1 and k_2. Once q_1 and q_2 are determined, k_0 is found from Eq. (3.28), and k_1 and k_2 are found from Eqs. (3.24) and (3.26). The final results are as follows:

$$\begin{aligned} & q_1 = (1/2)[B + \sqrt{B^2 - 4D}], q_2 = D/q_1, k_0 = E/D, \\ & k_1 = \{[A - (E/D)]q_1 - [C - (BE/D)]\}/(q_1 - q_2) \text{ and } k_2 = A - k_0 - k_1 \end{aligned} \tag{3.29}$$

3.9 Summary of the Procedure

The steps in the design of networks for impedance matching at multiple frequencies can be summarized as follows:

Step 1. Design the basic network D1 or D2 or both for the given R_L and R_S with $\omega_0 = 1$ rad/s.
Step 2. Choose the transformation function as follows. For $N = 2$, choose either the first two terms only or a single term in the summation in Eq. (3.8). For odd $N \geq 3$, choose either of the first two terms in Eq. (3.8) and the summation with $i = 1$ to $(N - 1)/2$. For even $N \geq 4$, choose the first two terms in Eq. (3.8) and the summation with $i = 1$ to $(N - 2)/2$; alternatively, use only the summation with $i = 1$ to $(N/2)$.
Step 3. Apply the transformation to D1 (or D2 or both) to obtain the circuit configuration (s) and element values in terms of $k_i's$ and $q_i's$.
Step 4. Put $S = j1$ and $s = j\omega$ in the transformation function and simplify to get an Nth order algebraic equation in ω.
Step 5. Identify the roots of the equation obtained above as $\omega_1, -\omega_2, \omega_3, \ldots$ or $-\omega_1, \omega_2, -\omega_3, \ldots$ depending on whether S has a pole at the origin or not. Write the equation in terms of $\omega_i's$ and compare its coefficients with those of the equation obtained in Step 4.
Step 6. Find the $k_i's$ and $q_i's$ in terms of $\omega_i's$ from the results of Step 5, and substitute them in the circuit (s) obtained in Step 3.
Step 7. Find some alternative circuits from the results of Step 6 by synthesizing the series and shunt LC impedances by Foster or Cauer techniques or a combination of them. Also, explore other possible transformations, if necessary, and finally, choose the best one according to a predefined criterion.

3.10 Concluding Discussion

In this chapter, we have presented the application of frequency transformation technique to impedance matching at multiple frequencies in formal network theoretic terms. This allows a multiplicity of solutions to be found out, from which one can make a selection based on an appropriate criterion. Explicit design equations have been given for $N = 2$ to 5.

While we have confined the discussion to only L-configuration, it should be obvious that the technique can be applied to other basic configurations like T or π. In all cases, however, the solution is far from being canonic. In fact, in the circuits considered in this chapter, we require 2 N number of components for matching at N frequencies. A reduction in component count may be possible by using other configurations of the basic network or by using two-port synthesis techniques.

We have considered only resistive source and load impedances and have assumed lossless inductances and capacitances. Some useful discussions on complex load

impedances, and component losses have been included in [4], but there is scope for formalizing the procedures in network theoretic terms and evolving design procedures for the general case.

In this chapter, we have not considered the bandwidth over which impedance matching takes place with a certain level of tolerance. There is a short discussion on this aspect in [4], but again, there is scope for further research, particularly because of the fact that if the $\omega_i's$ are not far apart, then the response around any ω_i is affected by that around other $\omega_i's$ close by. The positive aspect of this interaction is that one may be able to design wideband-matching circuits by placing $\omega_i's$ close together.

Acknowledgements This research was supported by the Honorary Scientist scheme of the Indian National Science Academy. The author thanks Shoubhik Dutta Roy for his help in the preparation of this manuscript.

References

1. L. Wienberg, *Network Analysis and Synthesis* (Wiley, Hoboken, New Jesrsey, 1962)
2. E.A. Guillemin, *Synthesis of Passive Networks* (Wiley, Hoboken, New Jesrsey, 1964)
3. S.C. Dutta Roy, On the design of multiple pass-band filters by frequency transformation. IETE J. Res. **58**, 20–23 (2012)
4. N. Nallam, S. Chatterjee, Multi-band frequency transformations, matching networks and amplifiers. IEEE Trans. Circuits Syst. I. Regular Papers, 60, 1635–1647 (2013)

Chapter 4
Triple Frequency Impedance Matching by Frequency Transformation

Abstract The purpose of this chapter is to bring to the attention of students and teachers of circuit theory, as well as of practising design engineers, a recently reported technique for impedance matching at multiple frequencies. This will be done through a specific example, namely that of design of a lumped network for matching a given load resistance R_L to a given source resistance R_S at three given frequencies ω_1, ω_2 and ω_3. The method is based on the application of the frequency transformation technique, commonly used in analog filter design, to a basic single-frequency matching network. By comparison with the existing approach of assuming a network configuration and solving a set of simultaneous equations, as reported in a recent paper, the new method is shown to have the advantages of design by synthesis, elegance and simplicity. Further, by virtue of the synthesis approach, the method allows a number of alternative networks to be designed, so that the best one can be selected on the basis of a predetermined criterion.

Keywords Frequency transformation · Network design · Impedance matching · Multiple frequency matching

4.1 Introduction

Modern multi-band mobile communication has created the need for microwave systems which can handle several frequencies by the same hardware, without the necessity for switching. This, in turn, requires networks which can match a given load to a given source at multiple frequencies. Most of the work reported in the literature on this subject has focussed on dual-band matching, using distributed networks [1–5]. Triple frequency matching has also been reported [6] using such networks. Distributed network systems become bulky, particularly when the operating frequencies are not very high, and lumped or hybrid systems have therefore been developed for compact devices [7, 8].

Source: S.C.Dutta Roy, "Triple Frequency Impedance Matching by Frequency Transformation." *IETE Journal of Education*, vol 55, pp 47–51, January–June 2014.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_4

Most of the network topologies used so far have been based on a design by analysis, i.e. by assuming a possible network configuration from heuristic considerations and experience, analysing it, and then solving a set of simultaneous equations to derive the component values. The purpose of this chapter is to bring to the attention of students and teachers of circuit theory as well as of practising design engineers a recently proposed method [9, 10] for the same purpose, which is based on the frequency transformation technique, as commonly used in analog filter design [11]. This will be done through a specific example, namely that of designing a network for matching a load resistance R_L to a given source resistance R_S at three given frequencies ω_1, ω_2 and ω_3. The reason for this specific choice is to compare the method with the existing approach for the same design, as reported in a recent paper [8]. It is shown that the new method is characterized by design by synthesis, elegance and simplicity. The circuit configuration evolves automatically, and by virtue of the synthesis approach, a number of alternative circuits can be easily found, from which the best one can be selected in accordance with a predetermined criterion.

First, for ready reference, we give a brief review of the general theory, and then discuss the design of the triple frequency transformer.

4.2 General Theory

4.2.1 Basic L-Network for Impedance Matching

We first consider the general L-network shown in Fig. 4.1 (called network D2 in [10]), where X_s is a reactance and B_p is a susceptance. For matching, we require the relation

$$R_s = 1/\left[jB_p + 1/(R_L + jX_s)\right] \tag{4.1}$$

to be satisfied at the frequency or frequencies of interest. Simplifying Eq. (4.1) and equating the real and imaginary parts on both sides give the following equations:

$$R_s\left(1 - X_s B_p\right) = R_L \tag{4.2a}$$

and

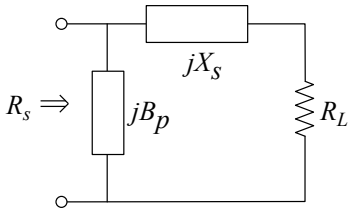


Fig. 4.1 The basic L-network for impedance matching

$$B_p R_L R_s = X_s \tag{4.2b}$$

Equation (4.2b) shows that B_p and X_s should be of the same sign. Also, combining Eq. (4.2a, b) to eliminate B_p, we get

$$X_s^2 = R_L(R_s - R_L) \tag{4.3}$$

Equation (4.3) shows that the configuration is applicable only if $R_s > R_L$. This is no restriction, however, because for the other case, i.e. $R_s < R_L$, the positions of the source and the load are to be interchanged. Equation (4.3) gives

$$X_s = \pm\sqrt{[R_L(R_s - R_L)]} = \pm R_1, \text{ say} \tag{4.4}$$

and correspondingly,

$$B_p = \pm R_1/(R_L R_S) = \pm G_2, \text{ say.} \tag{4.5}$$

As mentioned earlier, B_p and X_S should both be either positive or negative.

As discussed in the comprehensive network theoretic formulation of the technique by Dutta Roy [10], one could have a parallel development with the L-network having the shunt arm B_p across R_L and the series arm X_S in series with the source (this is called network D1 in [10]). As mentioned later in this chapter, this alternative is important in deriving other alternative equivalent circuits for matching.

4.2.2 Single-Frequency Matching

For matching at a single frequency ω_0, one choice is to have an inductor L_1 for X_s and a capacitor C_2 for B_p, while the other choice is to have C_1 for X_s and L_2 for B_p. Using the first choice, the element values are: $L_1 = R_1/\omega_0$ and $C_2 = G_2/\omega_0$.

4.2.3 Application of Frequency Transformation

Frequency transformation is a well-known technique in analog filter design [11] by which a normalized lowpass filter with a cut-off at 1 rad/sec can be transformed to any other kind of filter, be it high-pass, bandpass, band-stop or multi-bandpass, with the same tolerances as in the prototype. Basically, it transforms the frequency of 1 rad/sec to the cut-off frequencies of the desired filter. In more general terms, the performance at 1 rad/sec is reproduced in the transformed filter at the desired frequencies. There is no reason why performance cannot be impedance matching, as demonstrated in [9, 10] and this chapter.

Let S denote the complex frequency variable in the single-frequency matching network and let ω_0 be normalized to 1 rad/sec. Then the problem reduces to that of transforming $S = \pm j1$ to the multiple frequencies $s = \pm j\omega_k$, where, in general, $k = 1, \ldots, N$, s being the complex frequency variable in the transformed network. As is well known [12, 13], the transformation function is an LC immittance (impedance or admittance) function of the form

$$S = k_0/s + k_\infty s + \sum k_i s/\left(s^2 + q_i\right) \tag{4.6}$$

where the limits of the summation are determined by the number of frequencies at which matching is required. For example, if $N = 2$, then the first two terms in Eq. (4.6) suffice. Alternatively, one can use only the first term in the summation, which, in fact, is of the same form as the reciprocal of the first two terms. For $N = 3$, as is the case in the example to follow, either of the first two terms and one term in the summation serve the purpose. In general, for N odd and ≥ 3, either of the first two terms is to be taken along with the summation with i ranging from 1 to $(N-1)/2$. For N even and ≥ 4, there are two choices, viz. (1) take the first two terms and the summation with i ranging from 1 to $(N-2)/2$ or (2) take only the summation with i ranging from 1 to $(N/2)$. In addition, as already mentioned, there are two choices for the basic single-frequency matching network. Clearly, with the flexibility in choosing the basic network and the transformation function, one can generate a number of matching networks having the same performance. Some more alternative networks can be generated by synthesizing each of the reactance and susceptance functions by the other Foster or Cauer or mixed canonic forms. From this multiplicity of networks, one can choose the best one using some appropriate criterion like the total inductance or capacitance or the spread in the values of the components.

4.3 Triple Frequency Matching Network

We shall now illustrate the simplicity, elegance and flexibility of the frequency transformation technique by designing a triple-frequency matching network. The basic network is shown in Fig. 4.2a, where the fact that ω_0 has been normalized to 1 rad/sec has been taken into account. One appropriate transformation for this case is

$$S = k_0/s + k_1 s/\left(s^2 + q_1\right) \tag{4.7}$$

The impedance of L_1 (=R_1) thus transforms into

$$SL_1 = (k_0 R_1)/s + k_1 R_1 s/\left(s^2 + q_1\right) \tag{4.8}$$

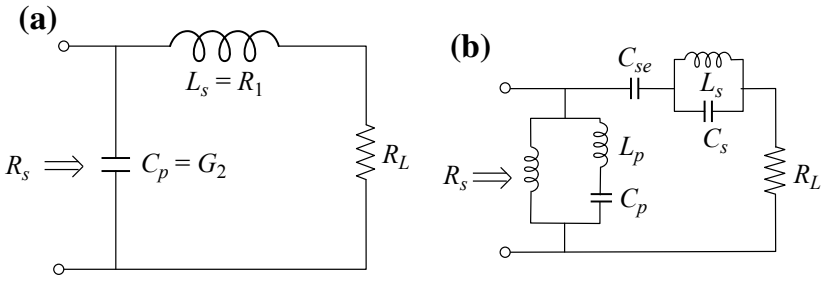


Fig. 4.2 Triple-frequency impedance transformer: **a** basic single-frequency matching circuit and **b** triple-frequency matching circuit with element values given in the text

The right-hand side of Eq. (4.8) represents a series connection of a capacitance $C_{\text{se}} = 1/(k_0\, R_1)$ and a parallel L_sC_s circuit with $L_s = k_1R_1/q_1$ and $C_s = 1/(k_1R_1)$, as shown in Fig. 4.2b.

Similarly, the admittance of C_2 (=G_2) transforms to the circuit shown in the shunt arm of Fig. 4.2b with the following element values: $L_{pe} = 1/(k_0\, G_2)$, $C_p = k_1\, G_2/q_1$ and $L_p = 1/(k_1\, G_2)$.

Now the only task that remains to be done is to find the values of k_0, k_1 and q_1. For this purpose, put $S = j1$ and $s = j\omega$ in Eq. (4.8); after simplification, we get the following equation in ω:

$$\omega^3 + (k_0 + k_1)\omega^2 - q_1\omega - k_0q_1 = 0 \tag{4.9}$$

The roots of this equation should give ω_1, ω_2 and ω_3, but not all with positive signs. To find which one(s) would be positive, we sketch the reactance function of Eq. (4.8) as shown in Fig. 4.3 and find the frequencies at which $S/j = 1$. Clearly, the roots of Eq. (4.9) are $-\omega_1$, ω_2 and $-\omega_3$. Figure 4.3 also shows that had we taken $S/j = -1$ in Eq. (4.8), the resulting cubic equation would have the roots ω_1, $-\omega_2$ and ω_3. Returning to Eq. (4.9), we see that the left-hand side should be equal to

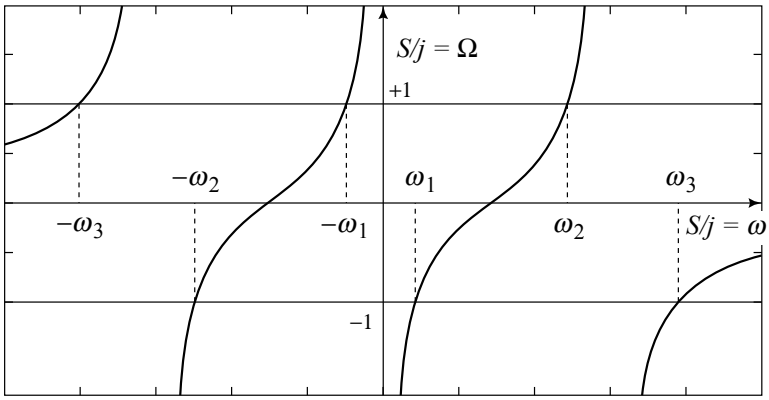


Fig. 4.3 Reactance sketch of Eq. (4.8)

$$\begin{aligned}&(\omega+\omega_1)(\omega-\omega_2)(\omega+\omega_3)\\&=\omega^3+(\omega_1-\omega_2+\omega_3)\omega^2-(\omega_1\omega_2+\omega_2\omega_3-\omega_3\omega_1)\omega-\omega_1\omega_2\omega_3\end{aligned}\tag{4.10}$$

Comparing the right-hand side of Eq. (4.10) with the left-hand side of Eq. (4.9), we get the following three equations:

$$\begin{aligned}k_0+k_1&=\omega_1-\omega_2+\omega_3,\\q_1&=\omega_1\omega_2+\omega_2\omega_3-\omega_3\omega_1,\ \text{and}\ k_0q_1=\omega_1\omega_2\omega_3\end{aligned}\tag{4.11}$$

Solving these equations, we finally get

$$\begin{aligned}q_1&=\omega_1\omega_2+\omega_2\omega_3-\omega_3\omega_1,\\k_0&=\omega_1\omega_2\omega_3/q_1,\ \text{and}\ k_1=\omega_1-\omega_2+\omega_3-k_0\end{aligned}\tag{4.12}$$

The circuit of Fig. 4.2b was previously derived by Lin, Chen and Zhao [8] following the conventional approach, i.e. by assuming the circuit, analysing it and then setting up and solving the equations obtained at the three frequencies. As a numerical example, they considered the following specifications: $R_L = 25\ \Omega$, $R_S = 50\ \Omega$ and $\{\omega_1, \omega_2, \omega_3\} = \{0.225, 0.425, 0.625\}$ GHz. Calculations from Eq. (4.12) and the expressions for the element values given under Eq. (4.8) give $C_{\text{se}} = 23.5$ pF, $L_s = 2.78$ nH, $C_s = 41.31$ pF, $L_{\text{pe}} = 29.38$ nH, $C_p = 2.22$ pF and $L_p = 51.64$ nH. These are exactly the same as those obtained in [8]. Simulation and experimental results given in [8] validate the design.

4.4 Alternative Circuits for Triple Frequency Matching

As mentioned in the previous section, several alternative circuits for the triple frequency impedance matching can now be obtained. Figure 4.2b network, as it is, has the shunt arm in Foster 2 form (henceforth abbreviated as F2) and the series arm in Foster 1 form (henceforth abbreviated as F1). The shunt arm can also be realized in F1 and the series arm can also be realized as F2. Accordingly, we have four alternative circuits, viz. (F2, F1), (F2, F2), (F1, F2) and (F1, F1). Note that for a three-element network, Cauer forms will reduce to either F1 or F2.

Second, using an inductance for B_p and a capacitance for X_s in the basic network of Fig. 4.1, and then applying the transformation of Eq. (4.7), one obtains a different circuit which also allows four different circuits to be generated.

Third, if the basic network is chosen as D1 of [10], in which there is a shunt arm B_p across R_L and a series arm X_s in series with R_S, then as in the case of D2 discussed so far, another eight networks can be generated, thus making a total of 16 alternative circuits for the same impedance matching purpose.

As mentioned in Sect. 2.2, an alternative transformation for triple-frequency impedance matching is

$$S = k_\infty s + k_1 s/\left(s^2 + q_1\right) \tag{4.13}$$

The values of the constants have been worked out in [10] along with a possible circuit by transforming the basic network D1, in which the series arm is F1 and the shunt arm is F2. Following the same procedure as in the case of Eq. (4.7), another 16 different circuits can be generated, thus making a grand total of 32 alternative circuits for any specified triple-frequency impedance matching problem.

All these circuits can now be compared in terms of total inductance, total capacitance, maximum inductance, maximum capacitance, inductance spread, and capacitance spread, as done in [10] for a dual-frequency impedance matching circuit, and the best one can be chosen in accordance with a predetermined criterion.

4.5 Conclusions

As discussed and demonstrated in this chapter, the application of frequency transformation on the basic single-frequency impedance matching network eases the problem of designing multiple frequency matching networks. It gives multiple solutions and facilitates the choice of the most suitable network in a given situation.

Acknowledgements This work was supported by the Indian National Science Academy through the Honorary Scientist scheme. The author thanks Shoubhik Dutta Roy and Yashwant V. Joshi for their help in the preparation of this chapter. Thanks are also due to the reviewer for constructive comments.

References

1. C. Monzon, A small dual frequency transformer in two sections. IEEE Trans. Microw. Theory Tech. **51**, 1157–61 (2003)
2. Y. Wu, Y. Liu, S. Li, A dual frequency transformer for complex impedances with two unequal sections. IEEE Microw. Wirel. Compon. Lett. **19**, 77–79 (2009)
3. X. Liu, Y. Liu, S. Li, F. Wu, Y. Wu, A three section dual band transformer for frequency dependent complex load impedance. IEEE Microw. Wirel. Compon. Lett. **19**, 611–613 (2009)
4. F. Paredes, G. Gonzales, J. Bonache, F. Martin, Dual band impedance matching networks based on split-ring resonators for application in RF identification (RFID). IEEE Trans. Microw. Theory Tech. **58**, 1159–1166 (2010)
5. M.L. Chuang, Dual band impedance transformer using two section shunt stubs. IEEE Trans. Microw. Theory Tech. **58**, 1257–1263 (2010)
6. C.S. Lee, C.-L Yang, A coupling matrix based design of triple band matching network. IEEE Microw. Wirel. Compon. Lett. **23**, 391–393 (2013)
7. Y. Liu, Y.J. Hao, Y. Hao, Lumped dual frequency transformer for frequency dependent complex loads. Prog. Electromagn. Res. **126**, 121–138 (2012)
8. Y. Liu, Y. Chen, Y.J. Zhao, Lumped triple-frequency impedance transformer. Electron. Lett. **48**, 1193–1194 (2012)

9. N. Nallam, S. Chatterjee, Multi-band frequency transformations, matching networks and amplifiers. IEEE Trans. Circuits Syst.-I: Regular Papers. **60**, 1635–1647 (2013)
10. S.C. Dutta Roy, Network design for multiple frequency impedance matching by the frequency transformation technique. IETE J. Res. **59**, 698–703 (2013)
11. S.C. Dutta Roy, On the design of a multiple pass-band filter by frequency transformation. IETE J. Res. **58**, 20–23 (2012)
12. L. Wienberg, *Network Analysis and Synthesis* (Wiley, Hoboken, NJ, 1962)
13. E.A. Guillemin, *Synthesis of Passive Networks* (Wiley, Hoboken, NJ, 1964)

Part II
Analog Signal Processing: Signal Processing by Distributed Networks

This part contain three chapters. In Chap. 5, we show that the familiar topological formulas to representation network can be extended to distributed networks also as a limiting case. Chapter 6 deals with equivalent circuits of lossy transmission lines. It is shown that short circuited lossy transmission lines are equivalent to a lossless line terminated in a resistance and that the equivalence is valid in two port sense also. Finally, in Chap. 7, we deal with the approximating a variable delay using a tapped delay line.

Chapter 5
Application of Topological Formulas to Distributed Parameter Networks

Abstract Topological formulas for lumped electrical networks are extended to distributed parameter networks, and explicit expressions are obtained for the two-port parameters of a non-uniform transmission line. The results agree with those derived earlier by Protonotarios and Wing, following a different procedure.

Keywords Topological formulas · Distributed network · Non-uniform transmission line

Nomenclature

$T_{2_{i_1 i_2 \ldots j_1 j_2}}$	a two-tree in which the sets of vertices $\{i_1, i_2, \ldots\}$ and $\{j_1, j_2, \ldots\}$ are in different connected parts.
$U_{i_1 i_2 \ldots, j_1 j_2 \ldots, k_1 k_2 \ldots}$	sum of admittance products of three-trees in each of which the sets of vertices $\{j_1, j_2, \ldots\}$, $\{j_1, j_2, \ldots\}$ and $\{k_1, k_2, \ldots\}$ are in different connected parts.
$V(Y)$	sum of admittance products of trees.
$W_{i_1 i_2 \ldots j_1 j_2 \ldots}(Y)$	sum of admittance products of two-trees $T_{2_{i_1 i_1 \ldots, j_1 j_2 \ldots}}$.
$[z]$	matrix of open-circuit impedance parameters of a two-port network.

5.1 Introduction

Topological formulas have now been well established for lumped electrical networks, both passive and active [1–3]. However, no attempt appears to have been made so far to apply these formulas to distributed parameter networks. In this chapter, topological formulas are used to obtain explicit expressions for the open-circuit

Source: A.Prabhakar & S.C. Dutta Roy, "Application of Topological Formulas to Distributed Parameter Networks," *JITE*, vol 18, pp 197–199, May 1972.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_5

Fig. 5.1 Passive two port

impedance parameters of one example of a distributed parameter network, viz. a non-uniform transmission line. Although essentially the same results have already been derived by Protonotarios and Wing [4] by a different procedure, it is believed that the approach adopted here is new, and it is hoped that it will stimulate further interest in the application of graph-theoretic techniques to distributed parameter networks.

5.2 Topological Formulas

We first recall the well-known topological formulas [1, p. 170] for the open-circuit impedance parameters of a two-port network (Fig. 5.1)

$$[z] = \frac{1}{V(Y)} \times \begin{bmatrix} W_{1,1'}(Y) & W_{12,1'2'(Y)} - W_{12',1'2}(Y) \\ W_{12,1'2'}(Y) - W_{12',1'2}(Y) & W_{2,2'}(Y) \end{bmatrix} \tag{5.1}$$

where

$$\det[z] = U(Y)/V(Y) \tag{5.2a}$$

and

$$U(Y) = U_{12',2,1'} + U_{1,2,1',2'} + U_{12,2',t'} + U_{1,2',1'2} \tag{5.2b}$$

5.3 Formulas for Ladder Networks and Extension to Non-uniform Transmission Lines

Consider a section of non-uniform transmission line of length l, for which $z(x)$ and $y(x)$ denote the series impedance and shunt admittance, respectively, per unit length of the line, at a distance x from the input port. We approximate the line by an n-section ladder as shown in Fig. 5.2, with

$$\Delta x = l/n,\ Z_k = z(k\Delta x)\Delta x \text{ and } Y_k = y(k\Delta x)\Delta x \tag{5.3}$$

In applying Eqs. (5.1) and (5.2a, 5.2b) to Fig. 5.2, we first note that

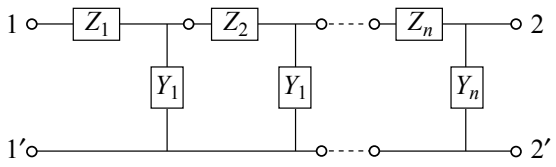


Fig. 5.2 Non-uniform line approximated by ladder

$$W_{12,1,2}(Y) = U_{12,2,1} = U_{12,2,1} = U_{1,2,12} = 0 \tag{5.4}$$

We may, therefore, rewrite $[z]$ and $\det[z]$ in the form

$$[z] = \frac{1}{\left(\prod_{i=1}^{n} Z_i\right) V(Y)} \times \begin{bmatrix} \left(\prod_{i=1}^{n} Z_i\right) W_{1,1'}(Y) & \left(\prod_{i=1}^{n} Z_i\right) W_{12,1'2'}(Y) \\ \left(\prod_{i=1}^{n} Z_i\right) W_{12,1'2'}(Y) & \left(\prod_{i=1}^{n} Z_i\right) W_{2,2'}(Y) \end{bmatrix} \tag{5.5}$$

and

$$\det[z] = \frac{\left(\prod_{i=1}^{n} Z_i\right) U_{1,2,1'2'}}{\left(\prod_{i=1}^{n} Z_i\right) V(Y)} \tag{5.6}$$

where the factor $\prod_{i=1}^{n} Z_i$ has been introduced in the numerator and denominator for later convenience.

Now, every tree of the ladder must contain at least one of the shunt arms. A tree is formed by any non-empty subset of the shunt arms, together with all the series arms preceding the first shunt arm in the subset, all the series arms succeeding the last shunt arm in the subset and all but one of the series arms between any two successive shunt arms in the subset. Finding out all the trees, multiplying each tree admittance product by $\prod_{i=1}^{n} Z_i$, and summing, we obtain

$$\left(\prod_{i=1}^{n} Z_i\right) V(Y) = \sum_{\substack{j=1 \\ i_1 < i_{2'} \le i_2 < \cdots < ij' \le ij}}^{n} Y_{i_1} Z_{i_{2'}} Y_{i_{2'}} Z_{i_3} \ldots Z_{ij'} Y_{ij} \tag{5.7}$$

Thus, a typical term in the right-hand side of Eq. (5.7) is obtained by choosing an arbitrary set of shunt admittances and exactly one of the series impedances between each pair of successive shunt admittances of this set, and taking their product. We may expand Eq. (5.7) in the form

$$\left(\prod_{i=1}^{n} Z_i\right) V(Y) = \sum_{i} Y_i + \sum\sum\sum_{i<j\le k} Y_i Z_i Y_k + \sum\sum\sum\sum\sum_{i<j\le k<m\le p} \times Y_i Z_j Y_k Z_m Y_p + \cdots \tag{5.8}$$

Applying similar arguments to the two-trees and three-trees, we get

$$\left(\prod_{i=1}^{n} Z_i\right) W_{1,1'}(Y) = 1 + \Sigma\Sigma_{i\le j} Z_i Y_j + \Sigma\Sigma\Sigma\Sigma_{i\le j<k\le m} Z_i Y_j Z_k Y_m + \cdots \tag{5.9}$$

$$\left(\prod_{i=1}^{n} Z_i\right) W_{1,2,1',2'}(Y) = 1 \tag{5.10}$$

and

$$\left(\prod_{i=1}^{n} Z_i\right) U_{1,2,1',2'}(Y) = \sum_i Z_i + \Sigma\Sigma\Sigma_{i\le j<k} Z_i Y_j Z_k + \sum\Sigma\Sigma\Sigma\Sigma_{i\le j<k\le m<p} Z_i Y_j Z_k Y_m Z_p + \cdots \tag{5.11}$$

The formula for $W_{2,2'}(Y)$ can be written in the same fashion as $W_{1,1'}(Y)$. In calculating $W_{12,1,'2'}(Y)$, we note that there is exactly one two-tree $T_{2_{12,1',2'}}$ and this two-tree is comprised of all of the series arms and none of the shunt arms, and hence the result of Eq. (5.10).

The formulas given above for the n-section ladder are equivalent to those proposed earlier [5, 6] for non-recurrent ladder networks. However, the above formulas have a simple topological interpretation, as stated in Eq. (5.7), and, in addition, the determinant of the z-parameters has been written in an explicit form, so that conversion to other types of parameters will be simple and routine.

We shall now extend the above formulas to the situation under consideration, viz. the non-uniform transmission line. For this, we allow Δx to approach zero, so that n approaches infinity. Consider the expansion on the right-hand side of Eq. (5.8), for example. It is easy to see that

$$\operatorname*{Lim}_{x\to\infty} \sum_i Y_i = \int_0^l y(x)\cdot dx \tag{5.12}$$

$$\operatorname*{Lim}_{\substack{x\to\infty \\ \Delta x\to 0}} \Sigma\Sigma\Sigma_{i<j\le k} Y_i Z_j Y_k = \int_0^l \int_0^{x_3} \int_0^{x_2} y(x_1) z(x_2) y(x_3) \times dx_1 dx_2 dx_3 \tag{5.13}$$

and so on. In a similar fashion, the integral equivalents of Eqs. (5.9) through (5.11) can be obtained. Finally, we have the following explicit formulas for the z-parameters of a non-uniform transmission line:

$$[z] = \frac{1}{a}\begin{bmatrix} b & 1 \\ 1 & c \end{bmatrix} \tag{5.14}$$

and

$$\det[z] = d/a \tag{5.15}$$

where

$$a = \int_0^l y(x) + \int_0^l \int_0^{x_3} \int_0^{x_2} y(x_1)z(x_2)y(x_3)dx_1dx_2dx_3 + \int_0^l \int_0^{x_5} \int_0^{x_4} \int_0^{x_3} \int_0^{x_2} y(x_1)z(x_2)y(x_3)z(x_4)y(x_5)dx_1 \times dx_2dx_3dx_4dx_5 + \cdots \tag{5.16}$$

$$b = 1 + \int_0^l \int_0^{x_2} z(x_1)y(x_2)dx_1dx_2 + \int_0^l \int_0^{x_4} \int_0^{x_3} \int_0^{x_2} z(x_1)y(x_2)z(x_3)y(x_4)dx_1dx_2 \times dx_3dx_4 + \ldots \tag{5.17}$$

$$c = 1 + \int_0^l \int_0^{x_2} y(x_1)z(x_2)dx_1dx_2 + \int_0^l \int_0^{x_4} \int_0^{x_3} \int_0^{x_2} y(x_1)z(x_2)y(x_3)z(x_4)dx_1dx_2 \times dx_3dx_4 + \ldots \tag{5.18}$$

and

$$d = \int_0^l z(x)dx + \int_0^l \int_0^{x_3} \int_0^{x_2} z(x_1)y(x_2)z(x_3)dx_1dx_2dx_3 + \int_0^l \int_0^{x_5} \int_0^{x_4} \int_0^{x_3} \int_0^{x_2} z(x_1)y(x_2)z(x_3)y(x_4)z(x_5) \times dx_1dx_2dx_3dx_4dx_5 + \cdots \tag{5.19}$$

In order to demonstrate the similarity of conversion to other parameters, we now obtain the chain parameters of the line

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \frac{1}{z_{21}} \begin{bmatrix} z_{11} & \det[z] \\ 1 & z_{22} \end{bmatrix} = \begin{bmatrix} b & d \\ a & c \end{bmatrix} \tag{5.20}$$

These expressions for the chain parameters are seen to agree with those derived earlier by Protonotarios and Wing [4].

References

1. S. Seshu, M.B. Reed, *Linear Graphs and Electrical Networks* (Addison-Wesley)
2. M.T. Jong, G.W. Zobrist, IEEE Trans. CT-15 **251** (1968)
3. A. Prabhakar, Generalized topological formulas for linear network functions. Ph.D. thesis, Indian Institute of Science, Bangalore, 1967
4. E.N. Protonotarios, O. Wing, IEEE Int. Conv. Rec. Part 7 **1** (1965)
5. F. Walker, Proc. IEEE **52**, 860 (1964)
6. S.C. Dutta Roy, Proc. IEEE **52**, 738 (1964)

Chapter 6
On Equivalent Representation of Lossy Transmission Lines

Abstract For short-duration surges, it was demonstrated that a short-circuited lossy transmission line is approximately equivalent to a lossless line terminated in a resistance, in so far as the output current is concerned. The same result is derived here in an alternative manner, which is believed to be simpler and more revealing. Further, the equivalence is also shown to hold in the two-port sense.

Keywords Lossy transmission lines · Equivalent representation

6.1 Introduction

In [1], Shah and Yavin demonstrated that for short-duration surges, a short-circuited dissipative transmission line can be equivalently represented as a lossless line terminated in a resistance $Z_0 \tanh [Rd/(2Z_0)]$. In deriving this result, the authors assumed an exciting voltage $v(0, t) = V_m \sin(\omega t + \theta)$ and compared the time domain solutions for the output current $i(d, \mathrm{t})$ in both the cases. We present here an alternative derivation of this result by carrying out analysis in the frequency domain only, starting from the expressions for the line voltage and current in terms of the sending end quantities. Our derivation is believed to be simpler and more revealing than that in reference [1]; simpler, because it does not require inversion of Laplace transforms, and more revealing, because it brings out the conditions for the validity of the equivalence in a straightforward manner and, in addition, also suggests a modification to obtain a closer equivalence. Finally, in this chapter, we point out that although the term 'equivalence' has been used by Shah and Yavin [1] to mean equality of y_{21} and y_{12}, (by reciprocity) parameters of the two lines only, it also holds good with regard to the other two-port parameters, viz. y_{11} and y_{22}, to the same extent of approximation.

Source: S. C. Dutta Roy, "On Equivalent Representation of Lossy Transmission Lines," *JIETE*, vol 19, pp 335–336, June 1973.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_6

6.2 Derivation

Consider a uniform transmission line of length d, characterized by the per unit length parameters R, L, G and C, which has been terminated in an impedance of Z_L. The voltage and current at any point x measured from the source end are given by Potter and Fich [2]

$$V(x) = V_s \cosh(\gamma x) - I_s Z_0 \sinh(\gamma x) \tag{6.1}$$

$$I(x) = I_s \cosh(\gamma x) - (V_s/Z_0) \sinh(\gamma x) \tag{6.2}$$

where V_s and I_s are the sending end voltage and current; $Z_0 = [(R + sL)/(G + sC)]^{1/2}$ is the characteristic impedance; and $\gamma = [(R + sL)(G + sC)]^{1/2}$ is the propagation constant of the line. At $x = d$, the boundary condition is

$$V(d) = Z_L I(d) \tag{6.3}$$

Combining Eqs. (6.1), (6.2) and (6.3), it is a matter of simple algebra to show that

$$I(d)/V_s = 1/\left[Z_0 \sinh(\gamma d) + Z_L \cosh(\gamma d)\right] \tag{6.4}$$

For a lossless line, $R = G = 0$, $Z_0 = (L/C)^{1/2}$ and $\gamma = s(LC)^{1/2}$; thus

$$I(d)/V_s|_{\text{lossless}} = 1/\left\{(L/C)^{1/2} \sinh\left[sd(LC)^{1/2}\right] + Z_L \cosh\left[sd(LC)^{1/2}\right]\right\} \tag{6.5}$$

For a lossy line with $G \cong 0$ (as in reference [1]),

$$[Z_0' = (L/C)^{1/2}[1 + R/(sL)]^{1/2} \text{ and } \gamma' = s(LC)^{1/2} \times [1 + /(sL)^{1/2} \tag{6.6}$$

where the prime has been used to distinguish this case from the previous one. For short-duration surges, $|s| \gg R/L$, and Eq. (6.6) can be approximated as

$$[Z_0' \cong (L/C)^{1/2}[1 + R/(2sL)] \text{ and } \gamma' \cong s(LC)^{1/2} \times [1 + R/(2sL)] \tag{6.7}$$

If this transmission line is shorted at the output end, i.e. $Z_L = 0$, then from (6.4) and (6.7), we get

$$I(d)/V_s|_{\text{lossy,shorted}} = 1/\left\{(L/C)^{1/2}[1 + R/(2sL)] \times \sinh\left[sd(LC)^{1/2} \cdot (1 + R/(2sL))\right]\right. \tag{6.8}$$

This can be put in the form

$$I(d)/V_s|_{\text{lossy,shorted}} = \frac{1}{[1 + R/(2sL)]\cosh[Rd/(2Z_0)]}$$

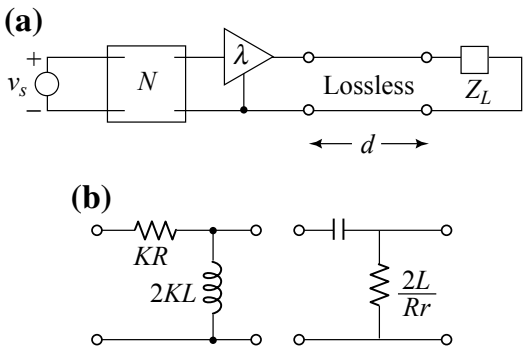


Fig. 6.1 **a** An equivalent circuit of a shorted lossy transmission line, where N stands for one of the two networks shown in (**b**) [$\lambda = 1$ if only condition (6.10) is to be eliminated, whereas $\lambda = \text{sech}\,[Rd/(2Z_0)]$ if both (6.10) and (6.11) are to be eliminated]. **b** Forms of N required in (**a**) [K is an arbitrary constant and r is an arbitrary resistance]

$$\times \frac{1}{(L/C)^{1/2}\sinh[sd(LC)^{1/2}] + (L/C)^{1/2}\tanh[Rd/(2Z_0)] \times \cosh[sd(LC)^{1/2}]} \tag{6.9}$$

Comparing Eqs. (6.9) with (6.5), we see that the shorted lossy line is equivalent to the lossless terminated line if

$$1 + R/(2sL) \cong 1 \tag{6.10}$$

$$\cosh\,[Rd/(2Z_0)] \cong 1 \tag{6.11}$$

and

$$Z_L = Z_0 \tanh\,[Rd/(2Z_0)] \tag{6.12}$$

where Z_0 refers to the lossless line.

The restriction given by Eq. (6.10) can be avoided by using the modified equivalent model shown in Fig. 6.1a, where N stands for either of the two networks shown in Fig. 6.1b, and $\lambda = 1$. It is obvious that the condition Eq. (6.11) can also be eliminated by using $\lambda = \text{sech}\,[Rd/(2Z_0)]$ in Fig. 6.1a. The resulting equivalent circuit is then exact as long as the approximation $[1 + R/(sL)]^{-1/2} \cong 1 + R/(2sL)$ is valid, and can be used for analog modelling as well as for digital computer studies of transients.

6.2.1 Two-Port Equivalence

The term equivalence as used in Ref. [1] and this chapter so far is in a restricted sense only. Considering the lossy and the lossless lines as two-port networks, as shown in Fig. 6.2, we have so far attempted to equate the y_{21} and y_{12} parameters only. Two-port equivalence will be valid if, in addition to y_{21}, the y_{11} and y_{22} parameters of the two

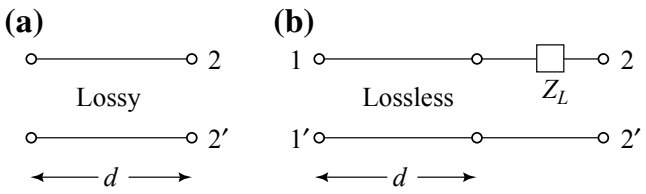


Fig. 6.2 Lossy and lossless lines considered as two ports

networks are also equal. To investigate this, we consider Fig. 6.2b first. Putting $x = d$ in Eqs. (6.1) and (6.2), combining with Eq. (6.3) and eliminating $I(d)$, we get

$$\left.\frac{1}{y_{11}}\right|_{\text{Fig.6.2b}} = Z_0 \frac{Z_L \cosh\left[sd(LC)^{1/2}\right] + Z_0 \sinh\left[sd(LC)^{1/2}\right]}{Z_0 \cosh\left[sd(LC)^{1/2}\right] + Z_L \sinh\left[sd(LC)^{1/2}\right]} \tag{6.13}$$

For the lossy case (Fig. 6.2a), it can be easily shown that

$$1/y_{11}|_{\text{Fig.6.2a}} = Z_0' \tanh(\gamma' d) \tag{6.14}$$

Combining this with Eq. (6.7) and simplifying, it can be shown that

$$\left.\frac{1}{y_{11}}\right|_{\text{Fig 6.2(a)}} \cong Z_0\left(1 + \frac{R}{2sL}\right) \times \frac{Z_0 \sinh\left[sd(LC)^{1/2}\right] + Z_0 \tanh\left[Rd/(2Z_0)\right] \times \cosh\left[sd(LC)^{1/2}\right]}{Z_0 \cosh\left[sd(LC)^{1/2}\right] + Z_0 \tanh\left[Rd/(2Z_0)\right] \times \sinh\left[sd(LC)^{1/2}\right]} \tag{6.15}$$

Comparing this with Eq. (6.13), we see that the y_{11} parameters of the two networks of Fig. 6.2 will be equal if Eqs. (6.10) and (6.12) are satisfied.

By symmetry, $y_{22} = y_{11}$ for Fig. 6.2a; however, for Fig. 6.2b,

$$1/y_{22}|_{\text{Fig.6.2.(b)}} = Z_L + Z_0 \tanh\left[sd(LC)^{1/2}\right] \tag{6.16}$$

Putting the value of Z_L from Eq. (6.12), we get

$$\left.\frac{1}{y_{22}}\right|_{\text{Fig.6.2b}} = \frac{Z_0 \sinh\left[sd(LC)^{1/2}\right] + Z_0 \tanh\left[Rd/(2Z_0)\right] \times \cosh\left[sd(LC)^{1/2}\right]}{Z_0 \cosh\left[sd(LC)^{1/2}\right]} \tag{6.17}$$

Comparing this with Eq. (6.15), we see that they are not equal even if Eq. (6.10) is satisfied; however, with Eq. (6.11) also satisfied, $\sinh\left[Rd/(2Z_0)\right] \cong \tanh\left[Rd/(2Z_0)\right) \cong 0$ and y_{22} of the two networks of Fig. 6.2 would be approximately equal.

References

1. K.R. Shah, Y. Yavin, "Equivalent representation of lossy transmission lines"—Part I. Proc. IEEE **59**, 1258 (1971)
2. J.L. Potter, S.J. Fich, *Theory of Networks and Lines* (Prentice Hall, New Delhi, 1965), p. 302

Chapter 7
Efficient Approximation of Variable Delay Using Tapped Delay Line

Abstract A variable delay is required for beam steering in an adaptive antenna array. This chapter presents an efficient design for approximating such a delay for broadband applications by using only one phase shifter and real coefficients. Exact mathematical formulas for determining the required weighting coefficients have also been derived.

Keywords Variable delay · Tapped delay line · Approximation

7.1 Introduction

Adaptive array systems have the ability to automatically respond to an unknown interference environment, by steering nulls and reducing the sidelobe levels in the direction of interference. The pattern of the array is steered by applying weighting across the antenna array elements. A tapped delay line filter, realizing a variable delay, is used in each channel of wideband adaptive arrays.

We, in this chapter, derive explicit as well as recursive mathematical formulas for determining the weights required for the structure, using the same criterion of approximation as used by Ko and Lim [1]. For calculation of N weights, the approach of reference [1] requires the inversion of an $N \times N$ matrix, needs N phase shifters, besides other involved evaluations. Our suggested formulas are in elegant forms and are shown to be computationally faster. Moreover, the proposed realization requires only one fixed phase shifter of 90 degrees, thereby making the structure simple and efficient.

Source: Shailey Minocha, S.C. Dutta Roy & Balbir Kumar, "Efficient Approximation of Variable Delay Using Tapped Delay Line," *JIETE*, vol 36, pp 129–131, 1990.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_7

7.2 Derivation of Formulas for the Weighting Coefficients

7.2.1 Recursive Formula

The ideal frequency response of delay τ is given by

$$\overset{\approx}{H}(\omega) = \exp(-j\omega\tau) \tag{7.1}$$

where ω is the radian frequency. We approximate Eq. (7.1) by the frequency response, *H(ω)*, of a tapped delay filter with *N* complex weights, i.e. by

$$H(\omega) = \sum_{i=1}^{N} c_i \exp[-j\omega(i-1)T] \tag{7.2}$$

where c_i are the desired weighting coefficients, T is the intertap delay, usually chosen to equal $m/(4f_o)$, where m is an integer and $f_o = \omega_o/(2\pi)$ is the centre frequency of the array; this choice of T, it can be shown, minimizes the complexity of the design.

We define the error of approximation by

$$E(\omega) \underline{\underline{\Delta}} |\overset{\approx}{H}(\omega) - H(\omega)| \tag{7.3}$$

The aim is to minimize the error in the frequency band of interest centred around fo. Choosing the criterion of approximation as the minimization of *E(ω)* in the maximally flat sense [1]. That is, by forcing

$$E(\omega)\,|_{\omega=\omega_o} = 0 \tag{7.4a}$$

and

$$\mathrm{d}^i E(\omega)/\mathrm{d}\omega^i\,|_{\omega=\omega_o} = 0, \quad i = 1, 2, \ldots, N-1 \tag{7.4b}$$

the following matrix equation results:

$$[a_{mn}][x_m] = [b_m], \quad m, n = 1, 2, \ldots, N \tag{7.5a}$$

where

$$a_{mn} = (n-1)^{m-1}$$

$$b_m = (\tau/T)^{m-1} \exp(-j\omega_o\tau) \tag{7.5b}$$

and

$$x_m = c_m \exp[-j\omega_o T(m-1)] \tag{7.5c}$$

This system of linear equations can be solved by Crout's method [4]. Following the steps of this method and after involved algebraic manipulations, Eq. (7.5a) is transformed to

$$[a'_{mn}][x_m] = [b'_m] \tag{7.6a}$$

where

$$a_{mn} = \begin{cases} 0, & n < m \\ 1, & n = m \\ \binom{n-1}{m-1}, & n \geq m \end{cases} \tag{7.6b}$$

$$b'_m = [1/(m-1)!](\tau/T)^{m-1}[\exp(-j\omega_o\tau)]\prod_{p=1}^{m-2}[1-(pT/\tau)] \tag{7.6c}$$

On solving Eq. (7.6), we finally get the following recursive formula for the weighting coefficients $c_{i:}$

$$c_i = \exp[j\omega_o T(i-1)]\left[\begin{array}{l} \frac{\exp(-j\omega_o\tau)}{(i-1)!}(\tau/T)^{i-1}\prod_{p=1}^{i-2}[1-(pT/r)] \\ \sum_{r=1}^{N-1}\binom{i+r-1}{i-1}c_{i+r}, \exp[-j\omega_o T(i+r-1)] \end{array}\right] \tag{7.7}$$
$$i = N, N-1_2, \ldots 1 \text{(descending order)}$$

7.2.2 *Explicit Formula*

Equation (7.6a) may also be written as

$$[x_m] = [a'_{mn}]^{-1}[b'_m] \tag{7.8}$$

where $[a'_{mn}]^{-1}$ is the inverse of the matrix $[a'_{mn}]$. Using Eqs. (7.5), (7.6) and (7.8), we obtain the following explicit formula for $c_{i:}$

$$c_i = \exp\{j\omega_o[(i-1)T-\tau]\}\left[\sum_{k=i}^{N}\frac{(-1)^{i+k}(\tau/T)^{k-1}}{(k-i)!(i-1)!}\prod_{p=1}^{k-2}[1-(pT/\tau)]\right] \tag{7.9}$$
$$i = 1, 2, 3 \ldots, N$$

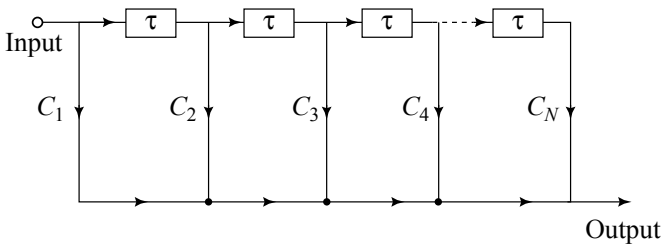


Fig. 7.1 General tapped delay line filter with N complex weights for realizing a delay of τ

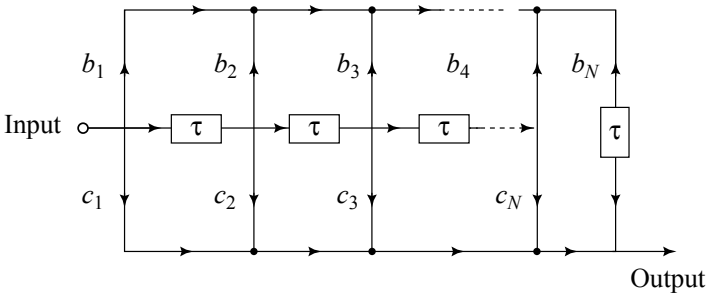


Fig. 7.2 Realization of the proposed design for the approximation of a variable delay τ, by using only one phase shifter

7.3 Implementation

The tapped delay line filter with N complex weights (c_i) is shown in Fig. 7.1. Let $c_i = a_i + j\, b_i$ $i =$ I, 2, 3,…, where a_i and b_i are real. Thus, the structure of Fig. 7.1 can be implemented as shown in Fig. 7.2 where only one phase shifter of 90 degrees (corresponding to multiplication by j) is required. Such an arrangement is highly efficient, besides being economical. The design proposed in reference [1], incorporating N phase shifters, suffers from inaccuracies of the phase shifters [2] (e.g. attenuation, phase instability, non-linearities, etc.). Our structure is relatively free from such drawbacks. The j factor, used in Fig. 7.2, can be efficiently realized by various methods, e.g. by the use of a Hilbert transformer [3].

7.4 Performance

The computation time of the weighting coefficients with the ICL 2960 system, for $N = 1$ to 7, has been plotted in Fig. 7.3 for $f_o\ \tau = 0.125$ and $f_o T = 0.25$. Curves (a) and (b) of Fig. 7.3 give, respectively, the computation time in milliseconds using the explicit formula and the recursive formula. Recursion, when used for evaluation of the weighting coefficients, is computationally faster than the direct or the explicit formula since a recursive formula makes use of the previously generated weight values in the algorithm. It is apparent (as expected) from Fig. 7.3 that the proposed recursive formula, Eq. (7.7), is more efficient than the explicit formula, Eq. (7.9), when the speed of computation is of prime importance. However, in recursion, there is a need to set aside memory storage arrays for the previously computed weights. Therefore, the explicit formula may be preferred when the economy of memory of the

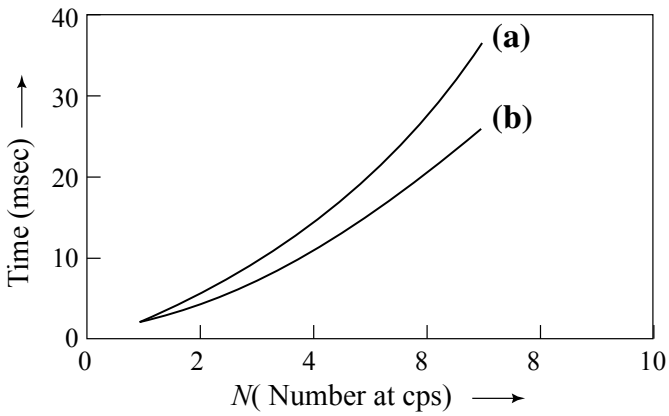


Fig. 7.3 Computation time for coefficients, c_i for delay of 0.5T for $N = 1$ to 7, **a** by using explicit formula (9), **b** by using recursive formula (7)

computer is the main concern. As is well known [4], it takes about $2\,N^3/3$ operations to solve a linear system of equations.

$\underset{\sim}{A}\,x = ba$, of order N, by Crout's method and about $8\,N^3/3$ operations for finding the matrix inversion $\overset{-1}{\underset{\sim}{A}}$. It is thus obvious that the computation time, if we adopt the matrix inversion approach of [1], is approximately four times the values given by curve (b) of Fig. 7.3.

7.5 Conclusions

Exact mathematical formulas for computation of the weighting coefficients, required for the variable delay structure, have been derived. The proposed design is efficient and requires only one fixed phase shifter in contrast to N phase shifters needed in the structure suggested in reference [1].

Acknowledgements The authors acknowledged the careful and constructive comments of the reviewer.

References

1. C.C. Ko, Y.C. Lim, Approximation of a variable-length delay line by using tapped delay line processing. Signal Process. **14**, 363–369 (1988)
2. M.I. Skolnik, *Radar Handbook* (McGraw-Hill Book Company Inc, New York, 1970)
3. L.R. Rabiner, R.W. Schafer, On the behaviour of minimax FIR digital Hilbert transformers. Bell. Sys. Tech. J. **53**, 363–391 (1974)
4. M.J. Maron, *Numerical Analysis—A Practical Approach*, 2nd edn. (Macmillan Publishing Company, New York, 1987)

Part III
Analog Signal Processing: Signal Processing by Active Networks

This part has six chapters, dealing with realization of various types of filters and an inductor, using Op-amp RC networks. Chapter 8 deals with such filters using a single operational amplifier. Tolerance minded design of an active RC filter form the contents of Chap. 9, while Chap. 10 is a supplement to Chap. 8, with more circuits using a single Op-amp and its different inputs. Second order active RC filters using Op-amps in low gain connection form the subject of Chap. 11. Inductor simulation and its applications are dealt with in Chap. 12. The effect of frequency limitation of Op-amp active RC filters is explored by the dominant pole technique in Chap. 13. Finally, Chap. 14 deals with 'Shadow' filters which is a recent innovation.

Chapter 8
Second-Order Active RC Filters Using a Single Operational Amplifier

Abstract The chapter deals with economic realizations of second-order active *RC* filters employing a single differential input operational amplifier as the active element. All the basic types of filters are considered and the sensitivities of Q and undamped natural frequency with respect to various parameters are evaluated. Further, it is shown that one circuit configuration can lead to several types of filter realizations having the same Q, simply by short circuiting or open circuiting certain elements, thereby reducing the manufacturing cost.

Keywords Active RC filter · Operational amplifier

8.1 Introduction

In [1], Dutta Roy proposed a procedure, alternative to that proposed by Mitra [2, 3] for realization of any rational transfer function with a single differential input operational amplifier. The new procedure requires the predetermination of three resistances and synthesis of three RC one ports, whereas Mitra's method requires synthesis of six RC one ports. Dutta Roy's method also allows much more flexibility in the choice of element values compared to that due to Mitra. The purpose of this chapter is to develop various second-order active RC filter realizations with as few passive components as possible. Starting with a basic circuit containing fifteen elements, as obtained with Dutta Roy's procedure for a general second-order transfer function. Further, it is shown that with one circuit configuration, low pass and bandpass realizations having the same Q, and, similarly from another circuit configuration, a high-pass, bandpass and band stop realizations having Q of the same order, may be obtained by short circuiting or open circuiting certain elements, thereby reducing the manufacturing cost.

Source: S. C. Dutta Roy, M. C. Srivastava & R. K. Patney, "Second Order Active RC Filters using a Single Operational Amplifier," *JIETE*, vol. 17, pp 408–414, November 1971.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_8

8.2 Basic Circuit

The basic circuit from which the various second-order realizations will be derived is shown in Fig. 8.1. The transfer function of this circuit can be expressed as

$$T(s) = \frac{V_0}{V_i} = \frac{Y_A(Y_D + Y_F) - Y_B(Y_C + Y_E)}{Y_F(Y_A + Y_C) - Y_E(Y_B + Y_D)} \tag{8.1}$$

In our realizations $Y_D = 0$; therefore, the transfer function reduces to

$$T(s) = \frac{Y_A Y_F - Y_B(Y_C + Y_E)}{Y_F(Y_A + Y_C) - Y_E Y_B} \tag{8.2}$$

By choosing the time constants of series RC impedances equal, removing two series RC impedances or by removing one series RC impedance and setting the time constants of remaining two series RC impedances equal, the transfer function $T(s)$ becomes a biquadratic one

$$T(S) = \frac{a_2 s^2 + a_1 s + a_0}{b_2 s^2 + b_1 s + b_0} \tag{8.3}$$

Next, by setting the appropriate numerator coefficient(s) to zero, a particular filter realization, with elements less than those hitherto used, can be obtained.

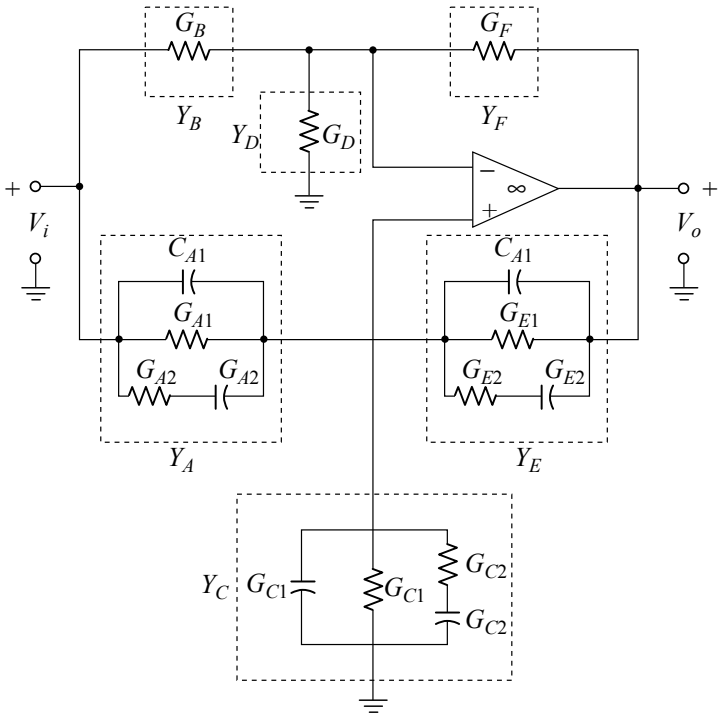


Fig. 8.1 Basic circuit for realizing various second-order transfer functions

8.3 Low-Pass Realization

To realize a low pass transfer function, we set in Fig. 8.1

$$\left.\begin{array}{l} G_{A2} = G_{C1} = G_{C2} = G_D = G_{E1} = 0 \\ C_{A2} = C_{C2} = C_{E1} = 0 \end{array}\right\} \tag{8.4}$$

with the admittances

$$\left.\begin{array}{ll} Y_A = C_{A1}s + G_{A1} & Y_B = G_B \\ Y_C = C_{C_1}s & Y_D = 0 \\ Y_E = \frac{G_{E2}s}{(G_{E2}/C_{E2})+8} & Y_F = G_F \end{array}\right\} \tag{8.5}$$

as shown in Fig. 8.2a. The coefficients of the transfer function, given by Eq. (8.3), become

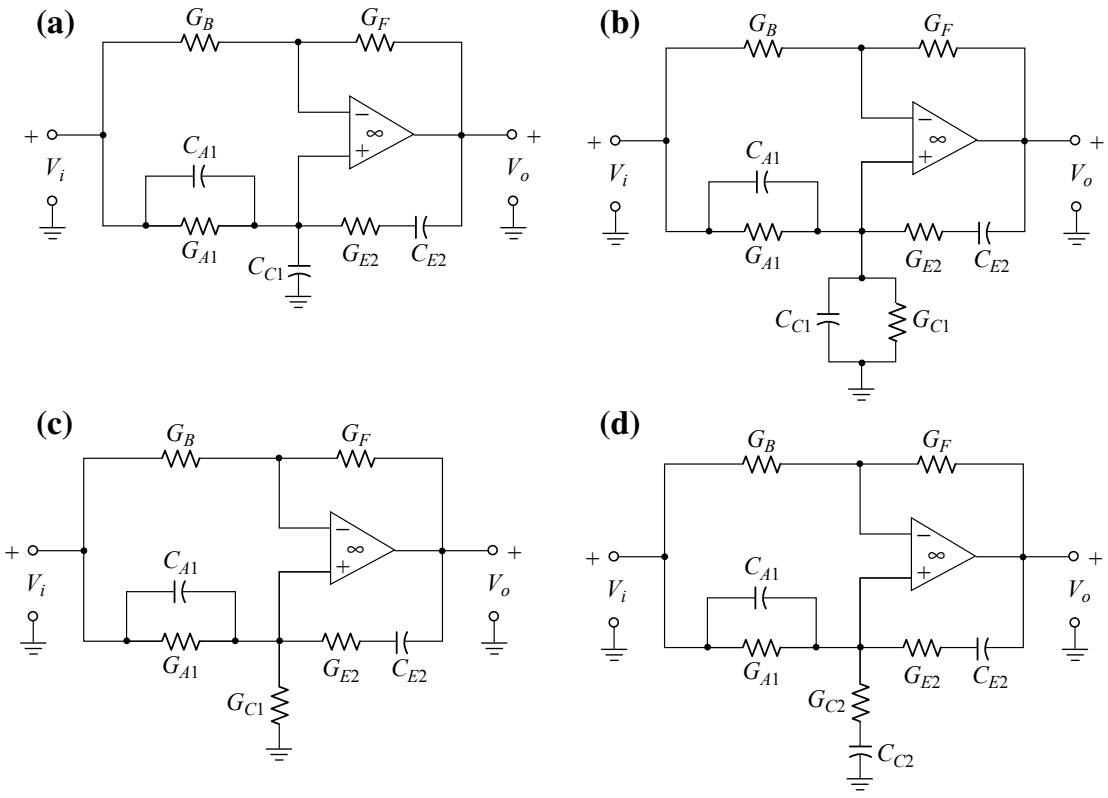


Fig. 8.2 **a** Low pass filter realization and second band stop filter realization with $a_0' > b_0'$. **b** Bandpass filter realizations. **c** High-pass filter realization and third band stop filter realization with $a_0' < b_0'$. **d** First band stop filter realization with $a_0' > b_0'$

$$\left.\begin{aligned} a_2 &= C_{A1}G_F - C_{C1}G_B \\ a_1 &= G_{A1}G_F - G_{E2}G_B + \tfrac{G_{E2}}{C_{E2}}(C_{A1}G_F - C_{C1}G_B) \\ a_0 &= \tfrac{G_{E2}}{C_{E2}}G_{A1}G_F \\ b_2 &= (C_{A1} + C_{C1})G_F \\ b_1 &= G_{A1}G_F - G_{E2}G_B + \tfrac{G_{E2}}{C_{E2}}(C_{A1} + C_{C1})G_F \\ &\text{and} \\ b_0 &= \tfrac{G_{E2}}{C_{E2}}G_{A1}G_F \end{aligned}\right\} \tag{8.6}$$

The general expressions for quality Q and natural frequency ω_n of this realization of seven elements are given by

$$\omega_n = \left[\frac{G_{A1}G_{E2}}{C_{E2}(C_{A1} + C_{C1})}\right]^{1/2} \tag{8.7}$$

and

$$Q = \frac{G_F[G_{A1}G_{E2}C_{E2}(C_{A1} + C_{C1})]^{\frac{1}{2}}}{G_{E2}(C_{A1} + C_{C1})G_F + C_{E2}(G_{A1}G_F - G_{E2}G_B)} \tag{8.8}$$

Setting $a_2 = 0$ and $a_1 = 0$ for low pass realization, we have

$$\left.\begin{aligned} &C_{A1}G_F = C_{C1}G_B \\ &\text{and} \\ &G_{A1}G_F = G_{E2}G_B \end{aligned}\right\} \tag{8.9}$$

If $C_{Al} = C_{Cl}$, $G_B = G_F$, $G_{A1} = G_{E2}$, the nominal values of Q and ω_n become

$$\omega_n = \frac{G_{A1}}{(2C_{A1}C_{E2})^{\frac{1}{2}}} \tag{8.10a}$$

$$Q = \left(\frac{C_{E2}}{2C_{A1}}\right)^{\frac{1}{2}} \tag{8.l0b}$$

and the transfer function reduces to

$$T(s) = \frac{G_{A1}^2/2C_{A1}G_{E2}}{s^2 + \frac{G_{A1}}{C_{E2}}s + \frac{G_{A1}^2}{2C_{A1}C_{E2}}} \tag{8.11}$$

From Eqs. 8.10a, b, it is evident that a very high Q may be obtained by appropriate choice of C_{E2} and C_{A1}

The sensitivities of Q and ω_n to the various parameters for this realization are

$$\left.\begin{array}{ll} S^Q_{C_{A1}} = S^Q_{C_{C1}} = -\frac{1}{4}, & S^Q_{CE2} = \frac{1}{2} \\ S^Q_{G_B} = Q^2, & S^Q_{GF} = -Q^2 \\ S^Q_{G_{A1}} = \frac{1}{2} - Q^2, & S^Q_{GE2} = -\frac{1}{2} + Q^2 \end{array}\right\} \tag{8.12}$$

and

$$\left.\begin{array}{ll} S^{\omega_n}_{C_{A1}} = S^{\omega_n}_{C_{C1}} = -\frac{1}{4}, & S^{\omega_n}_{C_{E2}} = -\frac{1}{2} \\ S^{\omega_n}_{G_{A1}} = S^{\omega_n}_{G_{ER}} = \frac{1}{2} & \end{array}\right\} \tag{8.13}$$

8.4 Bandpass Filter Realizations

The bandpass realization of eight elements shown in Fig. 8.2b may be obtained by setting

$$\left.\begin{array}{l} G_{A2} = G_{C2} = G_D = G_{E1} = 0 \\ C_{A2} = C_{C2} = C_{E1} = 0 \end{array}\right\} \tag{8.14}$$

With the admittances

$$\left.\begin{array}{ll} Y_A = C_{A1}s + G_{A1} & Y_B = G_B \\ Y_C = C_{C1}s + G_{C1} & Y_D = 0 \\ Y_E = \frac{G_{E_2}s}{\frac{G_{E2}}{C_{E2}}+s} & Y_F = G_F \end{array}\right\} \tag{8.15}$$

shown in Fig. 8.2b, the coefficients of the transfer function given by Eq. (8.3) and the general expressions for Q and ω_n may be written as

$$\left.\begin{array}{l} a_2 = C_{A1}G_F - C_{C1}G_B \\ a_1 = G_{A1}G_F - G_{C1}G_B - G_BG_{E2} \\ \qquad + \dfrac{G_{E2}}{C_{E2}}(C_{A1}G_F - C_{C1}G_B) \\ a_0 = \frac{G_{E2}}{C_{E2}}(G_{A1}G_F - G_{C1}G_B) \\ b_2 = (C_{A1} + C_{C1})G_F \\ b_1 = (G_{A1} + G_{C1})G_F - G_BG_{E2} + \frac{G_{E2}}{C_{E2}}(C_{A1} + C_{C1})G_F \\ b_0 = \frac{G_{E2}}{C_{E2}}(G_{A1} + G_{C1})G_F \end{array}\right\} \tag{8.16}$$

$$Q = \frac{G_F[(G_{A1} + G_{C1})/(C_{A1} + C_{C1})G_{E2} \cdot C_{E2}]^{\frac{1}{2}}}{G_{E2}(C_{A1} + C_{C1})G_F + C_{E2}(G_{A1}G_F + G_{C1}G_F - G_{E2}G_B)} \tag{8.17}$$

and

$$\omega_n = \left[\frac{G_{E2}(G_{A1} + G_{C1})}{C_{E2}(C_{A1} + C_{C1})}\right]^{1/2} \tag{8.18}$$

Setting $a_2 = 0$ and $a_0 = 0$, we obtain two BP realizations.

8.4.1 First Bandpass Realization

If we make $G_B = G_F$, $G_{A1} = G_{C1}$ and $C_{E2} = 2C_{A1} = 2C_{Cl}$, the nominal values of ω_n and Q become

$$\omega_n = \frac{1}{C_{A1}}\left[\frac{G_{E2}G_{A1}}{2}\right]^{1/2} \tag{8.19}$$

$$Q = \left(\frac{G_{E2}}{2G_{A1}}\right)^{1/2} \tag{8.20}$$

and the transfer function reduces to

$$T(s) = -\frac{(G_{E2}/2C_{A1})s}{s^2 + \frac{G_{A1}}{C_{A1}}s + \frac{G_{E2}G_{A1}}{C_{A1}^2}} \tag{8.21}$$

It may be seen that a very high Q may be obtained by appropriate choice of G_{E2} and G_{A1}.

The Q and ω_n sensitivities for this realization are

$$\left.\begin{aligned} &S_{C_{A1}}^{Q} = S_{C_{C1}}^{Q} = \tfrac{1}{2}\left(\tfrac{1}{2} - Q^2\right), \quad S_{C_{E2}}^{Q} = \left(Q^2 - \tfrac{1}{2}\right) \\ &S_{G_B}^{Q} = Q^2, \quad S_{G_F}^{Q} = -Q^2 \\ &S_{G_{A1}}^{Q} = S_{G_{C1}}^{Q} = -\tfrac{1}{4}, \quad S_{G_{E2}}^{Q} = -\tfrac{1}{2} \end{aligned}\right\} \tag{8.22}$$

$$\left.\begin{aligned} S_{C_{A1}}^{\omega_n} = S_{C_{C1}}^{\omega_n} = -\tfrac{1}{4}, \quad S_{CE2}^{\omega_n} = -\tfrac{1}{2} \\ S_{G_{A1}}^{\omega_n} = S_{G_{C1}}^{\omega_n} = \tfrac{1}{4}, \quad S_{G_{E2}}^{\omega_n} = \tfrac{1}{2} \end{aligned}\right\} \tag{8.23}$$

8.4.2 Second Bandpass Realization

If we set $G_B = G_F$, $G_{E2} = 2G_{A1} = 2G_{C1}$ and $C_{A1} = C_{C1}$, the nominal values of ω_n and Q become

$$\omega_n = G_{A1}\left(\frac{2}{CA_1\, C_{E2}}\right)^{\frac{1}{2}} \tag{8.24}$$

$$\text{and } Q = \left(\frac{C_{E2}}{2C_{A1}}\right) \tag{8.25}$$

and the transfer function reduces to

$$T(s) = -\frac{(G_{A1}/C_{A1})s}{s^2 + \frac{2G_{A1}}{C_{E2}}s + \frac{2G_{A1}^2}{C_{A1}C_{E2}}} \tag{8.26}$$

In this case, a high Q may be obtained by choosing $C_{E2} \gg C_{A1}$.
The Q sensitivities for this realization are

$$\left.\begin{aligned} &S_{C_{A1}}^{Q} = S_{C_{C1}}^{Q} = -\tfrac{1}{4}, \quad S_{C_{E2}}^{Q} = \tfrac{1}{2} \\ &S_{G_B}^{Q} = Q^2, \quad S_{G_F}^{Q} = -Q^2 \\ &S_{G_{A1}}^{Q} = S_{G_{C1}}^{Q} = \tfrac{1}{2}\left(\tfrac{1}{2} - Q^2\right), \quad S_{G_{E2}}^{Q} = Q^2 - \tfrac{1}{2} \end{aligned}\right\} \tag{8.27}$$

The ω_n sensitivities are same as in the case of first bandpass realization.

8.5 High-Pass Realization

To realize a high-pass transfer function, we set in Fig. 8.1

$$\left.\begin{aligned} G_{A2} = G_{C2} = G_D = G_{E1} = 0 \\ C_{A2} = C_{C1} = C_{C2} = C_{E1} = 0 \end{aligned}\right\} \tag{8.28}$$

With the admittances

$$\left.\begin{aligned} Y_A = C_{A1}s + G_{A1}, \quad & Y_B = G_B \\ Y_C = G_{C1}, \quad & Y_D = 0 \\ Y_E = \frac{G_{E2}s}{\frac{G_{E2}}{C_{E2}} + s}, \quad & Y_F = G_F \end{aligned}\right\} \tag{8.29}$$

as shown in Fig. 8.2c, the coefficients of the transfer function given by Eq. (8.3) and the general expressions for ω_n and Q may be obtained as

$$a_2 = C_{A1}G_F = b_2$$

$$\left.\begin{aligned} a_1 &= G_{A1}G_F - G_{C1}G_B - G_{E2}G_B + \tfrac{G_{E2}}{C_{E2}}C_{A1}G_F \\ a_0 &= \tfrac{G_E}{C_{E2}}(G_{A1}G_F - G_{C1}G_B) \\ b_1 &= (G_{A1} + G_{C1})G_F + \tfrac{G_{E2}}{C_{E2}}C_{A1}G_F - G_{E2}G_B \\ b_0 &= \tfrac{G_{E2}}{C_{E2}}(G_{A1} + G_{C1})G_F \end{aligned}\right\} \tag{8.30}$$

$$\omega_n = \left[\frac{G_{E2}(G_{A1} + G_{C1})}{C_{A1}C_{E2}}\right]^{\frac{1}{2}} \tag{8.31}$$

and

$$Q = \frac{G_F[(G_{A1} + G_{C1})G_{E2}C_{A1}C_{E2}]^{\frac{1}{2}}}{G_{E2}G_FC_{A1} + (G_{A1}G_F + G_{C1}G_F - G_{E2}G_B)C_{E2}} \tag{8.32}$$

Setting $a_1 = 0$ and $a_0 = 0$ for high-pass realization, we have C_{A1}= C_{E2}, G_B= G_F and G_{A1}= G_{C1} and nominal values of ω_n and Q become

$$\omega_n = \frac{(2G_{A1}G_{E2})^{1/2}}{C_{A1}} \tag{8.33}$$

$$Q = \left(\frac{G_{E2}}{2G_{A1}}\right)^{1/2} \tag{8.34}$$

The transfer function reduces to

$$T(s) = \frac{s^2}{s^2 + \frac{2G_{A1}}{C_{A1}}s + \frac{2G_{E2}G_{A1}}{C_{A1}^2}} \tag{8.35}$$

Q for this realization can be made very high by making $G_{E2} \gg G_{A1}$.

The sensitivities of Q and ω_n to the various parameters for this realization of seven elements are given by

$$\left.\begin{aligned} S_{C_{A1}}^{Q} &= \tfrac{1}{2} - Q^2, & S_{C_{E2}}^{Q} &= Q^2 - \tfrac{1}{2} \\ S_{G_B}^{Q} &= Q^2, & S_{G_F}^{Q} &= -Q^2 \\ S_{G_{A1}}^{Q} &= S_{G_{C1}}^{Q} = -\tfrac{1}{4}, & S_{G_{E2}}^{Q} &= \tfrac{1}{2} \end{aligned}\right\} \tag{8.36}$$

and

$$\left.\begin{aligned} S_{C_{A1}}^{\omega_n} &= S_{C_{E2}}^{\omega_n} = -\tfrac{1}{2} \\ S_{G_{E2}}^{\omega_n} &= \tfrac{1}{2}, & S_{G_{A1}}^{\omega_n} &= S_{G_{C1}}^{\omega_n} = \tfrac{1}{4} \end{aligned}\right\} \tag{8.37}$$

8.6 Band Stop Realization

Three band stop realizations have been obtained for the zeroes lying on imaginary axis at $s = \pm j\omega_n$, $s > \pm j\omega_n$ and $s < \pm j\omega_n$.

First Band Stop Realization with Zeroes at $s = \pm j\omega_n$

The band stop realization of eight elements shown in Fig. 8.2d may be obtained by setting

$$\left.\begin{aligned} G_{A2} = G_{C1} = G_D = G_{E1} = 0 \\ C_{A2} = C_{C1} = C_{E1} = 0 \end{aligned}\right\} \tag{8.38}$$

With the admittances

$$\left.\begin{aligned} Y_A &= C_{A1}s + G_{A1}, & Y_B &= G_B \\ Y_C &= \frac{G_{C2}s}{\frac{G_{C2}}{C_{C2}}+s}, & Y_D &= 0 \\ Y_E &= \frac{G_{E2}s}{\frac{G_{E2}}{C_{E2}}+s}, & Y_F &= G_F \end{aligned}\right\} \tag{8.39}$$

shown in Fig. 8.2d, the transfer function given by Eq. (8.2) may be written as

$$T(s) = \frac{s^3 + a_2 s^2 + a_1 s + a_0}{s^3 + \delta_2 s^2 + \delta_1 s + \delta_0} \tag{8.40}$$

where

$$\left.\begin{aligned} a_2 &= \left[\left(\frac{G_{C2}}{C_{C2}} + \frac{G_{E2}}{C_{E2}}\right) + \frac{G_{A1}}{C_{A1}} - \frac{G_B(G_{C2}+G_{E2})}{C_{A1}G_F}\right] \\ a_1 &= \left[\frac{G_{C2}}{C_{C2}} \cdot \frac{G_{E2}}{C_{E2}} + \left(\frac{G_{E2}}{C_{E2}} + \frac{G_{C2}}{C_{C2}}\right)\frac{G_{A1}}{C_{A1}} - \frac{G_B G_{C2} G_{E2}}{C_{A1}G_F}\left(\frac{1}{C_{C2}} + \frac{1}{C_{E2}}\right)\right] \\ a_0 &= \frac{G_{A1}G_{C2}G_{E2}}{C_{A1}C_{C2}C_{E2}} \\ \delta_2 &= \left[\left(\frac{G_{C2}}{C_{C2}} + \frac{G_{E2}}{C_{E2}}\right) + \frac{(G_{A1}+G_{C2})}{C_{A1}} - \frac{G_B G_{E2}}{C_{A1}G_F}\right] \\ \delta_1 &= \left[\left(\frac{G_{C2}}{C_{C2}} + \frac{G_{E2}}{C_{E2}}\right)\frac{G_{A1}}{C_{A1}} + \left(\frac{C_{A1}}{C_{C2}} + 1\right)\frac{G_{E2}}{C_{E2}} \cdot \frac{G_{C2}}{C_{A1}} - \frac{G_B G_{C2} G_{E2}}{C_{A1}C_{C2}G_F}\right] \\ \delta_0 &= \frac{G_{A1}}{C_{A1}} \cdot \frac{G_{C2}}{C_{C2}} \cdot \frac{G_{E2}}{C_{E2}} \end{aligned}\right\} \tag{8.41}$$

If we set

$$\left.\begin{array}{l} \frac{G_{E2}}{C_{E2}} = \frac{G_{C2}}{C_{C2}} = \alpha \\ G_{A1} = G_{C2},\ G_B = G_F \text{ and } C_{A1} = C_{E^2} \end{array}\right\} \tag{8.42}$$

the nominal values of coefficients δ_2, δ_1 and δ_0 and nominal values of ω_n and Q become

$$\left.\begin{array}{l} \delta_2 = \alpha + 2\frac{G_{A1}}{C_{A1}},\ \delta_1 = 3\alpha\frac{G_{A1}}{C_{A1}} \\ \delta_0 = \alpha^2\frac{G_{A1}}{C_{A1}} \end{array}\right\} \tag{8.43}$$

$$\omega_n = \frac{(G_{E2}G_{A1})^{\frac{1}{2}}}{C_{A1}} \tag{8.44}$$

$$Q = \frac{1}{2}\left(\frac{G_{E2}}{G_{A1}}\right)^{\frac{1}{2}} \tag{8.45}$$

and the transfer function reduces to

$$T(s) = \frac{s^2 + \alpha\frac{G_{A1}}{C_{A1}}}{s^2 + \frac{2G_{A1}}{C_{A1}}s + \alpha\frac{G_{A1}}{C_{A1}}} \tag{8.46}$$

This transfer function is of the form

$$T(s) = \frac{s^2 + a_0'}{s^2 + b_1's + b_0'} \tag{8.47}$$

where $a_0' = b_0'$, i.e. zeroes of transfer function are at $s = \pm j\omega_n$

In [4], Soderstrand and Mitra derived expressions relating to Q and ω_n, sensitivities of a third-order filter to the denominator polynomial coefficient sensitivities. Following them, the Q and ω_n sensitivities for a third-order system

$$T(s) = \frac{(s + r)(s^2 + a_1''s + a_0'')}{s^3 + \delta_2 s^2 + \delta_1 s + \delta_0} \tag{8.48}$$

may be expressed as

$$S_{x_i}^{Q} = \frac{[\delta_0(2\omega_n Q^2 + rQ - \omega_n)S_{x_i}^{\delta_0} + \delta_1(\omega_n^2 Q - 2\omega_n r Q^2)S_{x_i}^{\delta_i} + \omega_n^2(Qr - 2\omega_n Q^2)\delta_2 S_{x_i}^{\delta_2}]}{2\omega_n^2[(\omega_n^2 + r^2)Q - \omega_n r]} \tag{8.49}$$

$$S_{x_i}^{\omega_n} = \frac{(Qr - \omega_n)\delta_0 S_{x_i}^{\delta_0} + Q\omega_n^2\delta_1 S_{x_i}^{\delta_i} - Q\omega_n^2 r\delta_2 S_{x_i}^{\delta_2}}{2\omega_n^2[(\omega_n^2 + r^2)Q - \omega_n r]} \tag{8.50}$$

where

$S_{x_i}^{\delta_j}$ = denominator polynominal coefficient sensitivity with respect to an element x_i.

$$= \frac{\partial \delta_j / \delta_j}{\partial x_i / x_i} \tag{8.51}$$

and the coefficients δ_2, δ_1 and δ_0 are related to ω_n, Q and r as follows:

$$\left.\begin{aligned} \delta_2 &= (\omega_n/Q) + r \\ \delta_1 &= \omega_n^2 + (\omega_n/Q)r \\ \delta_0 &= \omega_n^2 r \end{aligned}\right\} \tag{8.52}$$

From Eqs. (8.43), (8.44), (8.45) and (8.52), we get

$$r = \alpha \tag{8.53}$$

Substituting the nominal values of coefficients δ_2, δ_1 and δ_0 and r as given by Eqs. (8.43) and (8.53) in Eqs. (8.49) and (8.50), we obtain

$$S_{x_i}^{Q} = S_{x_i}^{\delta_0} - (3/2)\mathrm{S}_{x_i}^{\delta_1} \tag{8.54}$$

and

$$S_{x_i}^{\omega_n} = \frac{(2Q^2-1)S_{x_i}^{\delta_0} + (3/2)S_{x_i}^{\delta_1} - (2Q^2+1)S_{x_i}^{\delta_2}}{(4Q^2-1)} \tag{8.55}$$

Thus Q and ω_n sensitivities with respect to elements of the circuit may be computed as

$$\left.\begin{aligned} S_{C_{A1}}^{Q} &= \tfrac{1}{2} - 2Q^2, \; S_{C_{C2}}^{Q} = -\tfrac{1}{2} \\ S_{C_{E2}}^{Q} &= 2Q^2, \qquad S_{G_{A1}}^{Q} = S_{G_{C2}}^{Q} S_{G_{E2}}^{Q} = 0 \\ S_{G_B}^{Q} &= +2Q^2, \qquad S_{G_F}^{Q} = -2Q^2 \end{aligned}\right\} \tag{8.56}$$

and

$$\left.\begin{aligned} S_{C_{A1}}^{\omega_n} &= -\tfrac{1}{2}, \quad S_{C_{C2}}^{\omega_n} = \tfrac{1}{2(4Q^2-1)} \\ S_{C_{E2}}^{\omega_n} &= -\tfrac{2Q^2}{(4Q^2-1)}, \quad S_{G_{A1}}^{\omega_n} = \tfrac{1}{2} \\ S_{G_{C2}}^{\omega_n} &= -\tfrac{(2Q^2+1)}{(4Q^2-1)}, \quad S_{G_{E2}}^{\omega_n} = \tfrac{2Q^2}{(4Q^2-1)} \\ S_{G_B}^{\omega_n} &= S_{G_F}^{\omega_n} = 0 \end{aligned}\right\} \tag{8.57}$$

8.6.1 Second Band Stop Realization

The low-pass filter configuration shown in Fig. 8.2a may also realize a band stop action with modified constraints on the elements. Relations (8.4)–(8.8) can be directly used for band stop realization as well.

By setting $a_1 = 0$, the following constraints on the elements of band stop filter may be obtained:

$$\left.\begin{array}{l} C_{A1}G_{A1} = G_{E2}C_{C1} \\ G_B = G_F,\ C_{A1} = C_{E2} \end{array}\right\} \tag{8.58}$$

The transfer function to given by Eq. (8.2) is reduced

$$T(s) = K\frac{s^2 + a_0'}{s^2 + b_1's + b_0'} \tag{8.59}$$

where

$$\left.\begin{array}{l} K = \frac{(1-G_{A1}/G_{E2})}{(1+G_{A1}/G_{E2})} < 1 \\ a_0' = \frac{G_{A1}G_{E2}}{C_{A1}^2[1-(G_{A1}/G_{E2})]} \\ b_0' = \frac{G_{A1}G_{E2}}{C_{A1}^2[1+(G_{A1}/G_{E2})]} \\ b_1' = \frac{2G_{A1}}{C_{A1}[1+(G_{A1}/G_{E2})]} \end{array}\right\} \tag{8.60}$$

$$a_0' > 0, a_0' > b_0', G_{E2} > G_{A1}$$

Thus, the zeroes of the transfer function are at $s > \pm j\omega_n$. The nominal values of ω_n and Q are given by

$$\omega_n = \frac{G_{E2}}{C_{A1}}\left[\frac{G_{A1}}{(G_{A1} + G_{E2})}\right]^{\frac{1}{2}} \tag{8.61}$$

and

$$Q = \frac{1}{2}\left(1 + \frac{G_{E2}}{G_{A1}}\right)^{\frac{1}{2}} \tag{8.62}$$

If $G_{E2} \gg G_{A1}$, a very high Q may be obtained.

Using general expressions given by Eqs. (8.7) and (8.8) for ω_n and Q and the constraint Eq. (8.58), Q and ω_n sensitivities for this realization may be obtained as

$$\left.\begin{array}{ll} S_{C_{A1}}^{Q} = 1 - \frac{1}{8Q^2} - 2Q^2, & S_{C_{C1}}^{Q} = -\frac{1}{2} + \frac{1}{8Q^2} \\ S_{C_{E2}}^{Q} = 2Q^2 - \frac{1}{2}, & S_{G_B}^{Q} = 2Q^2 - \frac{1}{2} \\ S_{G_F}^{Q} = -2Q^2 + \frac{1}{2}, & S_{G_{A1}}^{Q} = S_{G_{E2}}^{Q} = 0 \end{array}\right\} \tag{8.63}$$

and

$$\left.\begin{array}{l} S_{G_{A1}}^{\omega_n} = S_{G_{E2}}^{\omega_n} = \frac{1}{2} \\ S_{C_{E2}}^{\omega_n} = -\frac{1}{2}, \quad S_{C_{A1}}^{\omega_n} = -\frac{G_{E2}}{2(G_{A1}+G_{E2})} \\ S_{C_{C1}}^{\omega_n} = -\frac{G_{A1}}{2(G_{A1}+G_{E2})} \end{array}\right\} \tag{8.64}$$

8.6.2 Third Band Stop Realization

The high-pass filter configuration shown in Fig. 8.2c may also realize a band stop filter with modified constraints on the elements. Expressions given by Eqs. (8.28)–(8.32) can be directly carried over for band stop realization.

By setting $a_1 = 0$ in Eq. (8.30), the following constraints on the elements of band stop filter may be obtained:

$$\left.\begin{array}{l} G_{C1}C_{E2} = G_{E2}C_{A1} \\ G_B = G_F, G_{A1} = G_{E2} \end{array}\right\} \tag{8.65}$$

The transfer function for this realization is thus reduced to

$$T(s) = \frac{s^2 + a_0'}{s^2 + b_1' s + b_0} \tag{8.66}$$

where

$$\left.\begin{array}{l} a_0' = \frac{G_{A1}^2}{C_{A1}C_{E2}}\left(1 - \frac{G_{C1}}{G_{A1}}\right) \\ b_0' = \frac{G_{A1}^2}{C_{A1}C_{E2}}\left(1 + \frac{G_{C1}}{G_{A1}}\right) \\ b_1' = \frac{2G_{C1}}{C_{A1}} \\ a_0' < b_0' \text{ and } a_0 > 0 \\ \text{for} \\ \frac{G_{C1}}{G_{A1}} < 1 \end{array}\right\} \tag{8.67}$$

In this case, the zeroes of transfer function are at $s <\pm j\omega_n$.

The nominal values of ω_n and Q are given by

$$\omega_n = \left[\frac{G_{A1}(G_{A1} + G_{C1})}{C_{A1}C_{E2}}\right]^{\frac{1}{2}} \tag{8.68}$$

and

$$Q = \frac{1}{2}\left(1 + \frac{G_{A1}}{G_{C1}}\right)^{\frac{1}{2}} \tag{8.69}$$

If $G_{A1} \gg G_{C1}$, a very high Q may be obtained.

Using the general expressions given by Eqs. (8.31) and (8.32) for ω_n and Q and the constraints of Eqs. (8.65), Q and ω_n sensitivities may be obtained as

$$\left.\begin{array}{l} S_{C_{A1}}^{Q} = 0, \quad S_{C_{E2}}^{Q} = 0 \\ S_{G_B}^{Q} = 2Q^2 - \frac{1}{2}, \quad S_{G_F}^{Q} = \frac{1}{2} - 2Q^2 \\ S_{G_{A1}}^{Q} = 1 - \frac{1}{8Q^2} - 2Q^2 \\ S_{G_{C1}}^{Q} = -\frac{1}{2} + \frac{1}{8Q^2}, \quad S_{G_{E2}}^{Q} = 2Q^2 - \frac{1}{2} \end{array}\right\} \tag{8.70}$$

and

$$\left.\begin{array}{l} S_{C_{A1}}^{\omega_n} = S_{C_{E2}}^{\omega_n} = -\frac{1}{2} \\ S_{G_{E2}}^{\omega_n} = \frac{1}{2}, \quad S_{G_{A1}}^{\omega_n} = \frac{G_{A1}}{2(G_{A1}+G_{C1})} \\ S_{G_{C1}}^{\omega_n} = \frac{G_{C1}}{2(G_{A1}+G_{C1})} \end{array}\right\} \tag{8.71}$$

8.7 Economic Aspect of Filter Realizations

A few filter realizations described here are such that by addition or deletion of one or two elements one may obtain another realization having Q of the same order.

With the circuit configuration shown in Fig. 8.3a, second-order low pass filter realization with specifications may be obtained by short circuiting one resistor element G_{A1} of G_{E2} and open circuiting or blowing off the resistor G_{C1}. Bandpass filter having the same Q and the following other specifications may be obtained by open circuiting or blowing off the short circuit across G_{A1}.

$$\left.\begin{array}{l} \omega_n = \dfrac{G_{A1}}{(2C_{A1}C_{E2})^{\frac{1}{2}}}, \quad Q = \left(\dfrac{C_{E2}}{2C_{A1}}\right)^{\frac{1}{2}} \\ C_{A1} = C_{C1}, G_B = G_F \text{ and } G_{A1} = G_{E2} \end{array}\right\} \tag{8.72}$$

Bandpass filter having the same Q and the following other specification

$$\left.\begin{array}{l} \omega_n = G_{A1}\left(\dfrac{2}{C_{A1}C_{E2}}\right)^{\frac{1}{2}} \\ C_{A1} = C_{C1}, G_{E2} = 2G_{A1} = 2G_{C1} \end{array}\right\} \tag{8.73}$$

and

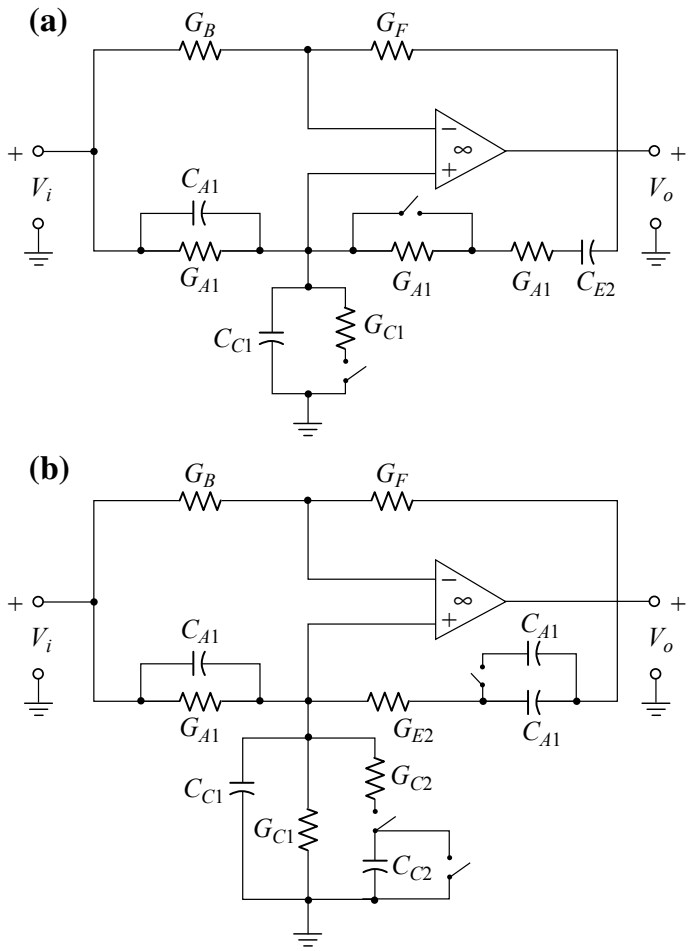


Fig. 8.3 **a** Circuit for low-pass or bandpass realization ($G_{E2} = G_{A1}$ or $2G_{A1}$). **b** Circuit for high-pass, bandpass or band stop filter ($a_0 = b_0'$) realizations with $C_{E2} = C_{A1}$ (or $2C_{A1}$, $G_{C1} = G_{C2} = G_{A1}$)

$$G_B = G_F$$

maybe obtained by open circuiting or blowing off the short circuit across E_{A1}.

Similarly, with another circuit configuration shown in Fig. 8.3b, high-pass realization, first bandpass and first band stop realization discussed earlier, may be obtained by appropriately short circuiting and open circuiting of elements.

Realization of various filter circuits from one common configuration thus reduces the cost.

8.8 Conclusion

It has been shown that a second-order filter having a high Q may be realized with seven or eight elements by employing a differential input operational amplifier. From the same circuit configuration, other filter realizations having Q of the same order may be obtained, thereby reducing the manufacturing cost. It has been observed that with these realizations, Q sensitivities are of high value and in certain cases, proportional to Q^2. Further, the Q sensitivity is not much affected by addition of other elements,

for example, series RC impedances of Fig. 8.1. The realizations would, therefore, be useful in low and moderate Q applications.

References

1. S.C. Dutta Roy, Active RC filters using a single differential-input operational amplifier. J. Instn Telecom. Engrs, **18**, 273–275 (1972)
2. S.K. Mitra, Active RC filters employing a single operational amplifier as the active element, in *Proceedings of Hawaii International Conference on System Sciences* (Honolulu, Hawaii, 1968)
3. S.K. Mitra, *Analysis and Synthesis of Linear Active Networks* (John Wiley & Sons Inc., New York, 1969)
4. M.A. Soderstrand, S.K. Mitra, Sensitivity analysis of third-order filters, in *Mexico 1971 IEEE International Conference on Circuits, Systems and Computers*, pp. 345–349 (1971)

Chapter 9
Tolerance Minded Design of Active RC Bandpass Filters

Abstract The chapter deals with a simple transformation which allows the result of Zeheb for tolerance-minded design of active RC low-pass filters to be conveniently extended to active RC bandpass filters.

Keywords Tolerance minded design · Active RC bandpass filter

9.1 Introduction

In his 1970 paper, Zeheb [1], suggested a new approach for the tolerance minded design of Butterworth and Chebyshev low-pass and high-pass filters, which are realized as RC active networks. It is shown here that with suitable transformations, the results of Zeheb can be extended for the tolerance minded design of active RC bandpass filters.

Whereas for LC networks, transformation from low-pass to bandpass filters are well known in the literature, no such transformation appears to be generally available for filters which are realized as RC networks. A simple and convenient transformation for this purpose is presented here, based on the approximations used in the theory of stagger tuned amplifiers and commonly applied to LC bandpass filters [2]. This development presupposes that the frequencies under consideration are relatively large and that the associated bandwidth covers a narrow range.

The underlying transformation is from $s = \sigma + j\omega$ to $p = \Sigma + j\Omega$, where $\Sigma = \sigma/(\omega_p - \omega_0)$, $\Omega = (\omega - \omega_0)/(\omega_p - \omega_0)$ and ω_0 is the algebraic mean of the upper and lower cut-off frequencies ω_p and $2\omega_0 - \omega_p$, respectively (defining ω_0 in this fashion provides symmetrical response specifications on either side of ω_0). It is seen that as a result of this transformation, practically all results of Zeheb for low-pass filter carry over to bandpass case by simply replacing Zeheb's $X = \omega/\omega_1$ and $\mu = \omega_2/\omega_1$ by Ω and the quantity $(\omega_2 - \omega_0)/(\omega_p - \omega_0)$, respectively, where ω_2 is the 'critical' frequency in the stopband.

Source: M.C. Srivastava, S.C. Dutta Roy & P.S. Satsangi, "Tolerance Minded Design of Active RC Band Pass Filters," *JIETE*, vol 18, pp 86–88, February 1972.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_9

9.2 Butterworth Approximation

The transfer function of n units of cascaded second-order bandpass filters can be expressed as

$$T(s) = \frac{s^n}{(s - s_1)(s - s_1^*)(s - s_2)(s - s_2^*) \ldots (s - s_n)(s - s_n^*)} \tag{9.1}$$

where asterisk denotes conjugate, and $s_i = \sigma_i + j\omega_i$, $i = 1, 2, \ldots n$.

As an example, the locations of poles and zeroes of a high Q bandpass filter with $n = 4$ are shown in Fig. 9.1 together with a phasor representation of Eq. (9.1) for $s = j\omega$. Note that for a high Q, $|\sigma_i| \ll |\omega_i|$, i.e. the poles are located close to the imaginary axis.

Let the frequencies under consideration be relatively large and the associated bandwidth cover a narrow range. Under such an assumption, the effect of single or multiple zeroes at the origin and the poles in the lower half s plane on the transfer function are insignificant relative to the effects of the poles in the upper half-plane in the proximity of the narrow range of frequency under consideration. Thus, the transfer function of the bandpass filter given by Eq. (9.1) can be approximated for $s = j\omega$ as

$$T(s) \cong \frac{k_0}{(j\omega - s_1)(j\omega - s_2) \ldots (j\omega - s_n)} \tag{9.2}$$

where

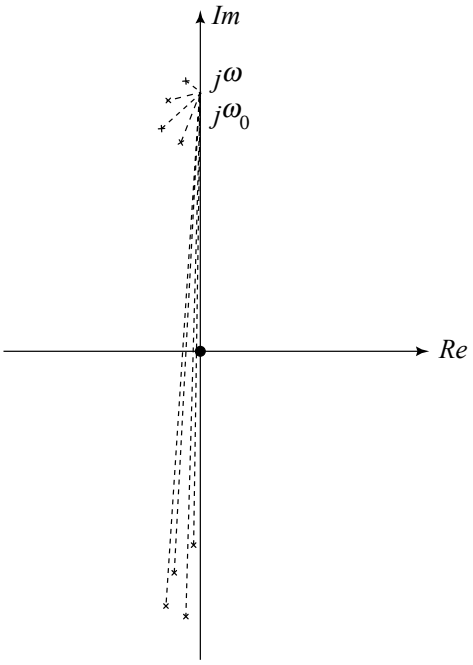


Fig. 9.1 Pole zero locations of a high Q bandpass filter with $n = 4$

$$k_0 = \frac{(j\omega_0)^n}{(j\omega_0 - s_1^*)(j\omega_0 - s_2^*)\ldots(j\omega_0 - s_n^*)} \tag{9.3}$$

and ω_0 is the algebraic mean of the upper and lower cut-off frequencies ω_p and $2\omega_0 - \omega_p$, respectively (defining ω_0 in this fashion provides symmetrical response specifications on either side of ω_0).

Let us make the response maximally flat at $\omega = \omega_0$ for Butterworth transfer function. To achieve this, we change the frequency variable from s to p such that

$$p = \frac{\sigma + j\bar{\omega}}{\omega_p - \omega_0} = \Sigma + j\Omega \tag{9.4a}$$

where

$$\bar{\omega} = \omega - \omega_0 \tag{9.4b}$$

This gives a change in the scale and a translation along the imaginary axis, such that when $s = 0 + j\omega_0, p = 0$ and the poles of the transfer function given by Eq. (9.1) are transformed to

$$p_i = \frac{\sigma_i + j(\omega_i - \omega_0)}{\omega_p - \omega_0}, \quad i = 1, 2, \ldots n \tag{9.5}$$

in the p-plane. The transfer function can now be approximated as

$$T(p) \cong \frac{k_0'}{(p - p_1)(p - p_2)\ldots(p - p_n)} \tag{9.6}$$

where

$$k_0' = \frac{k_0}{(\omega - \omega_0)^n} \tag{9.7}$$

For sinusoidal frequencies, the magnitude of this transfer function is of the form

$$|T(j\Omega)| = \frac{k'}{(\Omega^{2n} + A_1\Omega^{2n-1} + A_2\Omega^{2n-2} + \cdots + A_{2n-1}\Omega + A_{2n})^{\frac{1}{2}}} \tag{9.8}$$

where

$$k' = \left|k_0'\right|.$$

For maximally flat response at $\Omega = 0$ (i.e. $\omega = \omega_0$), $A_1 = A_2 = \ldots = A_{2n-1} = 0$, and hence

$$|T(j\Omega)| = \frac{k}{(1+\lambda\Omega^{2n})^{\frac{1}{2}}} \tag{9.9}$$

where

$$\lambda = \frac{1}{A_{2n}} = \left|\frac{1}{p_1 p_2 \ldots p_n}\right|$$

and

$$k = \frac{\omega_0^n/(A_{2n})^{\frac{1}{2}}}{\left[\{(\omega_0+\omega_1)^2+\sigma_1^2\}\ldots\{(\omega_0+\omega_n)^2+\sigma_n^2\}\right]^{\frac{1}{2}}(\omega_p-\omega_0)^n} \tag{9.10}$$

The magnitude function (attenuation) $\alpha(\bar{\omega})$ of the bandpass filter, having a narrow bandwidth remote from $\omega = 0$, may now be expressed as

$$\alpha(\bar{\omega}) = \ln(1+\lambda\Omega^{2n})^{\frac{1}{2}}$$

or

$$\exp[2\alpha(\bar{\omega})] = 1+\lambda\Omega^{2n} \tag{9.11}$$

Following Zeheb, for a first-order approximation, we have

$$\Delta\alpha(\bar{\omega}) = K\bar{\omega}\frac{\partial\alpha(\bar{\omega})}{\partial\bar{\omega}} \tag{9.12}$$

where

$$K = \frac{\Delta R}{R} + \frac{\Delta C}{C} \tag{9.13}$$

$\Delta R/R$ and $\Delta C/C$ being the extreme relative deviation from nominal value, overall resistors and all capacitors, respectively. $\Delta\alpha(\bar{\omega})$ is the deviation of $\alpha(\bar{\omega})$ from the theoretical magnitude which is arrived at if nominal value elements are used and no variations occur.

Equation (9.12) will now be applied to specific approximations for bandpass filter and deviation of α from the specifications will be expressed as an explicit function of K. Let the filter characteristic be approximated by delete this space and continue with Eq. (9.11).

Equation (9.11) and filter requirements be

$$\alpha(\bar{\omega}) \le \alpha_p \text{ for } \omega_0 - \omega_p \le \bar{\omega} \le \omega_p - \omega_0 \tag{9.14a}$$

and

$$\alpha(\bar{\omega}) \geq \alpha_a \text{ for } \bar{\omega}_2 \leq \bar{\omega} < \infty \text{ and } -\infty < \bar{\omega} \leq -\bar{\omega}_2 \tag{9.14b}$$

where ω_2 (ω_2 in $\bar{\omega}$ scale) is the 'critical' frequency in the stopband, α_a, α_p are in nepers ($\alpha_a \gg \alpha_p$) and ω_p, ω_0, ω_2, $\bar{\omega}$, $\bar{\omega}_2$ are in rad/sec.

From Eqs. (9.11) and (9.14a, b), we have

$$\alpha(\Omega)|_{\Omega=\pm 1} = \alpha_p \tag{9.15}$$

and

$$\alpha(\Omega)|_{\Omega=\pm\bar{\omega}_1/(\omega_p-\omega_0)} \geq \alpha_n \tag{9.16}$$

from which the constants λ and n of Eq. (9.11) be evaluated as

$$\lambda = \exp(2\alpha_p) - 1 \tag{9.17}$$

and

$$n = \frac{\left[\ln\{\exp(2\alpha_a) - 1\} - \ln\{\exp(2\alpha_p) - 1\}\right]}{2\ln\left[\bar{\omega}_2/(\omega_p - \omega_0)\right]} \cong \frac{2\alpha_a - \ln\{\exp(2\alpha_p) - 1\}}{2\ln\mu} \tag{9.18}$$

where

$$\mu = \frac{\bar{\omega}_2}{\omega_p - \omega_0} = \frac{\omega_2 - \omega_0}{\omega_p - \omega_0} \tag{9.19}$$

Using Eqs. (9.11)–(9.13), we have

$$\Delta\alpha(\Omega) = K\lambda n \frac{\Omega^{2n}}{1 + \lambda\Omega^{2n}} \tag{9.20}$$

Equation (9.20) is similar to the one derived by Zeheb for low-pass filter, with his $X = \omega/\omega_1$ replaced by Ω in our derivation. Further, it is seen that all the results and the design procedure of Zeheb for low-pass filter carry over to bandpass case by simply replacing Zeheb's X and $\mu = \omega_2/\omega_1$ by Ω and the quantity $(\omega_2 - \omega_0)/\left(\omega_p - \omega_0\right)$ respectively.

9.3 Conclusion

In this chapter, it has been shown that with suitable transformations, the results of Zeheb for the design of active RC low-pass Butterworth filter can be extended to

active RC bandpass filters. In a similar manner, one can obtain an expression for Chebyshev approximation of bandpass characteristics and extend Zeheb's results to this case also.

References

1. E. Zeheb, Int. J. Electr. **29**, 249 (1970)
2. W.R. Lepage, S. Seely, *General Network Analysis* (McGraw-Hill Book Co., Inc., New York, 1952)

Chapter 10
Active RC Filters Using a Single Differential Input Operational Amplifier

Abstract A procedure, alternative to that proposed by Mitra, is given for realizing any real rational transfer function by using RC one ports and a single differential input operational amplifier. It is shown, with the help of examples, that the new procedure admits of much more flexibility in the choice of element values than Mitra's method.

Keywords Active RC filter · Differential input operational amplifier · Mitra's method

10.1 Introduction

The commercial availability of high-quality differential input operational amplifiers in the integrated form makes it profitable to develop synthesis procedures using this device as the active element. Many such techniques have appeared in the recent literature, for both driving point and transfer functions [1, 2]. The present chapter is concerned with the circuit of Fig. 10.1, which was proposed by Mitra [3] for realizing any arbitrary real rational transfer function. Assuming the operational amplifier to be ideal, analysis of this circuit yields

$$T(s) = \frac{V_0}{V_i} = \frac{Y_A(Y_B + Y_D + Y_F) - Y_B(Y_A + Y_C + Y_E)}{Y_F(Y_A + Y_C + Y_E) - Y_E(Y_B + Y_D + Y_F)} \tag{10.1}$$

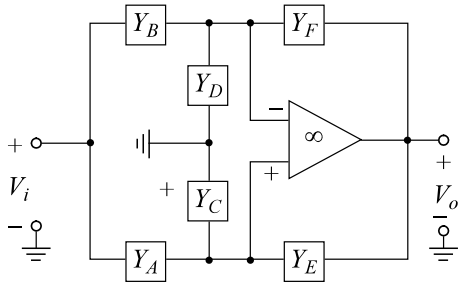


Fig. 10.1 Mitra's circuit for realizing an arbitrary real rational transfer function [Y_A, …, Y_F are RC one ports]

Source: S.C. Dutta Roy, "Active RC Filters Using a Single Differential Input Operational Amplifier," *JITE*, vol 18, pp 273–275 June 1972.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_10

In order to be able to express both numerator and denominator of Eq. (10.1) as the difference between two RC admittances, Mitra assumed [3].

$$Y_A + Y_C + Y_E = Y_B + Y_D + Y_F \tag{10.2}$$

The synthesis of a specified function

$$T(s) = N(s)/D(s) \tag{10.3}$$

starts by choosing a polynomial $Q(s)$ having distinct negative real roots and of degree satisfying

$$\text{deg } Q(s) + 1 \geq \text{higher of}[\text{deg } N(s), \text{deg } D(s)] \tag{10.4}$$

The next step is to make partial fraction expansions of $[N(s), D(s), D(s) - N(s)]/[sQ(s)]$; from these, the various admittances are identified as

$$\left.\begin{aligned} N(s)/Q(s) &= Y_A - Y_B \\ D(s)/Q(s) &= Y_F - Y_E \\ [D(s) - N(s)]/Q(s) &= Y_C - Y_D \end{aligned}\right\} \tag{10.5}$$

It can be shown that such identification of RC one ports satisfies the starting assumption given by Eq. (10.2).

We present here an alternative procedure for synthesizing any real rational $T(s)$ by the circuit of Fig. 10.1 without the need for assuming Eq. (10.2). The major advantage of the new procedure is shown to be a greater flexibility in the choice of element values, as compared to Mitra's method.

10.2 The New Method

We assume the upper tee of admittances in Fig. 10.1 to obey the relations

$$Y_D/Y_B = a \quad \text{and} \quad Y_F/Y_B = b \tag{10.6}$$

where a and b are arbitrary constants. In principle, Y_B may be an arbitrary RC admittance but, for practical convenience, we may choose Y_B to be a suitable conductance. Combining Eqs. (10.1) and (10.6), we get

$$T(s) = \frac{(a+b)Y_A - (Y_C + Y_E)}{b(Y_A + Y_C) - (1+a)Y_E} \tag{10.7}$$

Let the specified $T(s)$ be of the form of Eqs. (10.3) and let $Q(s)$ be chosen as in Mitra's method. Then we can write

$$N(s)/Q(s) = Y_1 - Y_2 \text{ and } D(s)/Q(s) = Y_3 - Y_4 \tag{10.8}$$

where $Y_1, \ldots, Y_4$ are realizable RC admittances. From Eqs. (10.7) and (10.8), we get the equations

$$(a+b)Y_A - Y_C - Y_E = Y_1 - Y_2 \tag{10.9}$$

and

$$Y_A + Y_C - [(1+a)/b]Y_E = Y_3/b - Y_4/b \tag{10.10}$$

Adding Eqs. (10.9) and (10.10) gives

$$Y_A - Y_E/b = [1/(1+a+b)][(Y_3/b + Y_1) - (Y_4/b + Y_2)] \tag{10.11}$$

Let

$$Y_A = [1/(1+a+b)][(1+m)Y_1 + (1+n)Y_3/b + pY_2 + qY_4/b] \tag{10.12}$$

Then Y_E is given by

$$Y_E = [b/(1+a+b)][mY_1 + nY_3/b + (1+p)Y_2 + (1+q)Y_4/b] \tag{10.13}$$

where m, n, p and q are constants. From Eqs. (10.9), (10.12) and (10.13), Y_C is obtained as

$$Y_C = [1/(1+a+b)][(ma-1)Y_1 + (ap+a+1)Y_2 + (an+a+b)Y_3/b + (aq-b)Y_4/b] \tag{10.14}$$

From Eqs. (10.12), (10.13) and (10.14), we see that *sufficient* conditions for RC realizability of Y_A, Y_C and Y_E are

$$p \geq 0, n \geq 0, m \geq 1/a \text{ and } q \geq b/a \tag{10.15}$$

That these conditions are not necessary will be shown with the help of an example in the last section. To simplify matters, we let conditions Eq. (10.15) to be satisfied with the equality signs. Then we get

$$\left.\begin{aligned} Y_A &= [1/(1+a+b)][(1+1/a)Y_1 + Y_3/b + Y_4/a] \\ Y_C &= [1/(1+a+b)][(1+a)Y_2 + (1+a/b)Y_3] \\ Y_E &= [1/(1+a+b)][Y_1/a + Y_2 + (1/b+1/a)Y_4] \end{aligned}\right\} \tag{10.16}$$

Note that the factor $1/(1+a+b)$ has the function of scaling the admittances. Any non-zero finite values of a and b can be chosen to complete the synthesis. For

example, if $a = b = 1$, then Eq. (10.16) become

$$Y_{\mathrm{A}} = (2Y_1 + Y_3 + Y_4)/3,\ Y_C = 2(Y_2 + Y_3)/3$$

and

$$Y_E = (Y_1 + Y_2 + 2Y_4)/3 \tag{10.17}$$

The examples worked out in the next section are based on the above identifications.

10.3 Two Examples

Example 1 Consider the synthesis of a notch filter characterized by

$$T(s) = N(s)/D(s) = \left(s^2 + 2.0\right)/\left(s^2 + 0.1s + 1.2\right) \tag{10.18}$$

Choosing $Q(s) = (s + 1)$ and making partial fraction expansions of $N(s)/[sQ(s)]$ and $D(s)/[sQ(s)]$, we identify

$$Y_1 = s + 2,\ Y_2 = 3s/(s + 1),\ Y_3 = s + 1.2 \text{ and } Y_4 = 2.1s/(s + 1) \tag{10.19}$$

Applying Eq. (10.17), we get

$$\begin{aligned} &Y_A = s + 1.733 + 0.7s/(s + 1),\ Y_C = 2s/3 + 0.8 + 2s/(s + 1) \text{ and} \\ &Y_E = s/3 + 2/3 + 2.4s/(s + 1) \end{aligned} \tag{10.20}$$

The synthesis of these one ports is elementary and the complete network is shown in Fig. 10.2a.

Choosing the same $Q(s)$ and applying Mitra's method, we get the network shown in Fig. 10.2b.

Example 2 Consider the all-pass function

$$T(s) = N(s)/D(s) = \left(-s^3 + 2s^2 - 2s + 1\right)/\left(s^3 + 2s^3 + 2s + 1\right) \tag{10.21}$$

Let

$$Q(s) = (s + 1)(s + 2) \tag{10.22}$$

Then by the usual procedure, we identify

$$\left.\begin{aligned} &Y_1 = \tfrac{1}{2} + (21/2)s/(s + 2),\ Y_2 = s + 6s/(s + 1), \\ &Y_3 = s + \tfrac{1}{2} \text{ and } Y_4 = (3/2)s/(s + 2) \end{aligned}\right\} \tag{10.23}$$

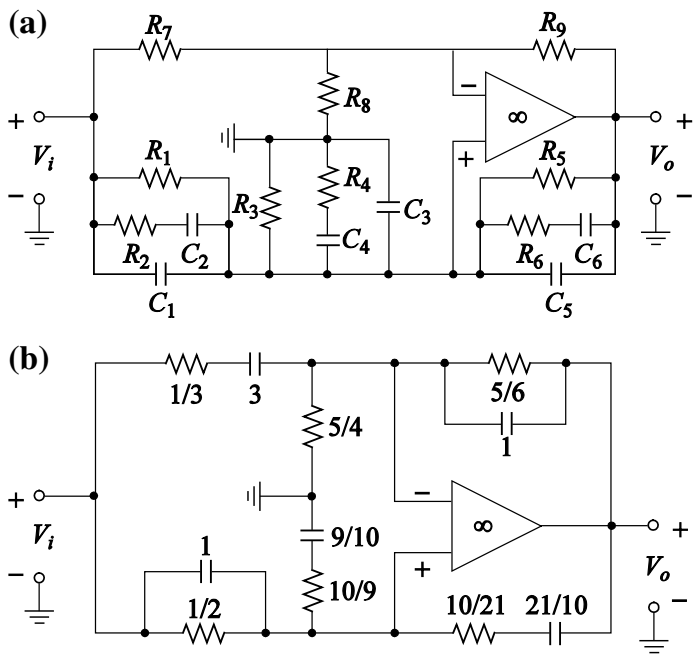


Fig. 10.2 **a** Example 1, synthesized by the new method (Element values are: $R_7 = R_8 = R_9 = R$ (arbitrary); $C_1 = 1$, $C_2 = 0.7$, $C_3 = 2/3$: $C_4 = 2$, $C_5 = 1/3$ and $C_6 = 2.4$: $R_1 = 3/5.2$, $R_2 = 1/0.7$; $R_3 = 1.25$, $R_4 = 1/2$. $R_5 = 1.5$ and $R_6 = 1/2.4$). **b** Example 1, synthesized by Mitra's procedure

Applying Eq. (10.17), we get

$$\left.\begin{aligned} Y_A &= \tfrac{1}{2} + s/3(15/2)s/(s+2),\ Y_C = 1/3 + (4/3)s \\ &\quad +4s/(s+1) \text{ and } Y_E = 1/6 + s/3 + 2s/(s+1) \\ &\quad +(9/2)s/(s+2) \end{aligned}\right\} \tag{10.24}$$

The complete network is shown in Fig. 10.3a. Figure 10.3b shows the network designed by Mitra's method to realize Eq. (10.21) with $Q(s)$ given by Eq. (10.22).

10.4 A Comparison of the Two Methods

In the new method, we require three predetermined resistances (which can be made equal) and three RC one ports to be synthesized, whereas Mitra's method requires the synthesis of six RC one ports. Compared to Mitra's method, a distinctive feature of the present approach is that it allows much more flexibility in the choice of clement values. This flexibility arises because in our method, besides $Q(s)$, one has the additional parameters a, b, m, n, p and q. These six parameters can be judiciously chosen for minimizing one or more of the following: sensitivity, number of passive components and the spread of element values. In Mitra's method, in contrast, the solution for RC one ports is unique, once $Q(s)$ is chosen. Another feature worth mentioning is that in our method, the two tees in the network of Fig. 10.1 can be independently scaled without changing the transfer function [this is because scaling

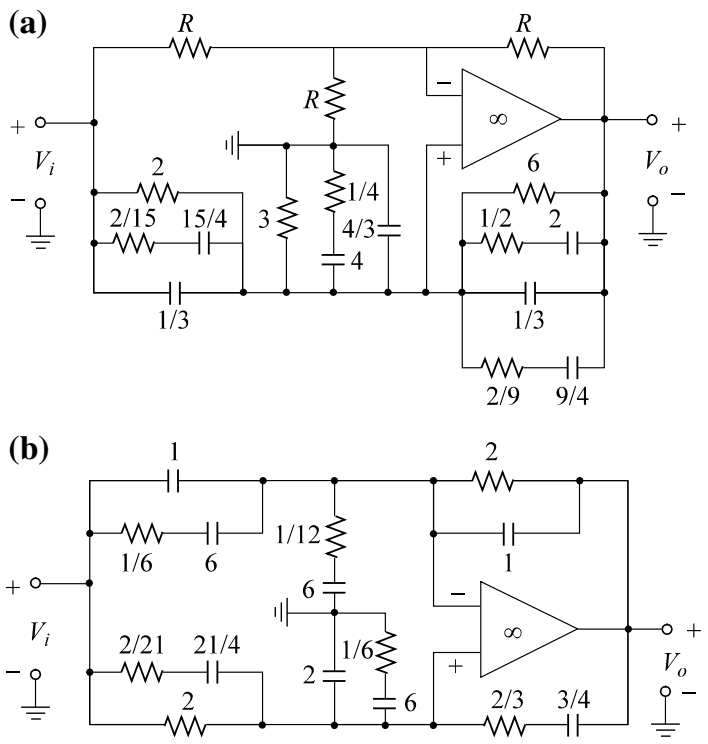


Fig. 10.3 **a** Example 2, synthesized by the new method; **b** Example 2, synthesized by Mitra's procedure

Y_B does not affect the transfer function given by Eq. (10.7)]. On the other hand, in Mitra's method, all the RC one ports have to be scaled by the same factor if $T(s)$ is not to be disturbed.

In the following section, we illustrate the flexibility of the new procedure with the help of Example 1.

10.5 An Illustration of the Flexibility of the New Method

The examples discussed previously show that our method requires more number of circuit elements than Mitra's method. Except the resistance spread in Example 1, the spreads are also seen to be higher in the networks designed by the new method. However, this is not true in general. The examples were worked out with the particular set of values $a = b = m = q = 1$ and $n = p = 0$. In what follows, we show that by taking advantage of the flexibility in choosing these six parameters, we are able to reduce the spreads to much lower values than the ones obtained in the circuits of Fig. 10.2a, b. Further, it will be shown that by choosing the six parameters properly, we may also reduce the number of circuit elements.

Consider Example 1 again. Combining Eq. (10.19) with Eqs. (10.12), (10.13) and (10.14) and letting $a = b = 1$, we get

$$\left.\begin{aligned} Y_A &= (1/3)[(2+m+n)s + (3.2+2m+1.2n)] + (3p+2.1q)s/(s+1)] \\ Y_B &= (1/3)[(1+m+n)s + (0.4+2m+1.2n) + (3p+2.1q+3.9)s/(s+1)] \\ Y_E &= (1/3)[(m+n)s + (2m+1.2n) + (5.1+3p+2.1q)s/(s+1)] \end{aligned}\right\} \tag{10.25}$$

We see from Eq. (10.25) that conditions in Eq. (10.15) are not necessary in this case. The complete network will be of the form shown in Fig. 10.2a with

$$\begin{aligned} &R_7 = R_8 = R_1 = R(\text{arbitrary}) \\ &C_1 = (2+m+n)/3,\ C_2 = P + 0.7q,\ C_3 = (m+n+1)/3 \\ &C_4 = p + 0.7q + 1.3,\ C_5 = (m+n)/3 \text{ and } C_6 = 1.7 + p + 0.7q; \end{aligned}$$

$$\left.\begin{aligned} &R_1 = 3/(3.2+2m+1.2n),\ R_2 = 1/(p+0.7q), \\ &R_3 = 3/(0.4+2m+1.2n),\ R_4 = 1/(p+0.7q+1.3), \\ &R_5 = 3/(2m+1.2n) \text{ and } R_6 = 1/(1.7+p+0.7q) \end{aligned}\right\} \tag{10.26}$$

If we choose $m = n = p = 10$ and $q = 0$, we get

$$\left.\begin{aligned} &R_7 = R_8 = R_9 = R(\text{arbitrary}) \\ &C_1 = 7.33,\ C_2 = 10,\ C_3 = 7,\ C_4 = 11.3,\ C_5 = 6.67 \text{ and } C_6 = 11.7, \\ &R_1 - 1/11.73,\ R_2 = 1/10,\ R_3 = 1/10.8,\ R_4 = 1/11.3, \\ &R_5 = 1/10.67, \text{ and } R_6 = 1/11.7 \end{aligned}\right\} \tag{10.27}$$

Choosing R anywhere between the highest and the lowest of the resistances R_1, …, R_6, we get resistance spread $\rho = 1.17$ and capacitance spread $\gamma = 1.755$, as compared to the figures $\rho = 3.6$ and $\gamma = 7.2$ in Fig. 10.2a and $\rho = 3.75$ and $\gamma = 3.33$ in Fig. 10.2b. Since the ideal values of ρ and γ are unity, we see that the design worked out above is indeed a very suitable one from the point of view of spread.

Next we consider minimizing the number of passive elements. In Eq. (10.26) if we put $m = n = p = q = 0$, then we get

$$\left.\begin{aligned} &R_7 = R_8 = R_9 = R(\text{arbitrary}) \\ &C_1 = 0.67,\ C_2 = C_5 = 0,\ C_3 = 0.33,\ C_4 = 1.3 \text{ and } C_6 = 1.7; \\ &R_1 = 1/1.07,\ R_2 = R_5 = \infty,\ R_3 = 1/0.133, \\ &R_4 = 1/1.3 \text{ and } R_6 = 1/1.7 \end{aligned}\right\} \tag{10.28}$$

Thus we get a total of 11 elements, as compared to 15 in Fig. 10.2a and 11 in Fig. 10.2b. Another desirable feature of the above design is that we have used only 4 capacitors instead of 6 in Fig. 10.2a and 5 in Fig. 10.2b.

References

1. S.K. Mitra, *Anaylsis and Synthesis of Linear Active Networks*, Chap. 11 (Wiley, New York, 1968)
2. W.J. Kerwin, in *Active Filters*, Chap. 2 ed. by L.P. Huelsman (McGraw-Hill Book Co. Inc., New York, 1970)
3. S.K. Mitra, in *Proceedings of the Hawaii International Conference on System Sciences*, Honolulu, Hawaii (1968)

Chapter 11
Second-Order Active RC Bandpass Filters Using Operational Amplifier in Low-Gain Connection

Abstract A critical study is made of a canonic second-order bandpass realization using a single-operational amplifier in low-gain (unity or two) connection. The circuit is suitable for low- and medium Q applications and features good stability, wide frequency range of operation and versatility of conversion to a high-pass filter. An all-grounded-capacitor version of the filter is also given; which is a desirable feature for integrated circuit (IC) implementation. The unity gain configuration presents component minimization over other similar circuits.

Keywords Active RC Bandpass filter · Second order · Operational amplifier · Low-gain connection

11.1 Introduction

A number of circuits have been proposed in the literature for the realization of second-order bandpass (BP) filters using one, two or more operational amplifiers (OA) along with an appropriate passive RC network [1–4]. All these circuits require a gain proportional to Q^2, thus posing stability problems and giving low-operating frequency range for high Q realizations. In order to obviate these disadvantages, Sedra [5] proposed a circuit using an OA in the unity gain connection (UGOA). In this chapter, a critical study is made of a different circuit which also requires a low-gain (unity or two) amplifier, but it enjoys component reduction over Sedra's circuit by two elements, thus making it a canonic structure in both active and passive components. Complex pole loci are plotted and an analysis is made of the frequency limitation based on the one-pole roll-off model of the OA. The upper bound on the frequency is calculated for the filter with OA in (a) unity gain (UG) (b) low-gain ($K_1 = 2$) connections, and it is shown that $K_1 = 2$ realization gives higher operating frequency range for $Q > 2\sqrt{2}$. The UGOA circuit has active sensitivity proportional to Q^2, but very low passive sensitivities. With a gain of 2, the active Q-sensitivity becomes proportional to Q, but passive sensitivities increase. In either

Source: S.C. Dutta Roy & Muslim T. Ahmed, "Second Order Active RC Band-Pass Filters using Operational Amplifier in low Gain Connection," *JIETE*, vol 21, pp 248–252, May 1975.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_11

case, the configuration has zero centre-frequency active sensitivity, $S_{K_1}^{\omega_0}$. An all-grounded version of the filter is also given. This feature coupled with the versatility of conversion to a high-pass (HP) filter makes the circuit attractive for IC hybrid realization for low and medium Q applications.

11.2 The Circuit Configuration

The circuit under consideration is shown in Fig. 11.1a and is essentially the Berndt and Dutta Roy's [6] inductor simulator (shown with the dotted lines), with a low-pass RC section cascaded at the input. The circuit has been briefly discussed by Russel [7]. It is canonic [8] and uses an OA with only five passive components. The transfer function obtained from the equivalent circuit in Fig. 11.1b is of the form

$$T(s) = Hs/\left(s^2 + \omega_0 s/Q + \omega_0^2\right) \tag{11.1}$$

with

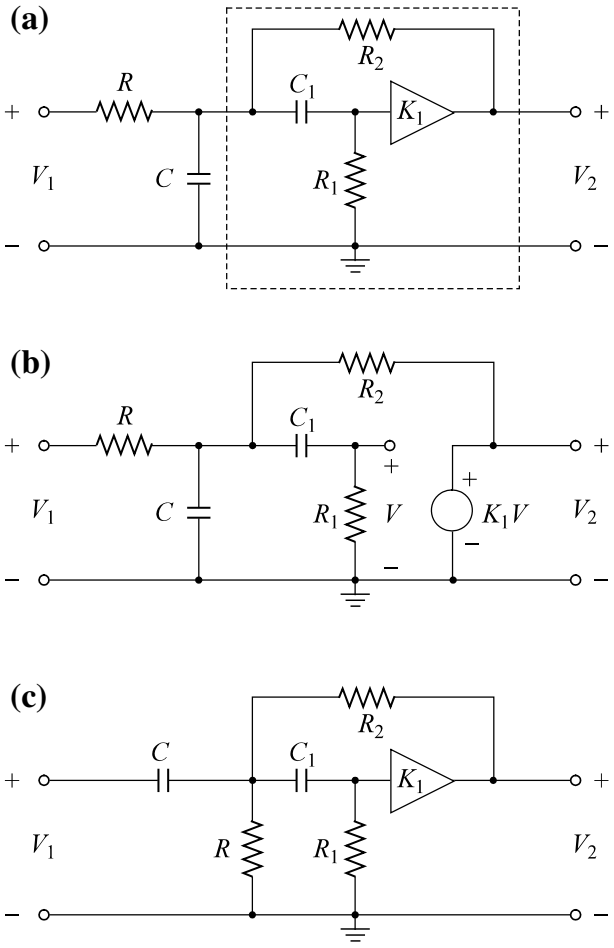


Fig. 11.1 **a** Bandpass filter with Berndt and Dutta Roy's inductor simulator shown within dotted lines. **b** Equivalent circuit of Fig. 11.1a. **c** High-pass filter from Fig. 11.1a

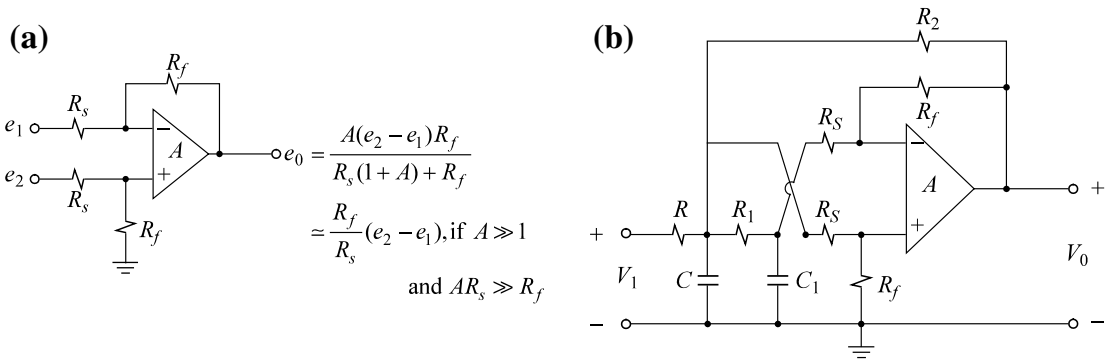


Fig. 11.2 **a** Difference amplifier. **b** All-grounded equivalent of the bandpass filter of Fig. 11.1a

$$H = K_1/RC \tag{11.2}$$

$$\omega_0 = [(R + R_2)/(CC_1RR_1R_2)]^{\frac{1}{2}} \tag{11.3}$$

and

$$Q = [CC_1RR_1R_2(R + R_2)]^{\frac{1}{2}}/[C_1R_2(R + R_1) + CRR_2 + C_1RR_1(1 - K_1)] \tag{11.4}$$

It can easily be verified that if R and C of the input section are interchanged, as in Fig. 11.1c, we get a second-order HP filter with the same ω_0 and Q as in Eqs. (11.3) and (11.4). If R in Fig. 11.1c is infinite, then one gets the Sallen and Key [4] HP circuit.

An all-grounded-capacitor version of the bandpass filter of Fig. 11.1a is shown in Fig. 11.2b, which is suitable for hybrid IC implementation [9]. This filter has the same ω_0 and Q as in Eqs. (11.3) and (11.4) and has been obtained by replacing the non-inverting OA stage by a differential-amplifier stage shown in Fig. 11.2a with its output $e_0 = K_1(e_2 - e_1)$, where $K_1 = R_f/R_s$. The structure is still canonic in capacitors and OA; however, the number of resistors has increased by two.

11.3 Sensitivity Considerations

The most important parameters for the filter performance are the pole Q and the pole-centre frequency ω_0. For the sensitivity of X to e_i, we use the standard definition

$$S_{e_i}^X = \frac{e_i}{X} \cdot \frac{\partial X}{\partial e_i} \tag{11.5}$$

It is apparent from Eq. (11.3) that $S_{K_1}^{\omega_0} = 0$, which is a desirable feature. Also from Eqs. (11.4) and (11.5), we get

$$S_{K_1}^{Q} = K_1 C_1 R R_1 / D \tag{11.6}$$

where D denotes the denominator in Eq. (11.4). The sensitivity of ω_0 to the passive elements are, from Eqs. (11.3) and (11.5)

$$S_R^{\omega_0} = -\frac{R_2}{2(R+R_2)}, \quad S_{R_2}^{\omega_0} = -\frac{R}{2(R+R_2)}, \quad S_{R_1}^{\omega_0} = S_C^{\omega_0} = S_{C_1}^{\omega_0} = -\frac{1}{2} \tag{11.7}$$

Thus $\left|S_{ei}^{\omega_0}\right| \leq \frac{1}{2}$. From Eqs. (11.4) and (11.5), the Q sensitivities can be calculated as

$$\left.\begin{aligned} S_R^Q &= (R + R_2/2)/(R+R_2) - [RR_2(C+C_1) + RR_1C_1(1-K)]/D \\ S_R^Q &= (R_2 + R/2)/(R_2+R) - [R_2C_1(R+R) + RR_2C]/D \\ S_{R_1}^Q &= \tfrac{1}{2} - R_1[C_1R_2 + RC_1(1-K_1)]/D \\ S_C^Q &= \tfrac{1}{2} - (CRR_2)/D \\ S_{C_1}^Q &= \tfrac{1}{2} - [C_1R_2(R+R_1) + C_1RR_1(1-K_1)/D] \end{aligned}\right\} \tag{11.8}$$

The order of S_{ei}^{Q} will be discussed in the following sections.

11.4 Design Equations

11.4.1 Case A: Unity Gain Configuration

Let $K_1 = 1$ and ω_0 be normalized to unity. Further, we preselect $C = C_1 = 1$; then Eqs. (11.3) and (11.4) can be written as

$$G_1(G + G_2) = 1 \text{ and } C = (1/Q) - 2G_1 \tag{11.9}$$

where $G_i = \mathrm{l}/R_i$. Eliminating G and solving for G_1 gives

$$G_1 = \frac{1}{4}\left\{(G_2 + 1/Q) \pm [(G_2 + 1/Q)^2 - 8]^{\frac{1}{2}}\right\} \tag{11.10}$$

For real G_1, we must have

$$G_2 \geq 2\sqrt{2} - (1/Q) \tag{11.11}$$

If G is to be positive, the second equation in (11.9) dictates that $G_1 < 1/(2Q)$. Combining this with Eq. (11.10) and taking negative sign (positive sign in Eq. (11.10) gives negative G) and simplifying, we get the additional restriction $G_2 > 2Q$. The final equations are

$$C = C_1 = 1, K_1 = 1, G_1 < 1/(2Q), G_2 > 2Q,$$

$$G = (1/Q) - 2G_1 \tag{11.12}$$

Under these conditions, Eq. (11.6) gives

$$S^Q_{K_1} = G_2/(G + 2G_1) > 2Q^2 \tag{11.13}$$

As is well known, when unity gain is obtained with an OA of gain A of the order of 10^6, $K = A/(1 + A) \approx 1$; so that $S^Q_A = S^Q_{K_1} \cdot S^K_A 1 > 2Q^2/A$. This shows a reduction of the sensitivity by the factor A. If $A = 10^6$, a Q of 500 will give S^Q_A of the order of $\frac{1}{2}$; the gain-sensitivity product (GS) is, however, of the order of $2Q^2$.

It has already been shown that $\left|S^{\omega_0}_{e_i}\right| \leq \frac{1}{2}$. Also, if $K = 1$, it is evident from Eq. (11.8) that $\left|S^Q_{e_i}\right| \leq 1$. For practical design values, $\left|S^Q_{e_i}\right|$ will be within $\frac{1}{2}$. As an example, let $C = C_1 = 1$, $K_1 = 1$, $R = R_1 = 3Q$, $R_2 = 1/3Q$; then the sensitivities are, for medium to high Q.

$$\left.\begin{array}{l} S^{\omega_o}_A = 0,\ S^Q_A = 3Q^2/A \\ S^{\omega_o}_R \simeq S^{\omega_o}_{R_2} \simeq -\frac{1}{2} = S^{\omega_o}_R = S^{\omega_o}_{C_1} = S^{\omega_o}_{R_1} \\ S^Q_R \approx 1/3,\ S^Q_{R_2} \simeq -\frac{1}{2},\ S^Q_{R_1} = S^Q_C = -S^Q_{C_1} = 1/6 \end{array}\right\} \tag{11.14}$$

The UGOA circuit has thus very low passive sensitivities and also presents component reduction. However, because of the high value of gain-sensitivity product (GS), the usefulness is limited to low and medium Q's only. It has been found that increasing the spread in capacitors does not lead to any marked improvement.

11.4.2 Case B: General K Configuration

For a general K_1, proceeding on similar lines, the following result is obtained for the case of minimum possible $S^Q_{K_1}$:

$$S^Q_{K_1 \min} = 2\sqrt{2}.Q - 1,\ K_1 > 2 - 1/(\sqrt{3}Q) \tag{11.15}$$

This implies that $K_1 = 2$ suffices to get any Q; and this can be obtained by using a high gain OA with equal resistors. The value of GS for this case is proportional to Q and the design equations are

$$K_1 = 2, C = C_1 = 1, G_2 = \sqrt{2} - 1/(2Q), G_1 = 1/\sqrt{2}, G = 1/(2Q) \tag{11.16}$$

Sensitivity calculation gives $\left|S^{\omega_o}_{ei}\right| \leq \frac{1}{2}$ as before; however, S^Q_C, $S^Q_{C_1}$, $S^Q_{R_1}$ and $S^Q_{R_2}$ become proportional to Q whereas $S^Q_R \simeq 3/2$. These passive Q sensitivities

agree with the order given by Mitra [10] for the denominator of the form $D(s) = (s+\alpha_1)(s+\alpha_2) - K_1 s$, to which Eq. (11.1) belongs. Thus only for $K_1 = 1$, there is a very significant reduction in the passive Q-sensitivity.

11.5 Frequency Limitation

In this section, frequency limitations are discussed for cases A and B and expression are derived for the highest usable frequency (ω_{0h}) for a given percentage deviation from nominal ω_0 and Q. The analysis and the pole-loci are based on the one-pole roll-off model of the OA. Assuming a finite gain-bandwidth product (GB), an infinite input impedance and $\{R_a A/(R_a + R_b)\} \gg 1$, the gain of the non-inverting amplifier is

$$K_1 = GB/\{s + (GB/K)\} \tag{11.17}$$

where A is the open-loop gain of OA and $K = 1 + R_b/R_a$ is the gain corresponding to infinite gain-bandwidth product (GB) and R_a, R_b are the biasing resistances.

11.5.1 Case A: OA in Unity Gain Configuration

With $K_1 = 1$, on substituting Eq. (11.17) in Eq. (11.1), the normalized denominator of the transfer function becomes

$$D(s_n) = (s_n^2 + s_n/Q + 1) + \frac{1}{GB_n} \cdot s_n[s_n^2 + (1/Q + L)s_n + 1] \tag{11.18}$$

where $s_n = s/\omega_o$, $GB_n = GB/\omega_o$, and

$$L = \{RR_1/[R_2(R+R)]\}^{\frac{1}{2}} \tag{11.19}$$

Using Budak–Peterela method [11], Eq. (11.18) can be put as $P_1(s_n) + (1/GB_n)P_2(s_n)$, with

$$\left.\begin{aligned} P_1(s_n) &= s_n^2 + (s_n/Q) + 1 \\ P_2(s_n) &= s_n[s_n^2 + (1/Q + L)s_n + 1] \end{aligned}\right\} \tag{11.20}$$

giving

$$\left.\begin{aligned} \frac{dP_n}{d(1/GB_n)} &= -P_2(s_n)/\frac{dP_1(s_n)}{ds_n}\Big|_{s_n=p_n} \\ &\triangleq -\frac{d\alpha_n}{d(1/GB_n)} + j\frac{d\beta_n}{d(1/GB_n)} \end{aligned}\right\} \tag{11.21}$$

where

$$p_n = -\alpha_n + j\beta_n = -(1/2Q) + j[1 - 1/4Q^2]^{\frac{1}{2}} \tag{11.22}$$

Following Budak et al. [11], the above equations can be manipulated to give

$$\frac{d\alpha_n}{d(1/GB_n)} = \frac{-L}{2Q}, \frac{d\beta_n}{d(1/GB_n)} = -\frac{(4Q^2-2)L}{8Q^2[1-1/4Q^2]^{\frac{1}{2}}} \tag{11.23}$$

$$\simeq -L/2 \tag{11.24}$$

where Eq. (11.24) holds for medium and high Q. Using Budak [11], Eqs. (11.13) and (11.14), and substituting Eq. (11.23) yields

$$\frac{d\omega_o}{d(1/GB_n)} = -\frac{\omega_o L}{2}; \frac{dQ}{d(1/GB_n)} = \frac{QL}{2} \tag{11.25}$$

Thus after the first-order correction to the nominal values of ω_0 and Q_1, the actual values are

$$\begin{aligned} \omega_{oa} &= \omega_o(1 - \omega_o L/2GB) \\ Q_a &= Q(1 + \omega_o L/2GB) \end{aligned} \tag{11.26}$$

Equations in (11.26) are the frequency limitation equations of the filter with UGOA. For a prescribed variation in ω_0 and Q of x percent from the nominal values, the highest usable frequency is

$$\omega_{o\hbar} \leq x \cdot GB/(50L) \tag{11.27}$$

11.5.2 Case B: OA in Gain = 2 Configuration

Proceeding as in case A it can easily be shown that Eq. (11.18) still represents the pole-defining equation with $GB'_n (= GB_n/2)$ replacing GB_n. The differential change in the real and imaginary parts of the complex poles and the actual ω_0 and Q can easily be shown to be

$$\frac{d\alpha_n}{d(1/GB'_n)} = \frac{-L}{Q}; \frac{d\beta_n}{d(1/GB'_n)} = -\frac{(4Q^2-2)L}{4Q^2[1-1/4Q^2]^{\frac{1}{2}}} \tag{11.28}$$

$$\left.\begin{aligned} \omega_o a &= \omega_o(1 - 3\omega_o L/GB) \\ Q_a &= Q(1 + 2\omega_o L/GB) \end{aligned}\right\} \tag{11.29}$$

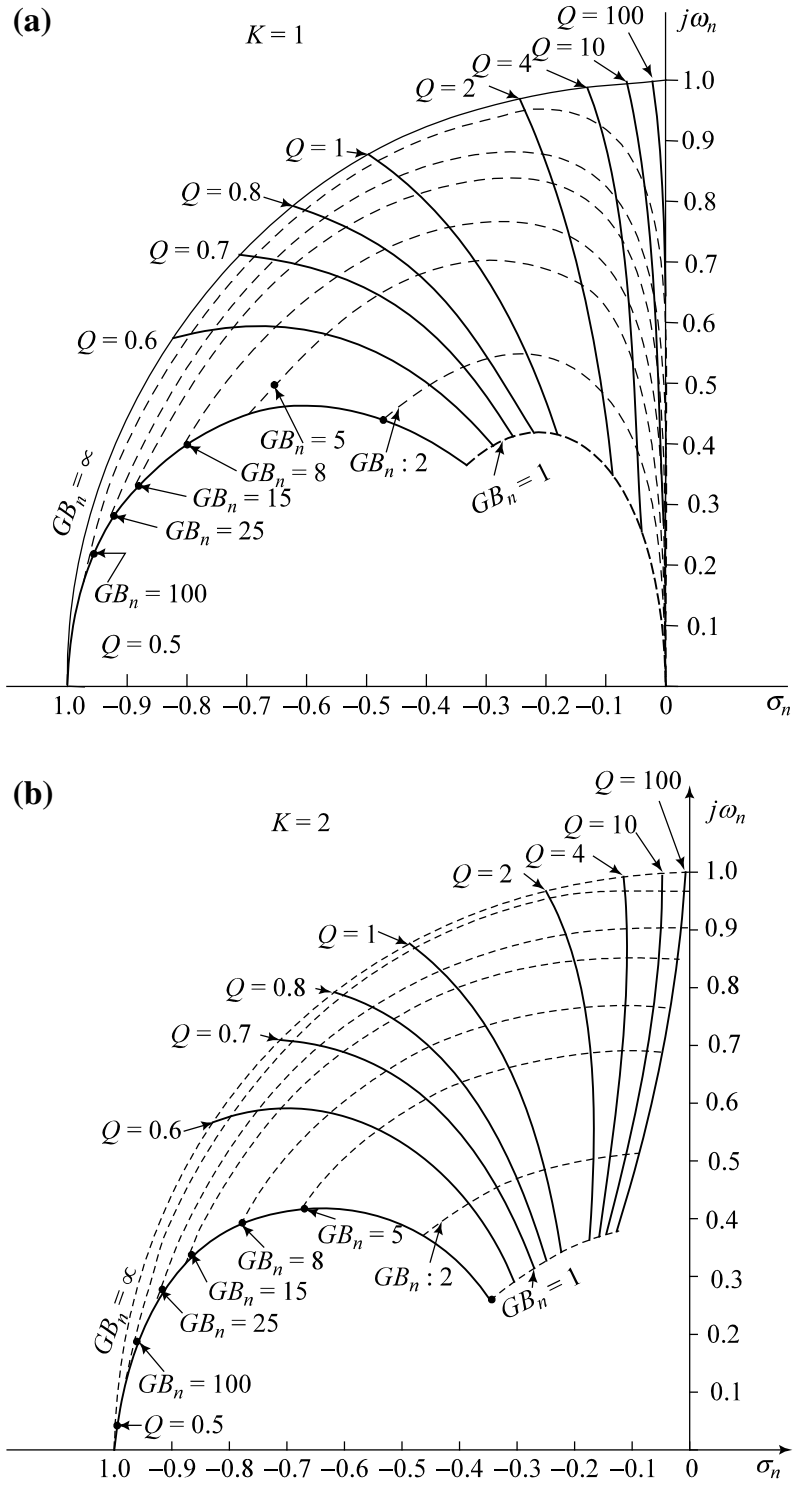


Fig. 11.3 **a** Pole-loci for UGOA realization. **b** Pole-loci for general low-gain ($K = 2$) realization

Thus the highest usable frequency for x percent allowed variation is

$$\omega_{0\hbar} < x \cdot GB/(200L) \tag{11.30}$$

The quantity given by Eq. (11.19) is an important design parameter for determining the value of ω_{0a}, Q_a and $\omega_{0\hbar}$. In order to compare the frequency limitations of the two cases, we find from the design Eqs. (11.12) and (11.16) developed in Sect. 11.4,

$$L\Big|_{K}^{\min} = 1 > 2Q \text{ and } L\Big|_{K}^{\max} = 2 < \sqrt{2} \tag{11.31}$$

giving

$$\left.\begin{aligned} \omega_{o\hbar}|K = 1 &< x/100(GB/Q) \\ \omega_{o\hbar}|K = 2 &> x/200(GB/\sqrt{2}) \end{aligned}\right\} \tag{11.32}$$

Then for $Q > 2\sqrt{2}$, OA in $K_1 = 2$ connection can realize higher $\omega_{0\hbar}$ compared with the UGOA case for the same percentage deviation in ω_o and Q from nominal values.

Equation (11.18) is used for plotting the complex pole-loci as shown in Fig. 11.3a, b for cases A and B, respectively, as a function of Q and GB_n for L determined from the design equations. The two sets of curves are similar in nature and show that a decrease in GB_n from the nominal value results in a new pole location having lower ω_0 (and probably also Q). With increasing Q, the curves of Fig. 11.3a approach the $j\omega_n$—axis faster than those of Fig. 11.3b. However, the poles are always confined to the left half-plane in both the cases, thus ensuring stable operation.

11.6 Conclusion

This chapter discusses canonic bandpass filters using OA in unity gain and low-gain configurations and thus giving a large frequency range of operation. A detailed study is made of the two cases with regard to sensitivity, gain-sensitivity product and the frequency limitations. The UGOA circuit has low passive sensitivity besides giving passive component minimization over similar circuits. By increasing the gain to $K_1 = 2$, a marked improvement is shown in the active Q-sensitivity, the gain-sensitivity product and the frequency limitations (for $Q > 2\sqrt{2}$) over the UGOA case. The improvement is however at the cost of passive Q-sensitivity, thus demanding higher tolerances in R's and C's. An all-grounded-capacitor equivalent of the filter is given which, coupled with the other advantages already discussed, makes it attractive for IC implementation. The circuit enjoys $S_{K_1}^{\omega_o} = 0$.

References

1. M.A. Soderstrand, S.K. Mitra, Extremely low sensitivity active RC filters. Proc. IEEE **57**(12), 2175 (1969)
2. P.R. Geffe, A Q-invariant active resonator. Proc. IEEE **57**(8), 1442 (1969)
3. T. Deliyannis, A low-pass filter with extremely low sensitivity. Proc. IEEE **58**(9), 1366–1367 (1970)
4. R.P. Sallen, E.L. Key, A practical method of designing RC active filters. Trans. IRE. CT **2**(1), 74–85 (1955)
5. A.S. Sedra, A class of stable active filters using unity gain voltage followers. IEEE J. SC **7**(4), 311–315 (1972)
6. D. Berndt, S.C. Dutta Roy, Inductor simulation using a unity gain amplifier. IEEE J. SC **4**(2), 161–162 (1969)
7. H.T. Russel, Single amplifier active filters give stable Q. EDN/EEE. **17**(1), 65–66 (1972)
8. F. Anday, Realization of 2nd–order transfer function using a minimum number of elements. Electron. Lett. **8**(24), 611–612 (1972)
9. G.S. Moschytz, Gain sensitivity product a figure of merit for hybrid—integrated filters using single operational amplifiers. IEEE J. SC **6**(2), 103–110 (1971)
10. S.K. Mitra, *Filter design using integrated operational amplifiers* (Proc. WESCON, San Francisco, California, 1969)
11. A. Budak, D.M. Petrela, Frequency limitations of active filters using operational amplifiers. IEEE Trans. CT **19**(3), 322–328 (1972)

Chapter 12
Some Recent Contributions to Inductor Simulation and Applications to Low-Sensitivity Biquad Design

Abstract Following a brief discussion on the advantages of inductor simulation approach to active filter design, some recent circuits for grounded as well as floating inductors are discussed. The application of simulated inductors to insensitive biquadratic design is illustrated by reference to two recently published techniques.

Keywords Inductance simulation · Active RC networks

12.1 Introduction

As is well known, active RC filters can be designed either in direct form or in cascade form. In the latter, the prescribed transfer function is realized as a cascade of lower order sections, usually active second order and passive first-order ones. In the direct form, no such decomposition is necessary and the realization is in terms of one structure. Of the various techniques available for this purpose, the more well-known ones are the inductance (L) simulation, the frequency-dependent negative resistance and the generalized impedance converter approaches. It is the first approach that the present chapter is concerned with. In this, a conventional RLC filter is first designed, and then each inductor is replaced by an equivalent active RC one port. There are two advantages of this simple minded procedure. First, the well-established passive synthesis procedures and extensive tables can be used conveniently. Second, as pointed out in [1], this technique often leads to a more stable and a less sensitive realization as compared to other methods. The advantages with regard to sensitivity can be illustrated by considering the doubly terminated lossless network of Fig. 12.1. Let

$$Z_i(j\omega) = R + jX \tag{12.1}$$

The power input P_i (and hence the power delivered to the load, P_L) is given by

Source: Suhash C. Dutta Roy, "Some Recent Contributions to Inductor Simulation and Applications to Low Sensitivity Biquad Design," *JIETE*, vol 22, pp 419–425, July 1976.
An up-dated version of a presentation at the Seminar on 'Recent Developments in Filter Theory and Design' held at Hyderabad, December 19–21, 1974.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_12

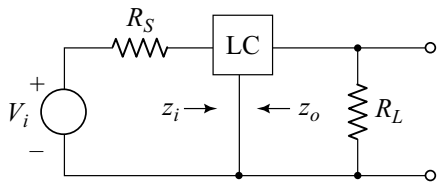


Fig. 12.1 A doubly terminated LC network

$$P_i = P_L = |V_i|^2 R/\left[(R + R_s)^2 + X^2\right]. \tag{12.2}$$

Let x be a parameter of the LC network. Then the sensitivity of P_L to x is

$$S_X^{P_L} = (x/P_L)[(\partial P_L/\partial R)(\partial R/\partial x) + (\partial P_L/\partial X)(\partial X/\partial x)] \tag{12.3}$$

From Eq. (12.2)

$$\partial P_L/\partial R = |V_i|^2(R_s^2 - R^2 + X^2)/[(R + R_s)^2 + X^2]^2 \tag{12.4}$$

and

$$\partial P_L/\partial X = -2|V_i|^2 RX/[(R + R_s)^2 + X^2]^2. \tag{12.5}$$

Normally, doubly terminated LC filters are designed to have maximum power transfer at frequencies of minimum loss in the passband. This implies that at these frequencies,

$$R = R_s \text{ and } X = 0$$

which, combined with Eqs. (12.4) and (12.5), lead to the conclusion that P_L is insensitive to small variations in the components of the LC network. Similar arguments can be given to show that the output impedance Z_0 should equal R_L for zero sensitivity in the passband.

The essential task in L simulation approach to active filter design is, obviously, that of designing insensitive active RC one ports having impedances proportional to the complex frequency s. Also, for successful implementation of this technique, the circuits for both grounded and floating inductor simulation are required.

12.2 Gyrator Based Methods

Essentially, all active RC circuits for simulating an inductor can be viewed as the realization of gyrators, a gyrator being a nonreciprocal two port, ideally described by the z-matrix

$$[z] = \begin{bmatrix} 0 & -r \\ r & 0 \end{bmatrix} \tag{12.6}$$

The symbolic representation of a gyrator and its use in simulating a grounded inductor are shown in Fig. 12.2. When two gyrators are cascaded with a capacitor interposed between the two, as shown in Fig. 12.3, one gets a floating inductor. It appears, therefore, that all problems of L simulation are solved if one can realize the z-matrix of Eq. (12.6). A large number of circuits using discrete components and integrated gyrators have been proposed in the literature, of which the Riordan circuit [2] is the most popular; a variation of the Riordan circuit, proposed by Antoniou [3] is shown in Fig. 12.4 and is considered as the best circuit for the purpose.

In general, a practical gyrator deviates from the ideal one, described by Eq. (12.6), in that $z_{11} \neq 0$, $z_{22} \neq 0$, and $z_{12} \neq -z_{21}$. The active element phase shifts cause additional complications. These lead to stability and sensitivity problems, but the most important effect is that the realized L becomes frequency dependent and lossy with a finite maximum Q [4].

The problem becomes much more serious in the case of floating-L simulation because, in addition to realizing two ideal gyrators, we also need an exact match between their gyration resistances. Many solutions have so far been proposed, but it

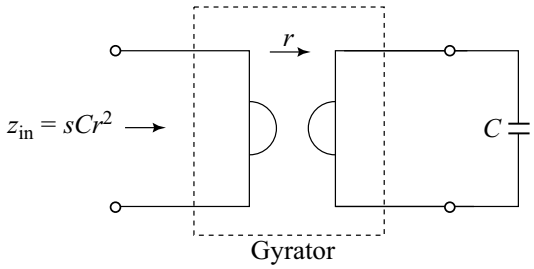


Fig. 12.2 Schematic of a gyrator and its use in L simulation

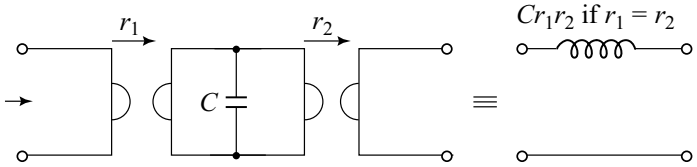


Fig. 12.3 Two gyrators used to simulate a floating L

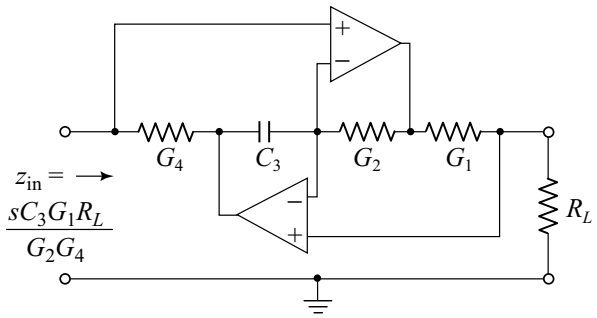


Fig. 12.4 Antoniou's circuit for a gyrator

has been appreciated that all such gyrator based realizations suffer from the need for accurate matching, in addition to excessive number of active and/or passive elements.

It has been amply demonstrated in recent years that what can be achieved by nonideal gyrators can also be obtained by using much simpler and more economic circuits. For example, a lossy grounded inductor circuit may typically consist of one operational amplifier (OA) connected in the unity gain (UG) configuration, two resistors and a capacitor. Similarly, a lossy floating inductor can be made from two UG amplifiers (UGA), three resistors and two capacitors. It is with some of these circuits and their applications that we concern ourselves in the rest of this chapter [5].

As already mentioned, we deal here with some selected contributions only; for other recent contributions, one may refer to the bibliography prepared by the author as referred to in [6], which lists papers published during the period 1971–early 1975. For literature published earlier, reference is made to Beddoes and Morin [7], Mitra [8] and Miller and Newcomb [9].

12.3 General Characteristics of Lossy, Grounded *L* Simulators

All L simulators using a single capacitor realize a bilinear RL impedance of the form

$$Z(s) = H(s\tau_1 + 1)/(s\tau_2 + 1),\ \tau_1 > \tau_2 \tag{12.7}$$

Two canonical passive realizations of Eq. (12.7) are shown in Fig. 12.5. An important special case arises when $\tau_2 = 0$, thus giving a series RL impedance; this will be designated as type A. A parallel RL impedance is obtained when $H \to 0$, $\tau_1 \to \infty$, but $H\tau_1$ is finite, this will be called type B. The general case of Fig. 12.5 will be referred to as type C. Setting $s = j\omega$ in Eq. (12.7), it is easy to show that the realized L and its Q are

$$L = H(\tau_1 - \tau_2)/(1 + \omega^2\tau_2^2)\ \&\ Q = \omega(\tau_1 - \tau_2)/(1 + \omega^2\tau_1\tau_2) \tag{12.8}$$

with the Q reaching a maximum value of

$$Q_0 = (1/2)\left(\sqrt{\tau_1/\tau_2} - \sqrt{\tau_2/\tau_1}\right) \tag{12.9}$$

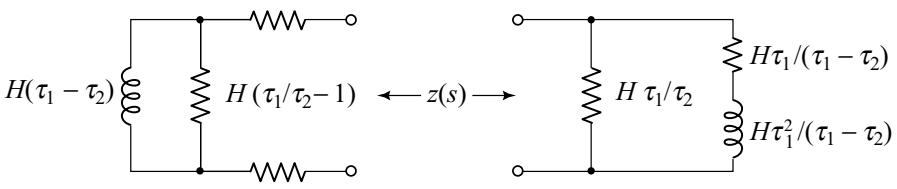


Fig. 12.5 Canonical representations of a bilinear RL impedance

at the frequency

$$\omega_0 = 1/\sqrt{\tau_1 \tau_2} \tag{12.10}$$

At $\omega = \omega_0$, the inductance value is

$$L_0 = L_{dc}/(1 + \tau_2/\tau_1) \cong L_{dc} \tag{12.11}$$

where L_{dc} is the dc value of L and the approximation is valid for high Q_0. For types A and B, L and Q are, respectively, given by

$$L = H\tau_1 \text{ and } Q = \omega\tau_1 \tag{12.12}$$

and

$$L = H\tau_1/(1 + \omega^2\tau_2^2) \text{ and } Q = 1/(\omega\tau_2) \tag{12.13}$$

The sensitivities of L and Q to any parameter can be found by routine calculations. For type C, evaluating the sensitivity at $\omega = \omega_0$, it can be shown that

$$S_X^Q = (1/2)[(\tau_1 + \tau_2)/(\tau_1 - \tau_2)](S_X^{\tau_1} - S_X^{\tau_2}). \tag{12.14}$$

Since a single capacitor, C, is used, both τ_1 and τ_2 must be proportional to C; thus

$$S_C^Q|_{\text{Type C},\omega=\omega_0} = 0 \tag{12.15}$$

12.4 Simple Circuits for Lossy Grounded Inductance

12.4.1 Four-Component Circuits

Two remarkably simple circuits were forwarded by Prestcott [10] and Berndt and Dutta Roy [11] and are shown in Fig. 12.6a and b, respectively. Note that the only

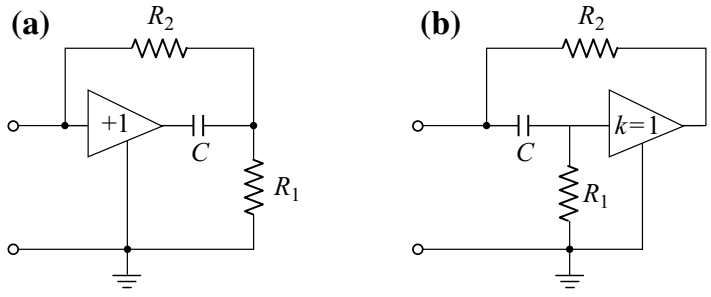
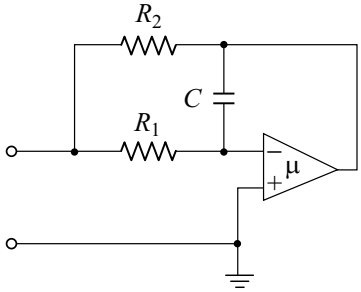


Fig. 12.6 **a** Prestcott circuit for grounded L simulation. **b** Berndt–Dutta Roy circuit for the same purpose

Table 12.1 Performance characteristics of Prestcott and Berndt–Dutta Roy circuits

	Prestcott	Berndt–Dutta Roy (at $\omega = \omega_0$)
Type	A	C
R-spread	May be unity	$4Q_0^2$
S_μ^L (μ = OA gain)	$\cong 2/\mu$	$\cong 3/\mu$
$S_{C,R_1R_2}^L$	1	$\cong 1$
S_C^Q	1	0
$\left\lvert S_{R_1,R_2}^Q \right\rvert$	1/2	$\cong 1/2$
S_μ^Q	$\cong 2Q^2/\mu$	$\cong 2Q_0^2/\mu$
Gain Sensitivity Product (GSP)	$\cong 2Q^2$	$\cong 2Q_0^2$

Fig. 12.7 Ford and Girling scheme for L simulation



difference between them is the position of the CR differentiating network. This, however, leads to quite different characteristics, as given in Table 12.1. A non-zero input admittance (G_i) and output impedance (R_0) of the active element can be absorbed in R_1 and R_2, respectively in Berndt–Dutta Roy circuit; the effect of either or both in Prestcott circuit is to convert this to a type C circuit.

Figure 12.7 shows a circuit proposed by Ford and Girling [12]. With $\mu = \infty$, which is a rather ideal situation, it behaves as type B, with the following characteristics:

R-spread may be unity;

$$S_\mu^{L,Q} = 0;\ S_{C,R^1,R^2}^L \cong 1;$$

$$S_C^Q = -1;\ \left\lvert S_{R^1,R^2}^Q \right\rvert - 1/2 \text{ and GSP} = 0. \tag{12.16}$$

where GSP = Gain Sensitivity Product.

With finite μ, however, we get a type C circuit with deteriorated characteristics, the major effect being that the GSP now becomes $2Q_0^2$. Further, since the OA is used here without any negative feedback other than that provided by the R_1 R_2 C network, the performance of the circuit will be severely frequency limited.

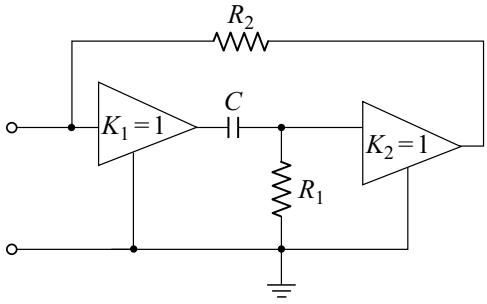


Fig. 12.8 Rao–Venkateswaran generalization of Prestcott and Berndt–Dutta Roy circuit

12.4.2 Five-Component Circuit

A generalization of Prestcott and Berndt–Dutta Roy circuits was proposed by Rao and Venkateswaran [13] and is shown in Fig. 12.8. This is a type *A* circuit with the following characteristics:

R-spread may be made unity;

$$\begin{aligned} S^{L}_{\mu_1,\mu_2} &= 1/(\mu_{1,2}+1); \\ S^{L}_{C,R_1,R_2} &= 1;\ S^{Q}_{\mu_1,\mu_2} = (1+Q^2)/(\mu_{1,2}+1);\ S^{Q}_{C,R_1} = 1; \end{aligned} \tag{12.17}$$

$S^{Q}_{R} = 0$; $(\mathrm{GSP})_1 \cong (\mathrm{GSP})_2 \cong Q^2$; and L can be adjusted independent of Q.

12.4.3 Circuits Using Six Components

The Berndt–Dutta Roy circuit of Fig. 12.6b realizes an impedence

$$Z(s) = R_2(sCR_1+1)/\{sC[R_2+R_1(1-k)]+1\} \tag{12.18}$$

If we let

$$k = 1 + (R_2/R_1) \tag{12.19}$$

in the manner shown in Fig. 12.9, we get a six-component type *A* circuit. However, this circuit is potentially unstable because, if due to any reason, for example, a change of temperature, k becomes greater than $1 + (R_2/R_1)$, i.e. $R_1\ R_A$ becomes greater than R_2R_B, a pole is created in the right half-plane.

The Caggiano circuit [14] has the same topology as that of Fig. 12.9 but the impedance is now measured across R_1. It suffers from the same stability problem as in the circuit just discussed. With k in the range $1 < k < (1 + R_2/R_1)$ in either of these two circuits, one obtains a type *C* performance with no improvement over the Berndt–Dutta Roy circuit.

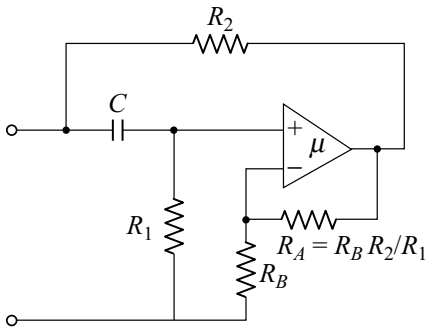


Fig. 12.9 Berndt–Dutta Roy circuit operated as type *A*

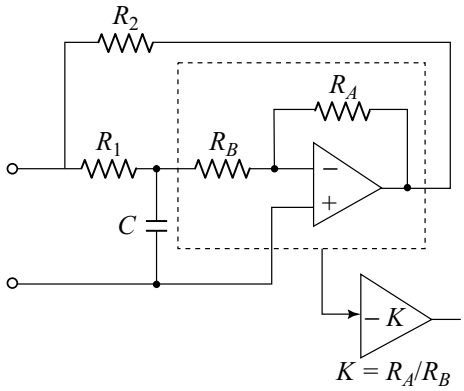


Fig. 12.10 Musiedlak and McPherson circuit

The circuit shown in Fig. 12.10 uses a negative gain amplifier and has been suggested by Musiedlak and McPherson [15]. Note that this circuit can be derived from the low-sensitivity Sallen and Key [16] resonator in Geffe [17], by interchanging the resistors and capacitors and short circuiting the input capacitor.[1] This is a type-C circuit with the following features.

Minimum gain required, $k = 4Q_0^2$; R-spread $= 16Q_0^4$;

$$S_\mu^L \cong -4Q_0^2/\mu;\ \left|S^L_{RA,B,C_1,2}\right| \cong 1;\ S_\mu^Q \cong 2Q_0^2/\mu;\ \left|S^Q_{RA,B}\right| \cong 1/2;\ \left|S^Q_{R_1,2}\right| \cong 1/(8Q_0^2);$$
$$S_C^Q = 0 \text{ and } GSP = 2Q_0^2 \tag{12.20}$$

The sensitivities are seen to be low, and the grounded capacitor is a desirable feature for IC implementation; however, the requirement of high gain severely limits the frequency range of application. Another disadvantageous feature is the high *R*-spread, and the design constraint of $R_B \gg 1/(\omega C)$ at the frequencies of interest.

By changing the position of the R_1C integrating network in Fig. 12.10 to the output, we obtain the circuit of Fig. 12.11, which has the same sensitivity and other performance characteristics as those of the parent circuit, but has the advantage that the finite input impedence R_B of the gain block does no longer put a design constraint.

If the capacitor in Fig. 12.10 is returned to the output instead of ground, as shown in Fig. 12.12, we get the finite gain version of Ford and Girling circuit. Compared to

[1]The Prestcott and Berndt–Dutta Roy circuits may also be so derived from the Sallen and Key catalog [17].

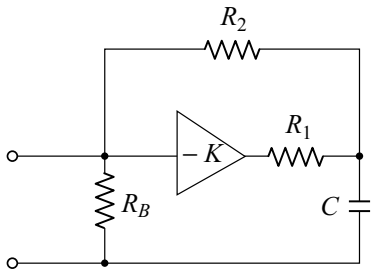


Fig. 12.11 Circuit derived from Fig. 12.10 by changing the position of the R_1C network

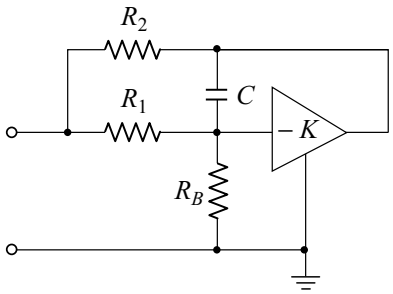


Fig. 12.12 Circuit derived from Fig. 12.10 returning C to amplifier output

Fig. 12.10, it loses the advantage of a grounded capacitor but has the advantage of reducing S^L_{μ,R_A,R_B} by the factor $1/(4Q_0^2)$.

12.4.4 Comments

A major conclusion of this presentation of grounded inductance simulators is that while the sensitivity of L and Q to passive elements can be made reasonably small, the sensitivity of Q to the active element and the gain itself cannot be reduced simultaneously. In fact, in all the circuits discussed, the gain sensitivity product is proportional to Q^2. An improvement in this figure is possible only at the expense of an excessive number of elements.

In concluding this section, some comments on the effect of finite OA gain and its low-pass frequency response on the realized L and its Q appear to be in order. This study has been conducted by Ahmed and Dutta Roy [18] on the tuned circuits formed by connecting a capacitor across the simulated L. With ideal OA, the zeros of the admittance of the tuned circuit are determined by the equation

$$sC_t + (1/H)(s\tau_2 + 1)/(s\tau_1 + 1) = 0 \tag{12.21}$$

where C_t is the tuning capacitor. On simplification, Eq. (12.21) yields

$$s^2 + [(\tau_2 + HC_t)/(HC_t\tau_1)]s + 1/(HC_t\tau_1) = 0 \tag{12.22}$$

from which the quality factor Q_t and the resonance frequency ω_t are determined to be

$$Q_t = (HC_t\tau_1)^{1/2}/(\tau_2 + HC_t) \tag{12.23}$$

and

$$\omega_t = (HC_t\tau_1)^{-1/2} \tag{12.24}$$

If each OA is represented by its one-pole roll-off model, *viz.*

$$\mu_i = \mu_{0i}/\{1 + (s/\omega_i)\} \tag{12.25}$$

then for an n-OA circuit, Eq. (12.22) gets changed to an equation of the form

$$S^{n+2} + \lambda_{n+1}s^{n+1} + \cdots + \lambda_1 s + \lambda_0 = 0 \tag{12.26}$$

where λ_i's are constants. Equation (12.26) must have a dominant complex root pair and can be written as

$$\left(s^2 + \{s\omega_{te}/Q_{te}\} + \omega_{te}^2\right)\left(s^n + \beta_{n-1}s^{n-1} + \ldots + \beta_0\right) = 0$$

where β_i's are constants and ω_{te} and Q_{te} are, to a good order of approximation, the effective ω_t and Q_t. A computer algorithm FLAC (Frequency Limitations of Active Circuits) has been written to find Q_{t0} and ω_{t0} and hence the percentage deviation, with Q_t, ω_t, μ_{0i} and ω_i as the input data. The results show that the Rao–Venkateswaran circuit has the best performance amongst all such circuits. The negative gain circuits, as expected, have a very low highest usable frequency compared to unity gain circuits.

12.5 Lossy Floating Inductance Simulation

A simple, lossy floating inductance simulator circuit is presented in [19]. Consider the Prestcott circuit of Fig. 12.6a with R_2 removed from ground, to form the two port of Fig. 12.13.

Routine analysis shows that the y-parameters of this circuit are

$$y_{11} = -y_{12} = 1/(R_1 + R_2 + sCR_1R_2) \tag{12.27}$$

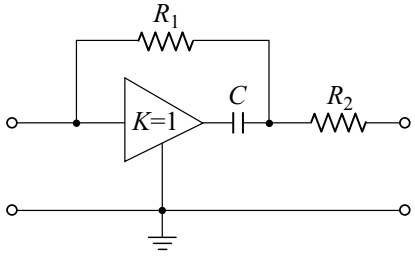


Fig. 12.13 Prestcott circuit converted to a two port

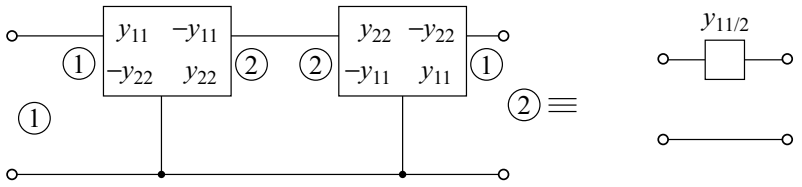


Fig. 12.14 Basic principle of floating immittance simulation

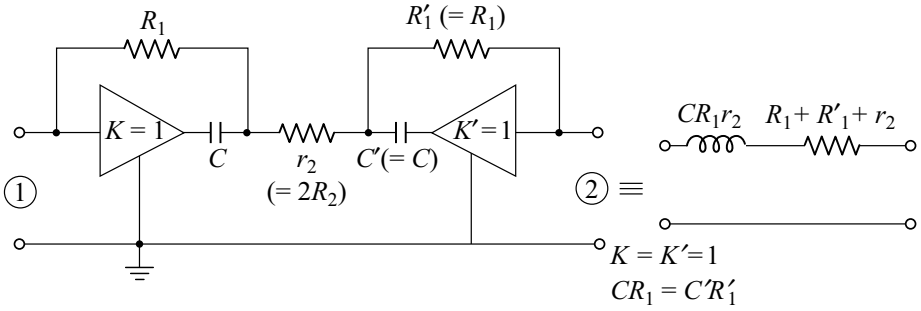


Fig. 12.15 Dutta Roy's floating-L simulator

and

$$y_{22} = -y_{21} = (sCR_1 + 1)(R_1 + R_2 + sCR_1R_2). \tag{12.28}$$

If two such two ports are cascaded back to back, as in Fig. 12.14, it is easy to show that the result is a floating inductor of impedence

$$Z(s) = R + sL = 2(R_1 + R_2) + 2sCR_1R_2. \tag{12.29}$$

The complete circuit is shown in Fig. 12.15, where the values within brackets represent nominal ones.

Although the two-component blocks in Fig. 12.14 were assumed identical, it can be shown that this is not essential for obtaining a floating L. For this, one has to analyse Fig. 12.15 with general element values; assuming $k \cong k' \cong 1$ (with OAs used to realize k and k', this assumption is quite well justified in practice), the only condition required to be satisfied for a floating L turns out to be

$$CR_1 = C'R'_1. \tag{12.30}$$

The resulting floating admittance is

$$Z(s) = R_1 + R'_1 + r_2 + sCR_1r_2. \tag{12.31}$$

where $r_2 = R_2 + R'_2$. In practice, Eq. (12.31) can be realized by adjusting a single resistance (R_1 or R'_1) so that the voltages at ports 1 and 2 are indistinguishable.

In order to study the sensitivity characteristics, a general analysis of Fig. 12.15 has been done. The resulting y-parameters are written as $Y_{mn}(j\omega) = \pm 1/(R_{mn} + j\omega L_{mn})$, where $m, n = 1, 2$ and the negative sign is used for $m \neq n$. The sensitivities of L_{mn} and $Q_{mn} = \omega L_{mn}/R_{mn}$ are evaluated with respect to various parameters and finally, the nominal values of the parameters are substituted. When this is done, it is found that for $\mu = \mu'$ and reasonably high Q's (> 10),

$$S_{\mu'}^{L_{11},Q_{11},L_{21},Q_{21}} \; S_{\mu}^{L_{22},Q_{22},L_{12},Q_{12}} = 0 \tag{12.32}$$

$$S_{\mu}^{L_{11}} = S_{\mu'}^{L_{22}} = (2 + 1/\Psi)/\mu$$

$$S_{\mu}^{Q_{11}} = S_{\mu'}^{Q_{22}} = (2 + 1/\Psi)Q^2/\mu$$

where $\Psi = R_1/r_2$. All other sensitivities are found to be less than or equal to unity, some approaching zero as Q increases. With $\mu = \mu' \leq 10^5$, L_{mn} and Q_{mn} sensitivities to respective μ's also approach zero provided Ψ is not very small and $Q^2 \ll \mu/(2 + 1/\Psi)$. In practice, $\Psi = 1$, i.e. $r_2 = R_1$ is a good choice because then one can fabricate the circuit with equal resistors and equal capacitors.

Compared to the earlier known circuits for floating inductor simulation, the circuit just discussed has the following attractive features : low component count; unity spread in resistors, capacitors and active elements; easy adjustment; low sensitivities; absolutely stable configuration; UG connection of OA's leading to widest possible frequency range of operation; and easy conversion to grounded inductance simulator.

Instead of the Prestcott circuit, one could also start from Berndt–Dutta Roy or Rao–Venkateswaran circuit and obtain floating-L simulations. These circuits have been studied by Ahmed [20] together with the frequency limitations of all the three circuits. The Prestcott derived circuit is found to have a superior high-frequency performance as compared to others. However, all three of them are found to be useful up to only a small fraction of the gain bandwidth product of the OA's.

The approach discussed here for floating-L simulation has been independently reported by Wise [21].

12.6 Applications of Simulated L in the Design of Insensitive Biquadratic Sections

Leach and Chan [22] proposed a method of insensitive biquadratic design using a perfect simulated L. Dutta Roy and Ahmed [23] showed that one could use a lossy simulated L in this situation with the low-sensitivity feature maintained, and with the added advantage of lower component count. Consider, for example, the circuit of Fig. 12.16, where the (r, L) combination is derived using the Prestcott circuit of Fig. 12.6a. Note that

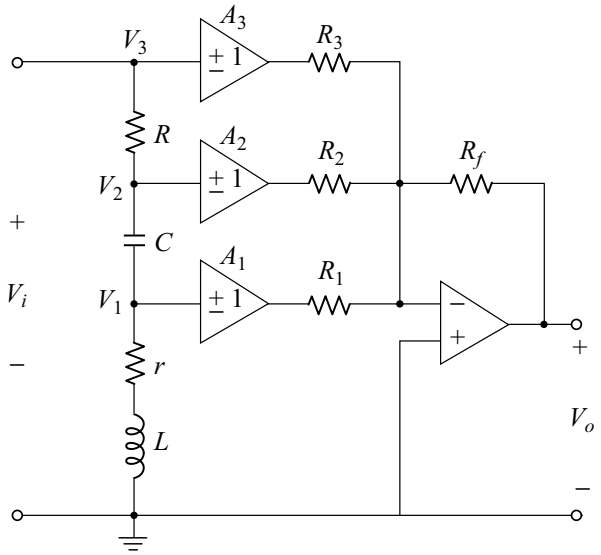


Fig. 12.16 Insensitive biquad scheme using an L-simulator

$$
\begin{aligned}
V_3 &= V_{\text{in}} \\
V_2 &= V_{\text{in}}\left[s^2 + sr/L + 1/(LC)\right]/\left[s^2 + s(r+R)/L + 1/LC\right)] \\
V_1 &= V_{\text{in}}\left(s^2 + sr/L\right)/\left[s^2 + s(r+R)/L + 1/(LC)\right]
\end{aligned} \tag{12.33}
$$

and that

$$
V_0 = -\sum_{i=1}^{3} V_i A_i R_f / R_i \tag{12.34}
$$

Denoting—$A_i\ R_f/R_i$ by a_i, and substituting for V_i/V_{in} from Eq. (12.33), it is easy to show that the transfer function $H(s) = V_0/V_{\text{in}}$ is given by

$$
H(s) = \left\{s^2(a_1 + a_2 + a_3) + s[r(a_1 + a_2 + a_3) + Ra_3]/L + (a_2 + a_3)/(LC)\right\}/\left[\left(s^2 + s(r+R)/L + 1/(LC)\right]\right. . \tag{12.35}
$$

Choosing a_i's, i.e. R_i's and A_i's appropriately, one obtains all types of second order transfer functions. For example,

$R_2 = \infty, R_1 = R_2, A_1 = -A_2 = \pm 1$ gives low-pass;
$R_2 = R_3 = RR_1/r, A_1 = A_2 = -A_3 = \pm 1$ gives high-pass;
$R_1 = \infty, R_2 = R_3, A_2 = -A_3 = \pm 1$ gives bandpass;
$R_1 = \infty, R_2 = R_3[1 + R/(r)], A_2 = -A_3 = \pm 1$ gives bandstop
and
$R_1 = \infty, R_2 = R_3[1 + R/(2r)], A_2 = -A_3 = \pm 1$ gives all-pass

filters. The Q and ω_0 sensitivities can be calculated in the usual manner and it is found that the passive sensitivities are all less than unity, some of them being zero and some 1/2. The multiparameter sensitivities are

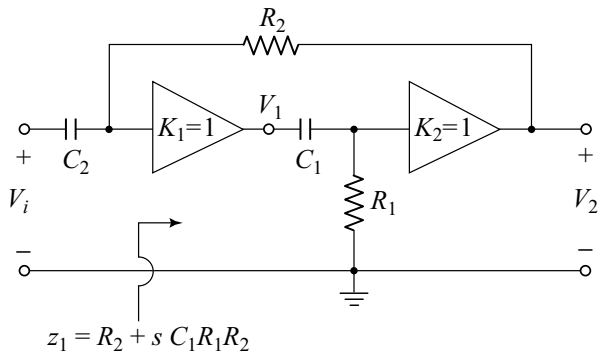


Fig. 12.17 Dutta Roy's insensitive biquad scheme using the circuit of Fig. 12.8

$$\sum\nolimits_{\omega_0} = \sum\nolimits_i \left|S_{X_i}^{\omega_0}\right| = 2 \text{ and } \sum\nolimits_Q < 3. \tag{12.36}$$

A different method of insensitive biquadratic design with a simulated L has been recently proposed by Dutta Roy [24]. The basic structure for this is the Rao–Venkateswaran (*loc. cit.*) circuit of Fig. 12.8 with an additional capacitor C_2 added at the input, as shown in Fig. 12.17. From the fact that $Z_i = R_2 + sC_1\,R_1\,R_2$, it is easy to show that

$$V_{1;z} = V_i\left[s^2 + s/\left(C_1R_1, s^2\right]/\left[s^2 + \mathrm{s}/(C_1R_1) + 1/(C_1R_1C_2R_2)\right].\right.$$

By appropriate linear combinations of $V_{i,1,2}$ in another OA, one can again obtain any specified biquadratic transfer function. The advantage with respect to the previous scheme is that no isolating UGA is needed in this case, thus leading to component saving. For example, the basic structure of Fig. 12.16 is a high-pass filter with V_2, as the output; comparing with the previous high-pass circuit, we get a saving of three OAs and five resistors.

12.7 Conclusion

From the account of some recent developments given here, it can be concluded that L simulation approach to active filter design is a fairly developed technique, and can be used for biquadratic or higher order filter designs with the important advantages already mentioned. Designing with lossy, but high Q simulated inductors create no additional problems as compared to the design of doubly terminated LC filters with lossy coils.

References

1. H.J. Orchard, Inductorless filters. Electron. Lett. **2**, 224–225 (1966)
2. R.H.S. Riordan, Simulated inductors using differential amplifiers. Electron. Lett. **3**, 50–51 (1967)
3. A. Antoniou, Realization of gyrators using operational amplifiers and their use in RC-active network synthesis. Proc. lEE. **166**(11), 1839–1850 (1969)
4. T.N. Rao, P. Gary, R.W. Newcomb, Equivalent inductance and Q of a capacitor loaded gyrator. IEEE J. SC **2**(1), 32–33 (1967)
5. S.C. Dutta Roy, On operational amplifier simulation of a grounded inductance. Arch. Elektr. Ubertragung. **29**(3), 107–115 (1975)
6. S.C. Dutta Roy, Bibliography of inductance simulation by active RC methods. Microelectron. Reliab. **15**, 637–639 (1976)
7. M.P. Beddoes, K.R. Morin, Bibliography on inductance simulation by gyrator methods. IEEE Trans. CT **14**(1), 107–111 (1967)
8. S.K. Mitra (ed.), *Active Inductorless Filters* (IEEE Press Book, New York, 1971), pp. 215–219
9. J.A. Miller, R.W. Newcomb, An annotated bibliography on gyrators in network theory: circuits and uses. Technical Report No. R-**72**-01 (Elec. Eng. Dept., University of Maryland, USA, 1972)
10. A.J. Prestcott, Loss compensated active gyrator using differential input operational amplifier. Electron. Lett. **2**, 283–284 (1966)
11. D. Berndt, S.C. Dutta Roy, lnductor simulation using a single unity gain amplifier. IEEE J. SC **4**, 161–162 (1969)
12. R.L. Ford, F.E.J. Girling, Active filters and oscillators using simulated inductance. Electron. Lett. **2**, 52 (1966)
13. K.R. Rao, S. Venkateswaran, Synthesis of inductors and gyrators with voltage controlled voltage sources. Electron. Lett. **6**, 29–30 (1970)
14. A.C. Caggiano, Operational amplifier simulates inductance. Electron **41**, 99 (1968)
15. J. Musiedlak, J. McPherson, Some inductance simulators utilizing a single differential operational amplifier*, in Proceedings of the 16th Midwest Symposium on Circuit Theory* (1973), pp. IX. 8.1-IX.8.10
16. R.P. Sallen, E.L. Key, A practical method of designing RC active filters. IRE Trans. CT **2**, 74–85 (1955)
17. P.R. Geffe, RC amplifier resonators for active filters. IEEE. Trans. CT **15**(4), 415–419 (1968)
18. M.T. Ahmed, S.C. Dutta Roy, Evaluating the frequency limitations of operational amplifier RC networks: the dominant pole technique and its application to inductance simulators, in *Proceedings of the 18th Midwest Symposium on Circuits and Systems*, Paper No. R 19.8 (1975)
19. S.C. Dutta Roy, A circuit for floating inductance simulation. Proc. IEEE **62**, 521–523 (1974)
20. M.T. Ahmed, Active RC synthesis of single and multivariable network functions using operational amplifiers, Ph.D. thesis, IIT, New Delhi, 1974
21. B. Wise, Active simulation of floating lossy inductances. Proc. IEEE **121**, 85–87 (1974)
22. D.P. Leach, S.P. Chan, A generalized method of active RC synthesis. IEEE. Trans. CT **18**, 643–650 (1971)
23. S.C. Dutta Roy, M.T. Ahmed, Synthesis of RC active networks using non-ideal simulated inductance. IEEE Trans. Circuits Syst. **21**, 250–257 (1974)
24. S.C. Dutta Roy, A versatile second order active filter. Electron. Eng. **46**, 64–68 (1974)

Chapter 13
Evaluating the Frequency Limitations of Operational Amplifier RC Networks by the Dominant-Pole Technique

Abstract A general computer-based technique is given for investigating the high-frequency performance of second-order active RC networks employing an arbitrary number of operational amplifiers. It uses the dominant-pole technique for finding deviations in Q and ω_0 from their nominal values as a function of the gain bandwidth product of the operational amplifiers. The technique is applied to active RC circuits simulating a nonideal grounded inductance and curves are given for the highest usable frequency for permissible deviation in Q and ω_0 for some typical circuits.

Keywords Active RC filters · Dominant-pole technique · Frequency limitations

13.1 Introduction

The actual performance of an operational amplifier (OA)–RC network differs considerably from that predicted by conventional design, based on the assumption of ideal OA, due to two major reasons, viz. (i) the sensitivity of the network function to the variations of active and passive components from their nominal values and (ii) frequency limitations introduced by the practical OA. Sensitivity has been a subject of extensive research in the past two decades, and well-established methods are now available for its evaluation and minimization [1]. Frequency limitation due to OA is a relatively recent topic; some of the notable contributions in this direction are due to Budak and Petrela [2, 3], and Faulkner and Grimbleby [4, 5]. These are based on one-pole roll-off model of the OA and are particularly suitable for circuits using one OA. Applied to multiple OA networks, both the methods have some inherent limitations; in the method of Budak and Petrela [6], for example, one may have to sacrifice accuracy by neglecting the fourth and higher order terms in the denominator of the transfer function. In the Faulkner–Grimbleby method [5], the developed procedure is to be repeated as many times as are the number of OA's in the circuit.

Source: Muslim T. Ahmed & S.C. Dutta Roy, "Evaluating the Frequency Limitations of Operational Amplifier RC Networks by the Dominant-Pole Technique and Its Applications to Inductance Simulation," *JIETE*, vol 22, pp 703–708, November 1976.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_13

In the present chapter, a general technique, based on the dominant-pole approximation, is developed for the study of a second-order active RC network using an arbitrary number, n, of OA's. The technique is applied to low-component count, non-ideal, grounded, inductance simulator (IS) circuits reported by Dutta Roy [7]. The effect of finite gain bandwidth product (η) of the OAs has been displayed in the form of curves giving the highest usable frequency (ω_h) for prescribed deviations in Q and ω_0 from their nominal values.

13.2 Dominant-Pole Technique

Consider an active RC network containing n-OAs which realizes, nominally, the second-order transfer function

$$T(s) = \frac{N(s)}{D(s)} = \frac{N(s)}{s^2 + a_1 s + a_0}. \tag{13.1}$$

The degree of $N(s)$ is 0, 1 or 2 depending upon the type of filter desired. The two most important performance criteria characterizing Eq. (13.1) are

$$Q = (\sqrt{a_0})/a_1 \text{ and } \omega_0 = \sqrt{a_0}. \tag{13.2}$$

Identical OAs are assumed, each having an open-loop gain function

$$A(s) = -\frac{A_0\omega_a}{s + \omega_a} = -\frac{\eta}{s + \omega_a} \tag{13.3}$$

where $A_0 = A(0)$; ω_0 is the 3 dB open-loop bandwidth of the OA and $\eta = A_0\omega_a$ stands for the gain bandwidth product. When Eq. (13.3) is substituted for each OA gain in the n-OA circuit being considered, it is obvious that the denominator polynomial of the resulting transfer function will be changed to an $(n + 2)$th-order polynomial. Denoting this by $D'(s)$,

$$D'(s) = B_{n+2}s^{n+2} + B_{n+1}s^{n+1} + \cdots + B_1 s + B_0 \tag{13.4a}$$

$$= (s^2 + a_1' s + a_0')(b_n s^n + b_{n-1}s^{n-1} + \cdots + b_1 s + b_0) \tag{13.4b}$$

where, nominally (i.e. when $A_0 = \infty$, $\omega_a = \infty$), $a_1' = a_1$, $a_0' = 0$, $b_n = b_{n-1} = \ldots = b_1 = 0$ and $b_0 = 1$. The quadratic factor in Eq. (13.4b) represents the effect of perturbation of the quadratic $D(s)$ due to finite η of the OA's and is the dominant complex root pair of $D'(s)$ (Ghausi [7]; Millman and Halkias [8]). If this factor can be found out, then the effective Q and ω_0 denoted by Q_e and ω_{0e}, respectively, can be calculated as

$$Q_e \simeq \left(\sqrt{a_0'}\right)/a_1', \ \omega_{0e} \simeq \sqrt{a_0'} \tag{13.5}$$

The percentage deviations in Q and ω_0 are, therefore,

$$\Delta Q = 100(Q_e - Q)/Q, \Delta\omega_0 = 100(\omega_{0e} - \omega_0)/\omega_0. \tag{13.6}$$

A computer program FLAC (Frequency Limitations of Active Circuits) has been written to evaluate the dominant quadratic factor of $D'(s)$ and hence ΔQ and $\Delta\omega_0$ for a large number of second-order circuits. The results for grounded inductance simulators will be presented in the next section. A flow diagram for FLAC is given in Fig. 13.1.

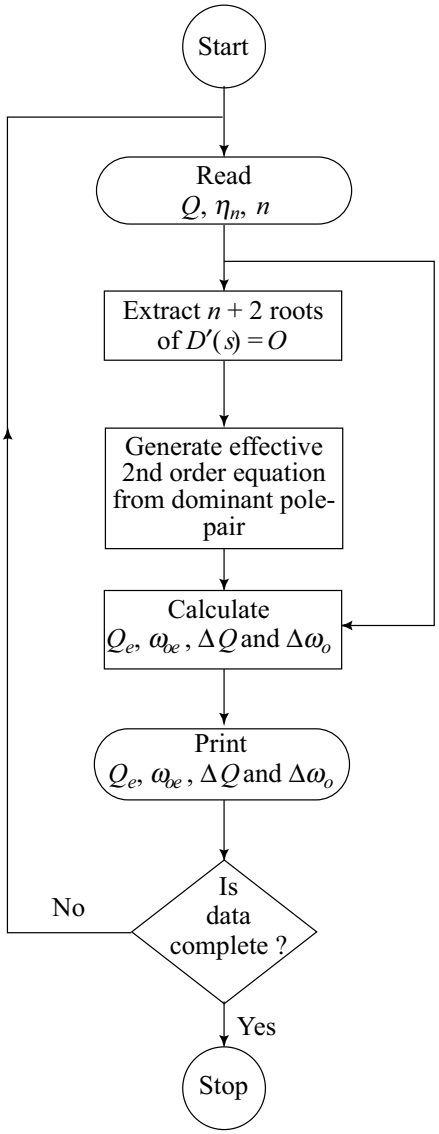


Fig. 13.1 Flowchart of FLAC

13.3 Frequency Limitations of Lossy Grounded Inductance Simulators

Figure 13.2 shows the various circuits for simulating a grounded nonideal inductance, as considered by Dutta Roy [6], where each voltage amplifier is realized by an OA, with or without suitable resistances. Each of these circuits, assuming a frequency-independent OA gain, realizes a bilinear RL-impedance of the form

$$Z(s) = H(s\tau_1 + 1)/(s\tau_2 + 1),\ \tau_1 \geq \tau_2 \tag{13.7}$$

For evaluating the high-frequency performance, we convert each inductance simulator to a tuned circuit by connecting a capacitance C_t across the input terminals. Then, nominally, the quality factor Q_t and the resonance frequency ω_t of the tuned circuit are given by

$$Q_t = (HC_t\tau_1)^{1/2}/(HC_t + \tau_2) \tag{13.8}$$

$$\omega_t = (HC_t\tau_1)^{-1/2} \tag{13.9}$$

The values of H, τ_1, τ_2, Q_t, ω_t and the design equations and constraints for each circuit have been given by Dutta Roy [6]. Taking the frequency dependence of OA gain into account, the polynomial $D'(s)$, which determines effective Q_t and ω_t becomes of the form of Eq. (13.4a). When this is equated to zero, the resulting equation can be written in the form

$$s^{n+2} + c_{n+1}, s^{n+1} + \cdots + c_1 s + c_0 = 0 \tag{13.10}$$

where $c_j = B_j/B_{n+2}$. The values of the coefficients are given in Table 13.1 for each circuit; these are functions of Q_t and the normalized gain bandwidth product $\eta_n = \eta/\omega_t$ both forming the input data to FLAC. The coefficients are used for evaluating the dominant complex root pair of Eq. (13.10) and hence finding Q_{te} and ω_{te} and ΔQ_t and $\Delta\omega_t$ for a given range of Q_t and η_n.

Computations using the dominant-pole technique (DPT) were carried out for a large range of Q_t and η_n. High values of Q, in general, lead to large element spread, and hence are not realizable if the circuit is to be fabricated in IC form; nevertheless, Q values up to 200 have been taken to demonstrate the frequency limitations at high Q's. Some of the results are shown in Table 13.2, which gives the maximum η_{n}^{-1} corresponding to a given Q_t and percentage deviation in either Q_t or ω_t. If all the three quantities Q_t, ΔQ_t and $\Delta\omega_t$ are specified, then one has to use the lower of the two values of η_n^{-1} corresponding to $(Q_t, \Delta Q_t)$ and $(Q_t, \Delta\omega_t)$. The highest usable frequency for a particular OA being used in the circuit can easily be found from

$$\omega_h = \eta \times \eta_n^{-1}. \tag{13.11}$$

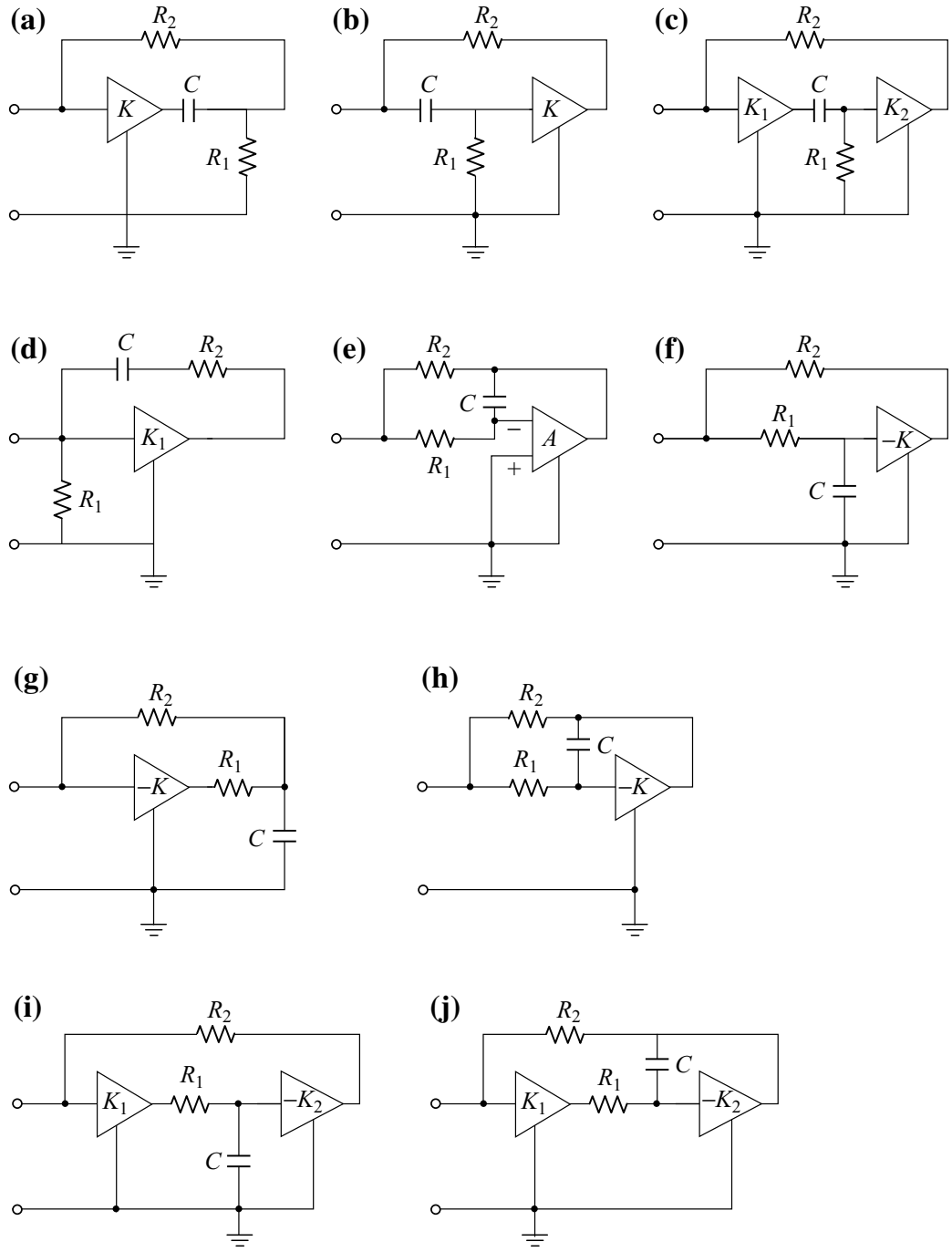


Fig. 13.2 Nonideal grounded inductors obtained with OA's in: Positive-gain connection **a** Prestcott circuit, **b** Berndt–Dutta Roy circuit, **c** Rao–Venkateswaran circuit, **d** Caggiano circuit: Infinite-gain connection, **e** Ford–Girling circuit; Negative-gain connection—circuits derived from **f** Sallen–Key's second-order active filter, **g** by shifting the RC network, **h** by returning capacitor to amplifier output; Positive and negative-gain connection—circuits derived from **i** Geffe's Q-invariant resonator, **j** Sanderson's Q-and ω_0-invariant resonator

Table 13.1 Coefficients of Eq. (13.10) for the various circuits of Fig. 13.2

Circuits referred to in Fig. 13.2	Coefficient (c_t)
A, B	$c_0 = \eta_n,\ c_1 = \eta_n/Q_t + 1,\ c_2 = \eta_n + \left(1 + 2Q_t^2\right)/Q_t$
C	$c_0 = \eta_n^2,\ c_1 = \eta_n^2/Q_t + 2\eta_n,\ c_2 = \eta_n^2 + 2\eta_n\left(1 + Q_t^2\right)/Q_t + 1,\ c_2 = 2\eta_n + \left(1 + Q_t^2\right)/Q_t$
D	$c_0 = \eta_n/2,\ c_1 = \eta_n/(2Q_t) + 1,\ c_2 = \eta_n/2 + \left(1 + 2Q_t^2\right)/(2Q_t)$
E	$c_0 = -\eta_n, c_1 = 1 - \eta_n/Q_t + 2Q_t\omega_{an} \cdot c_2 = \left(1 + 2Q_t^2\right)/Q_t - \eta_n$
F, G	$p \triangleq 1 + 4Q_t^2;\ c_0 = \eta_n/p$ $c_1 = \eta_n(2p - 1)/\left(2Q_t p^2\right) + 1/p$ $c_2 = [\eta_n + (2p - 1)/(2Q_t)]/p$
H	$p \triangleq 1 + 4Q_t^2;\ c_0 = \eta_n$ $c_1 = \left[p/(2Q_t)^3 + 2Q_t/p\right]\eta_n + 1,$ $c_2 = \eta_n + p/(2Q_t)^3 + 2Q_t]$
I	$p \triangleq 4Q_t^2;\ c_0 = \eta_n^2/p,\ c_1 = \eta_n(4\eta_n Q_t + p + 1)/p^2,$ $c_2 = \left(\eta_n^2 + 1\right)/p + 4\eta_n Q_t(p + 1)/p^2,$ $c_3 = [\eta_n(p + 1) + 4Q_t]/p$
J	$p \triangleq 4Q_t^2;\ c_0 = \eta_n^3,\ c_1 = \eta_n^2\left[\eta_n/(2Q_t) + 2\eta_n Q_t/p + p + 1 + 1/\mathrm{p}\right],$ $c_2 = \eta_n 2[\eta_n + (p + 2)/(2Q_t) + 2Q_t(2p + 1)/p] + \eta_n(p + 2)$ $c_3 = \eta_n[\eta_n(p + 2) + (2p + 1)/(2Q_t) + 2Q_t(p + 2)] + p,$ $c_4 = p[\eta_n + (p + 1)/(2Q_t)] + \eta_n(p + 1)$

In this sense, Table 13.2 gives universal values for the calculation of ω_h. A close glance at the Table indicates that, basically, the circuits of Fig. 13.2 can be classified into three groups depending upon their operational frequency range. Circuits characterized by medium and high-Q realizations at high frequencies will be called class I. Class II circuits are the ones for low Q (<25) realizations at low frequencies. Circuits showing a variable frequency range of operation will be called class III; their range depends on the ratio $\omega_{an} = \omega_a/\omega_t$. For low ω_{an}, these realizations behave as class I circuits, whereas for high ω_{an} they fall in class II circuits. For a typical representative of each class, universal curves will be given which show the variations of η_n^{-1} with Q_t at various $\Delta\omega_t$ and ΔQ_t as parameter. From these curves, ω_h can easily be calculated by using Eq. (13.11).

It is evident from Table 13.2 and Fig. 13.2 that the active inductors using non-inverting OA belong to class I. Using commercial OAs (with external compensation), the computer results show that these circuits are suitable for high-Q applications in the range of 500 kHz within 5% allowed deviations in Q_t and ω_t. Table 13.2 shows that Fig. 13.2c, using two non-inverting OAs, has the best performance amongst all

Table 13.2 η_n^{-1} (maximum) for calculating the highest usable frequency of inductance simulators

Inductance simulators	Q of the tuned of circuit	Values of η_n^{-1} (max.) for given Q, and permissible percentage deviation in						Class of IS
		Q_t			ω_t			
		1	5	10	1	5	10	
a, b	5	2.0 (–3)	1.1 (–2)	3.1 (–2)	2.0(–1)	1.1 (–2)	2.3 (–2)	I
	20	4.1 (–4)	2.4 (–3)	7.0 (–3)	4.1 (–4)	2.5 (–3)	4.7 (–3)	
	200	2.0 (–5)	2.4 (–4)	7.1 (–4)	1.8 (–5)	2.5 (–4)	4.8 (–4)	
c	5	3.0 (–3)	1.2 (–2)	1.0 (–1)	2.7 (–3)	1.3 (–2)	2.4 (–2)	I
	20	4.7 (–4)	2.9 (–3)	4.0 (–2)	4.3 (–4)	2.7 (–3)	4.7 (–3)	
	200	4.1 (–5)	3.3 (–4)	3.4 (–3)	4.2 (–5)	2.7 (–4)	5.0 (–4)	
d	5	2.1 (–3)	1.4 (–2)	1.2 (–1)	2.0 (–3)	1.1 (–2)	2.3 (–2)	I
	20	2.3 (–4)	3.6 (–3)	3.0 (–2)	2.3 (–4)	2.5 (–3)	4.8 (–3)	
	200	4.2 (–5)	3.6 (–4)	3.0 (–3)	4.0 (–5)	2.5 (–4)	5.0 (–4)	
$e(\omega_{an} = 10^{-2})$	5	2.0 (–3)	8.0 (–3)	1.3 (–2)	2.0 (–3)	9.1 (–2)	1.7 (–2)	III
	20	4.6 (–4)	2.3 (–3)	5.2 (–3)	4.2 (–4)	2.3 (–3)	1.4 (–3)	
	200	6.0 (–5)	3.1 (–4)	6.7 (–4)	4.0 (–5)	2.2 (–4)	5.4 (–4)	
f, g	5	1.0 (–5)	8.0 (–5)	1.4 (–4)	1.4 (–3)	2.8 (–3)	4.0 (–3)	II
	20	–	1.8 (–6)	2.8 (–6)	9.5 (–5)	2.4 (–4)	3.5 (–4)	
	200	–	–	–	–	–	–	
h	5	2.1 (–3)	1.2 (–2)	3.0 (–2)	2.0 (–3)	1.1 (–2)	2.2 (–2)	I
	20	4.1 (–4)	2.8 (–3)	6.7 (–3)	4.0 (–4)	2.3 (–3)	1.0 (–2)	
	200	4.1 (–5)	2.8 (–4)	6.7 (–4)	4.0 (–5)	2.3 (–4)	1.0 (–3)	
i	5	2.0 (–5)	1.0 (–4)	1.9 (–4)	1.4 (–3)	3.4 (–3)	5.0 (–3)	II

(continued)

Table 13.2 (continued)

Inductance simulators	Q of the tuned of circuit	Values of η_n^{-1} (max.) for given Q, and permissible percentage deviation in						Class of IS
		Q_t			ω_t			
		1	5	10	1	5	10	
	20		1.4 (–6)	2.2 (–6)	9.1 (–5)	2.2 (–4)	3.5 (–4)	
	200	–	–	–	–	–	–	
j	5	1.1 (–3)	5.4 (–3)	1.4 (–2)	2.0 (–3)	1.1 (–2)	2.2 (–2)	I
	20	2.5 (–4)	1.3 (–3)	2.5 (–3)	5.0 (–4)	2.7 (–3)	5.0 (–3)	
	200	2.5 (–5)	1.5 (–4)	6.4 (–4)	5.7 (–5)	5.0 (–4)	5.0 (–4)	

Note a(b) stands for 'a' multiplied by 10^b. A blank indicates that these inductance simulators are unsuitable for realizing the corresponding specifications

the active inductances investigated; the frequency limitation curves for this circuit are given in Fig. 13.3.

As expected, the effect of η on circuits using inverting OA is quite appreciable, especially, on the Q of the circuits. The DPT results indicate these realizations to be suitable only for low frequency, low Q (<20) applications, inspite of their excellent sensitivities [6]; consequently, these are class II active inductors. A typical representative of this class is Fig. 13.2f, and its frequency limitation curves are also given in Fig. 13.3. An exception to the inverting OA circuits is provided by Fig. 13.2h. The feedback via the condenser seems to improve the high-frequency performance considerably and makes it a class I instead of class III circuit.

Figure 13.2i, j, using one inverting and one non-inverting OA, show a marked difference from each other in their frequency limitation characteristics. Whereas the first circuit is found to behave as class II, the second one suits more to the family of class I.

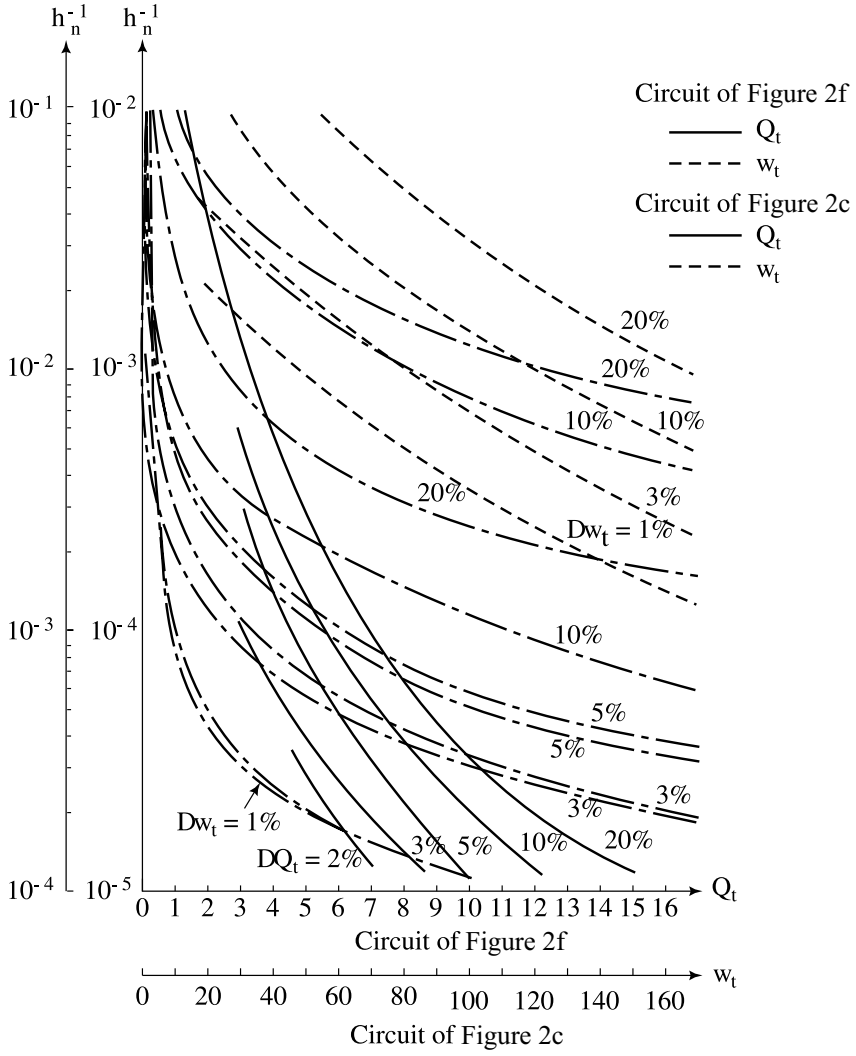


Fig. 13.3 Maximum $\eta_n^{-1}/(\omega_h)$–curves for allowable deviations in Q_t and ω_t for inductance simulators of Fig. 13.2c (Class I) and Fig. 13.2f (Class II)

Figure 13.2e, using OA in the infinite-gain connection is a class III active inductor. Here the coefficient c_1, being a function of the parameter ω_{an} (see Table 13.1), makes the frequency limitation characteristics variable and critically dependent on this parameter. Investigation with ω_{an} varying between 10 and 10^{-3} have been made to cover the practical frequency range of commercially available OA's. Figure 13.4 gives universal curves for this circuit with ω_{an} as parameter; these are similar to the curves of class II circuits at high ω_{an} ($\approx$10). For low ω_{an} ($\approx 10^{-3}$), the frequency limit improves considerably and the characteristics depicted are those of class I circuits. Moreover, the effect of ω_{an} variation is more prominent on the Q_t deviations as compared to ω_t deviations, the latter being noticeable only at high $\Delta\omega_t$ and low Q_t values.

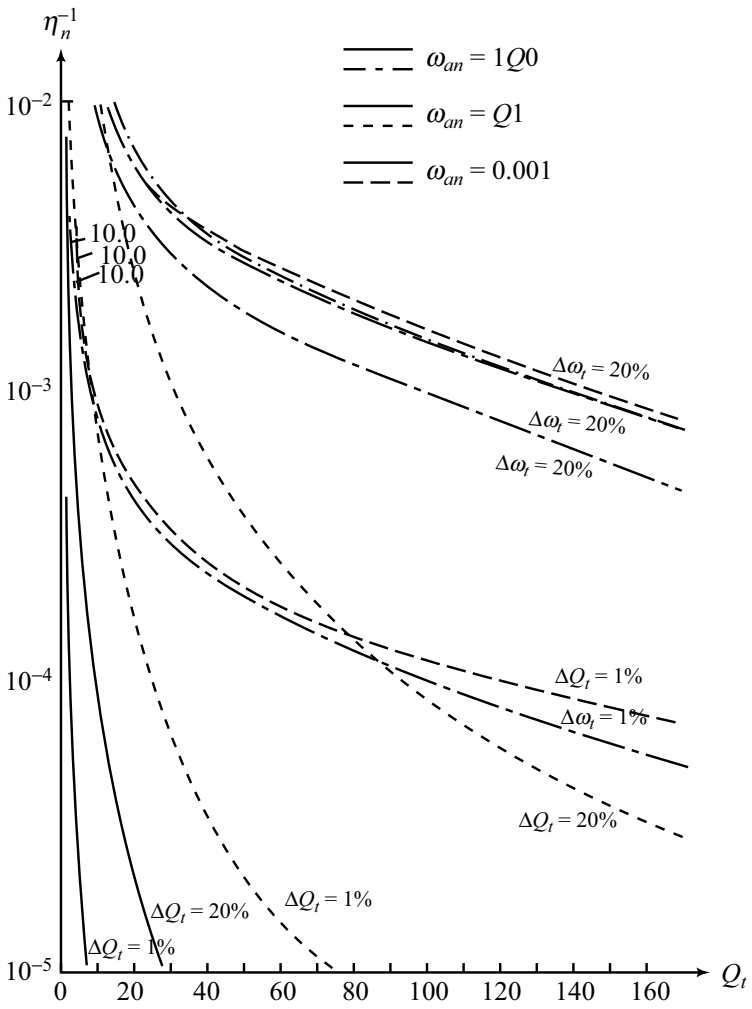


Fig. 13.4 Maximum $\eta_n^{-1}/(\omega_h)$–curves for allowable deviations in Q_t and ω_t for Fig. 13.2e (Class III)

13.4 Experimental Investigations

Experimental investigations were conducted on a bandpass filter circuit (Dutta Roy and Ahmed [9]), shown in Fig. 13.5a, which is, basically, the circuit of Fig. 13.2b, with a low-pass RC section cascaded at the input. The filter was designed and fabricated to realize a nominal Q of 10 and variable ω_0. Internally compensated operational amplifier SMC 741 was intentionally chosen to be of poor quality ($\eta \simeq 0.7$ MHz) with a view to avoid the difficulties and inaccuracies encountered in the measurements at high frequency.

The experimental curves along with the ones obtained by using Budak–Petrela [3] method and the dominant-pole technique are shown in Fig. 13.5b. They show the percentage deviations in Q and ω_0 from the nominal values at various frequencies. It is observed that there is a close agreement between the practical values and $\Delta\omega_0$-curves obtained by the DPT. However, the ΔQ-curves obtained from the DPT are optimistic, whereas, both the $\Delta\omega_0$ and ΔQ-curves predicted by Budak–Petrela method are pessimistic. It is important to note that the filter uses only one OA for

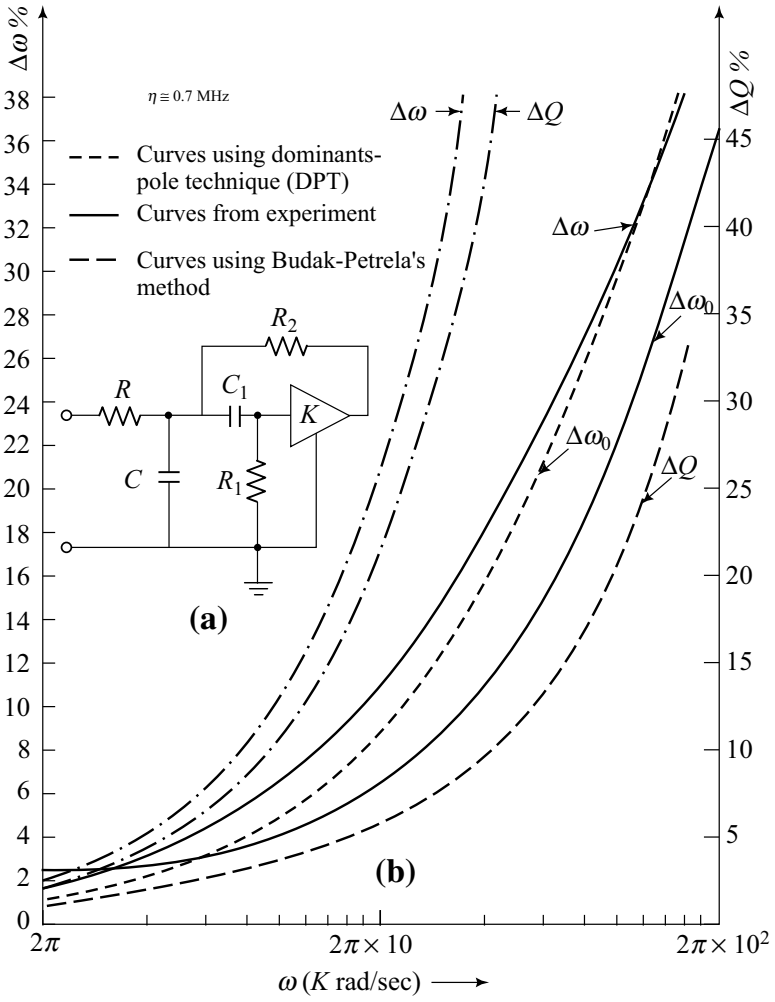


Fig. 13.5 **a** Bandpass filter, **b** comparative frequency limitation curves

which Budak–Petrela theory has least approximation (because there is no truncation of terms). But Fig. 13.5b indicates a closer agreement between practical results and the ones obtained by using DPT as compared to those obtained by Budak–Petrela method. As the DPT is equally accurate for multiple OA networks, it can logically be presumed that when used on circuits having more than one OA, it will continue to give closer results than those obtained from Budak–Petrela method.

13.5 Conclusion

A dominant-pole technique has been presented for the frequency limitations of second-order active RC networks using n operational amplifiers. The method is applied to nonideal grounded inductance simulators and based on their frequency limitation characteristics, they are classified into three groups of active inductances. The frequency behaviour and, hence, classifications are largely influenced by the mode in which the OA is used. A comparative frequency limitation study of a band-pass filter was made using the Budak–Petrela method and the DPT. The experimental results indicate a closer correlation with the dominant-pole technique.

References

1. S.K. Mitra (ed.), *Active Inductorless Filters* (IEEE Press, New York, 1971), pp. 205–208
2. D.M. Petrela, A. Budak, Open-loop gain sensitivity of active filters using a single finite bandwidth voltage amplifier, in *Proceedings of the 14th Midwest Symposium on Circuit Theory*, Paper 14.2 (1971)
3. A. Budak, D.M. Petrela, Frequency limitations of active filters using operational amplifiers. IEEE Trans. CT-**19**, 322–328 (1972)
4. E.A. Faulkner, J.B. Grimbleby, Active filters and gain-bandwidth product. Electron. Letts. **6**, 549–550 (1970)
5. E.A. Faulkner, J.B. Grimbleby, The effect of amplifier gain-bandwidth on performance of active filters. Radio Electron. Eng. **43**, 347–352 (1973)
6. S.C. Dutta Roy, On operational amplifier simulation of grounded inductance. Arch. Elektr. Ubertragung. **29**, 107–115 (1975)
7. M.S. Ghausi, *Electronic Circuits* (Van-Nostrand Reinhold, New York, 1971), pp. 216–220
8. J. Millman, C.C. Halkias, *Integrated Electronics* (McGraw Hill, New York, 1972), p. 382
9. S.C. Dutta Roy, M.T. Ahmed, Second-order active RC band-pass filters using operational amplifier in low gain connection. JIETE. **21**, 248–252 (1975)

Chapter 14
'Shadow' Filters—A New Family of Electronically Tunable Filters

Abstract It is shown that the 'shadow' filter, introduced recently by Lakys and Fabre, is capable of achieving not only a constant bandwidth bandpass filter, but also several other interesting characteristics depending upon the choice of the desired output port and the port from which feedback signal is derived. Two such characteristics are illustrated in this chapter, namely constant centre frequency, variable Q, and constant value of the product of centre frequency and Q. A problem is also posed for further investigation. Finally, a suggestion is made for the design of a simple but interesting experiment for undergraduate students, using only resistors, capacitors and op-amps, which illustrates the basic concepts of filters and what the shadow technique can do to improve their quality.

Keywords Shadow filters · Versatility · Open problems

14.1 Introduction

In their 2010 paper, Lakys and Fabre [1] introduced a new family of second-order filters, called 'shadow' filters for electronically tuning the characteristics of a basic filter. A general block diagram of the scheme is shown in Fig. 14.1, in which the authors of [1] considered $H(s)$ to be a bandpass transfer function while $H_1(s)$ is a low-pass transfer function, both being of second order with the same denominator. They show that thereby the characteristics of the modified transfer function $H'(s) = V_0/V_{\text{in}}$ retains the bandpass characteristic, but with variable centre frequency and variable Q, in such a manner that the bandwidth remains a constant.

The purpose of this chapter is to make a more general analysis of the scheme and to demonstrate that by properly choosing $H(s)$ and $H_1(s)$, various types of interesting characteristics of practical importance can be obtained. Two such characteristics are illustrated in this chapter, namely constant centre frequency, variable Q and constant value of the product of centre frequency and Q. A problem is also posed for further investigation.

Source: S.C. Dutta Roy, "'Shadow' Filters—A New Family of Electronically Tunable Filters," *IETE Journal of Education*, vol 51, pp 75–78, May–December 2010.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_14

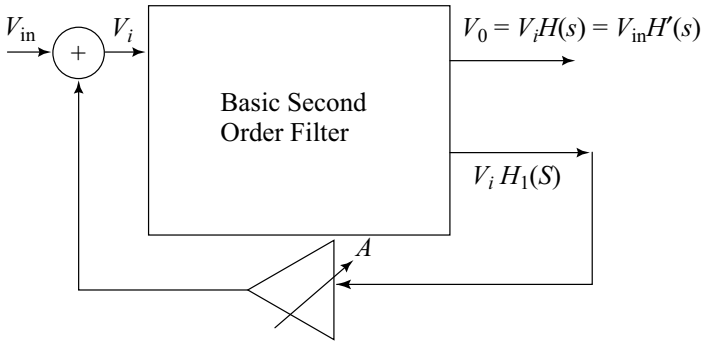


Fig. 14.1 Schematic diagram of the shadow filter

Finally, a suggestion is made for the design of a simple but interesting experiment for undergraduate students, using only resistors, capacitors and op-amps, which illustrates the basic concepts of filters and how the shadow technique can be used to improve their quality.

14.2 Analysis

Consider a basic second-order filter, which can be active or passive, and which is capable of realizing any of the four possible characteristics, viz. low-pass, high-pass, bandpass or band-stop or a combination of these. Such realizations can, for example, be a simple series *RLC* circuit [2] or an active biquad [3] or a simple combination of RC filters connected in parallel at the input as shown in Fig. 14.2. Similarly, the feedback signal is taken from any of these outputs or a combination of them. Thus, $H(s)$ and $H_1(s)$ are of the forms

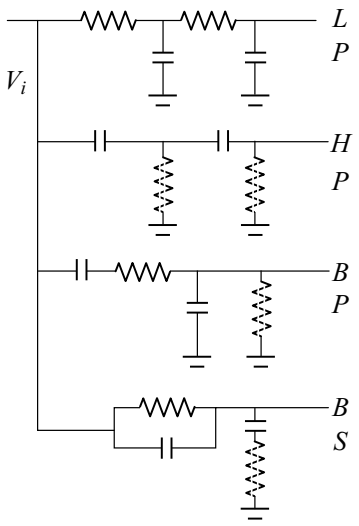


Fig. 14.2 A passive RC basic filter. All resistors are of value *R* and all capacitors are of value *C*

$$H(s) = \frac{d + es + fs^2}{s^2 + (\omega_0/Q)s + \omega_0^2} \tag{14.1}$$

and

$$H_1(s) = \frac{a + bs + cs^2}{s^2 + (\omega_0/Q)s + \omega_0^2} \tag{14.2}$$

By analysis of Fig. 14.1, we get

$$H'(s) = \frac{H(s)}{1 - AH_1(s)} \tag{14.3}$$

Combining Eqs. (14.1), (14.2) and (14.3) gives

$$H'(s) = \frac{d + es + fs^2}{s^2(1 - Ac) + [(\omega_0/Q) - Ab]s + \omega_0^2 - Aa} \tag{14.4}$$

This can be rewritten as

$$H'(s) = \frac{d' + e's + f's^2}{s^2 + (\omega_0'/Q')s + \omega_0'^2} \tag{14.5}$$

where

$$(d', e', f') = (d, e, f)/(1 - Ac);$$

$$\omega_0' = \sqrt{(\omega_0^2 - Aa)/(1 - Ac)};$$

$$Q' = \sqrt{(1 - Ac)(\omega_0^2 - Aa)}/(\omega_0 - AbQ) \tag{14.6}$$

14.3 Special Cases

In [1], the authors considered $b = 0$ and $c = 0$ (it would be wrong to write $b = c = 0$ because b and c are dimensionally different) so that

$$\omega_0' = \omega_0\sqrt{(1 - Aa/\omega_0^2)}; \qquad Q' = \sqrt{(1 - Aa/\omega_0^2)} \tag{14.7}$$

Thus the centre frequency and Q change proportionately and the bandwidth $\Delta\omega' = \omega_0'/Q'$ remains the same as the original bandwidth $\Delta\omega = \omega_0/Q$.

An interesting variation of the circuit is obtained by taking the high-pass output of the basic second-order cell for $H_1(s)$, i.e. by taking $a = 0$ and $b = 0$. Then we get

$$\omega_0' = \omega_0/\sqrt{(1 - Ac)}; Q' = Q\sqrt{(1 - Ac)} \tag{14.8}$$

In this case, the product $\omega_0' Q'$ remains invariant, i.e. we get higher Q at lower frequencies and vice versa. This is a remarkable feature because achieving high Q at low frequencies is a perennial problem in passive filters because of the large values of inductance and capacitance required and the difficulty of constructing high valued, high Q inductors. This is also a problem in active filters because of the large ratios of resistors and capacitors required for the purpose.

If the feedback signal is taken from the bandpass output itself in Fig. 14.1, then $a = 0$ and $c = 0$ so that from Eq. (14.6),

$$\omega_0' = \omega_0;\ Q' = Q/(1-AbQ/\omega_0) \tag{14.9}$$

In this case, one obtains a variable Q with fixed centre frequency. The corresponding bandwidth is given by

$$\Delta\omega' = \Delta\omega(1-Ab/\omega_0) \tag{14.10}$$

This case is also of practical interest.

14.4 Range of Variation

Obviously, A can be positive as well as negative. The range of positive A (assuming a, b and c to be positive) is determined from the constraint that ω_0', Q' and $\Delta\omega'$ cannot be negative or imaginary. The range of negative A is unlimited in theory and would be limited only from practical considerations. Accordingly, depending upon the particular case under consideration, the parameters ω_0', Q' and $\Delta\omega'$ can be varied from 0 to infinity in theory, and from very low to very high values in practice.

14.5 An Open Problem and a Suggestion

Is it possible to have a constant Q and a variable centre frequency (or bandwidth)? No combination of a, b and c was found to yield this feature. The question is therefore left as an open problem.

The shadow filter concept discussed here can be used, in a variety of ways, to design an experiment for undergraduate students which will illustrate the various basic filters and their tuning/quality improvement by adding feedback through the

amplifier A. For example, the basic second-order cell can be breadboarded with identical resistors and capacitors, as in Fig. 14.2, to give low-pass, high-pass, bandpass and band-stop characteristics. The relevant transfer functions are:

$$\begin{gathered} LP : \omega_0^2/D(s);\ HP : s^2/D(s);\ BP : \omega_0 s/D(s);\ BS :\ (s^2 + 2\omega_0 s + \omega_0^2)/D(s) \\ D(s) = s^2 + (\omega_0/Q)s + \omega_0^2;\ \omega_0 = 1/(CR);\ Q = 1/3 \end{gathered} \tag{14.11}$$

The gain block A can be a single op-amp in inverting or non-inverting configuration; in the latter case, the resistance from the input to the inverting terminal of the op-amp has to be large enough so as not to load the filter to which it is connected. The summer can be realized by another op-amp.

14.6 An Example

In the three special cases considered, $H(s)$ was a bandpass filter. Since the features discussed here apply to the denominator of $H'(s)$ it is clear that the same effects will occur if $H(s)$ is low-pass, high-pass, bandpass or band-stop, thus facilitating adjustment of bandwidth and sharpness of cutoff. As an example, consider Fig. 14.2 and let $H(s)$ be the transfer function of the band-stop filter. The magnitude characteristic of this filter is rather poor because the dc and infinite frequency transmissions are unity but the minimum transmission occurring at $\omega_0 = 1/(CR)$ is only 2/3, while ideally, one would like to have zero transmission at this frequency. To get a deeper minimum, apply feedback from the bandpass output. The minimum value then becomes $2/(3-A)$. For example, with $A = -47$, the minimum value becomes 0.04. Note also that interestingly, when A is positive and lies between 1 and 3, this minimum becomes a maximum. In fact, when A is close to 3, the maximum value can go very high and the filter becomes a bandpass one. It will also be instructive to see what happens when A is gradually increased towards the value 3 and beyond. At some point near 3, the circuit will break into oscillations and produce sinusoidal oscillations at the frequency $\omega_0 = 1/(CR)$.

14.7 Conclusion

The versatility of an interesting active filter, which can be electronically tuned by varying the gain of an amplifier, has been discussed in this chapter. The filter can be quickly breadboarded in the laboratory for experimentation and self-learning by the undergraduate student.

References

1. Y. Lakys, A. Fabre, Shadow filters—a new family of second order filters. Electron. Lett. **46**, 276–277 (2010)
2. S.C. Dutta Roy, The many faces of the single tuned circuit. IETE J. Education. **41**, 101–104 (2000)
3. N.J. Fliege, Multiple biquads. Ch. 13 in *The Circuits and Filters Handbook*, (W.K. Chen (Ed), 2007, CRC Press)

Part IV
Analog Signal Processing: Signal Processing by Discrete Time Networks

In discrete time networks, the signals are still analog, but are in the form of samples. These occur in charge transfer devices (CTD), which are of two types, *viz*. Bucket Brigade Device (BBD) and Charge Coupled Device (CCDs); when the sampled signals are quantized to a finite number of bits, then the signal becomes truly digital. The mathematical description and tools for signal processing are the same for sample data as well as digital signals.

This part contains three chapters, dealing with some aspects of CTD and digital signal processing, (Chap. 15), Signal Processing applications of CTDs (Chap. 16) and Linear phase versus minimum phase CTD transversal filters (Chap. 17).

Chapter 15
Some Aspects of Digital and CTD Signal Processing

Abstract The work done by me and my students in the fields of digital and charge transfer device signal processing is briefly described. Some general comments on the future scope of work in these two related fields are included.

Keywords DSP · CTD signal processing

15.1 Introduction

In this chapter, I shall discuss some aspects of digital and charge transfer device (CTD) signal processing with which I and my students had been involved in the past. In a personal narrative like this, it is neither possible to review the progress and discuss the future promises of the subject, nor shall I attempt to do this. I would therefore mention others' work only in the context of development of our ideas, and, of course, for critical comparison.

15.2 Our Early Work in Digital Signal Processing

Signal processing was one of the latest additions to my research interests at that point of time. Prior to, and also concurrently with signal processing, I and my research group (which, of necessity, has a time-varying composition) had been involved mostly with analysis and synthesis of active and passive circuits, of lumped as well as distributed types, with emphasis on applications in analog integrated circuits (ICs). It was in 1969, shortly after I joined IIT, Delhi that a colleague of mine, S. N. Hazra, expressed an interest in working with me on digital filters. I had, till then, only a casual interest in the subject through the research papers appearing in different professional journals, particularly in *IEEE Transactions on Audio and Electroacoustics*

Source: S.C. Dutta Roy, "Some Aspects of Digital and CTD Signal Processing," *JIETE*, vol 28, pp 252–260, June 1982.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_15

(whose title was later changed to *IEEE Transactions on Acoustics, Speech and Signal Processing*, and later to *IEEE Transactions on Signal Processing*, to reflect the importance of signal processing). I was, however, convinced in my mind that digital signal processing would be a very active and dynamic field of research during the 70s, and it appeared to be an appropriate time to start learning about it. I, therefore, agreed to study, with Hazra, some of the key papers in the field, and there followed long hours of discussions. It was about a year later, after making many mistakes, that we started appreciating the mathematical intricacies and practical importance of the subject, particularly in view of the commercial availability of digital hardware in the form of IC chips. My previous experience in active and passive filters naturally led me to suggest that we concentrate on digital filter design, starting from well-documented analog filters. One method of designing digital filters is to start with the magnitude-squared function of an analog low-pass filter and replace, ω/ω_c, where ω_c is the cutoff frequency, by either $(\tan \omega T/2)/(\tan \omega_c T/2)$ or $(\sin \omega T/2)/(\sin \omega_c T/2)$, T being the sampling period [31, 32]. This results in the so-called digital tan or sine filter. Our first study was concerned with a critical comparison of these two types of filters with the generating analog filter, in respect of performance in the transition as well as stopbands, using Butterworth, Chebyshev and Papoulis optimum filters as specific examples. We found [1] that the tan filter is superior to both analog and sine filters. The cutoff slope and asymptotic loss depend on ω_c and both of them increase in the case of the tan filter, and decrease in the case of the sine filter, as ω_c approaches the Nyquist frequency.

Stimulated by the work of [33], I had worked on sharpening the cutoff of analog filters by introducing multiple $j\omega$—axis zero [34]. We looked into the possibility of extending this technique to digital filters and succeeded. An interesting result emerging from this study is that the Butterworth tan filter is a special case of maximally flat sine filter with coincident transmission zeros at the Nyquist frequency, whose multiplicity equals the order of the filter [2].

15.3 Linear-Phase Recursive Filters

Our next effort in digital filtering was concerned with achieving linear phase in recursive digital filters. Here, the approach of transformation of analog filter does not serve the purpose and one has to resort to a direct approximation. We started from a well-established result in analog filters viz. that an infinite array of poles at equal intervals on a line parallel to the $j\omega$-axis in the left half of the s-plane leads to an equal ripple delay characteristic in the entire frequency range [35, 36]. We, therefore, argued that if all the poles of a digital filter are distributed uniformly on a circle of radius $r < 1$ in the z-plane, there should result an approximation to linear phase. Indeed, the transfer function having the above distribution of poles is

$$H_o(z) = 1/(r^m + z^m) \tag{15.1}$$

and its phase is given by

$$\phi_o(w) \cong -m\omega T + r^m \sin(m\omega T) \tag{15.2}$$

where the approximation is valid for large m. The error in linear phase is seen to be periodic with a maximum value of r^m, which can be made arbitrarily small by choosing r and m suitably. The magnitude response of Eq. (15.1) is

$$M_o(\omega) \cong 1 - r^m \cos(m\omega T) \tag{15.3}$$

and has an equal ripple all-pass characteristic. The filter can be made low-pass by cancelling the stopband poles by introducing zeros at these locations. Retaining linear phase requires their mirror image zeros also to be introduced. Using this simple idea, we worked out an analytical design procedure [3, 4] for a low-pass filter: satisfying prescribed specifications on delay, maximum delay error and asymptotic slope. In contrast to this, other known design procedures [37, 38] for similar specifications require numerical techniques for solving a set of simultaneous nonlinear equations.

15.4 Discrete Hilbert Transform and Its Applications in Digital Filtering

I went to England in 1973 as a Visiting Professor for a year, to work at the University of Leeds, where I interacted with two very renowned circuits specialists, J. O. Scanlan and J. D. Rhodes. I also delivered invited seminars at a number of UK Universities, a Canadian University, and at the IEE Headquarters, London, which gave me an opportunity of holding personal discussions with a number of experts in digital signal processing. This experience at UK inspired me to devote a major part of my research efforts to this field after I returned to IIT in 1974. One problem I worked on concerned a matrix formulation of the discrete Hilbert transform (DHT), which plays an important role in signal processing and communication systems. A matrix formulation of the DHT was given independently by Sabri and Steenaart [39] and Burris [40], in which the vector $[\phi] = [\phi(0)\phi(1)\ldots\phi(N-1)]^T$, representing the DHT of the periodic sequence $a(n)$, $n = 0$ to $N-1$, was expressed as the product of an $N \times N$ coefficient matrix $[c_{n,v}]$ and the vector $[A] = [A(0)\ A(1)\ldots\ A(N-1)]^T$, where $A(k)$ is the discrete Fourier transform (DFT) of $a(n)$. By exploiting certain properties of the coefficient matrix, I succeeded in showing that [5] an alternative formulation is possible in which $[\phi]$ is expressed as $[\hat{A}][\hat{C}]$ where $[\hat{C}]$ is a column vector and $[\hat{A}]$ has elements of the form $A_i - A_j$. The advantage of this formulation is that the number of multiplications, which accounts for most of the processing time, is reduced by a factor of two. Assisted by an M. Tech. student, A. Agarwal, I applied this formulation to a digital low-pass filter, and showed that this leads to less storage requirement and

improved noise performance, in addition to reduced computations [9]. This approach was later extended to two-dimensional filtering by Bose and Prabhu [41].

15.5 Roundoff Noise in Digital Filters

Digital signal processing algorithms such as linear filtering are realized either as programmes in a general purpose computer or by a special-purpose hardware. In either case, the signal and the coefficients are stored as binary numbers in finite length registers. This finite wordlength constraint gives rise to two main sources of error, viz. coefficient quantization error and multiplication roundoff error. Unless carefully analysed and accounted for, these errors can lead to very undesirable effects like limit cycles and overflow oscillations. These errors are intimately connected with the structure or the architecture of the processor. For a specific case, viz. a second-order digital notch filter, R. K. Patney, a doctoral student and I studied the roundoff noise properties of various canonical structures and established that depending upon the zeros of the noise transfer functions (NTFs), such structures fall in three distinct groups. The group showing low noise magnitudes has NTF zeros which are simply related to the coefficients of the transfer function [6].

For the general digital filter, the variance of the multiplication roundoff error at the output due to a noise source is given by

$$\sigma_o^2 = (q^2/12)[1/(2\pi j)] \oint_\Gamma H(z)H(z^{-1})z^{-1}dz \tag{15.4}$$

where q is the quantization step size, Γ is the unit circle and $H(z)$ is the NTF. Various methods are available for the evaluation of Eq. (15.4) (see, e.g. Mitra et al. [42] and the references cited there). We obtained two alternative methods [17] for this purpose, both of which are based on the decomposition of the given NTF in terms of digital all-pass transfer functions. In the first method, the NTF is expressed as the sum of first-order all-pass functions, the motivation being that for each component, the integral in Eq. (15.4) simply reduces to $q^2/12$. The method requires somewhat complex arithmetic calculations, but takes less computation time than that of [42]. The second method utilizes the concept of polyphase digital filtering [43] to express the NTF as a sum of all-pass functions, of first- and second orders. The accuracy obtained in this method is not as good as in the first method, but it has the advantage of a drastic reduction in computation time, specially for higher order NTF's. Both of these methods, we believe, offered new insights into the roundoff noise problem in digital filters.

15.6 Quantized Coefficent Design of Digital Filters

As to the coefficient sensitivity problem in digital filters, we investigated the possibility of reducing the deterioration in filter characteristics due to coefficient quantization by an optimal choice of the coefficients from a discrete parameter space, once the worldlength is decided by considerations of cost. Let $F(c_i', \omega)$ be a function describing some characteristic of a digital filter, where c_i', $i = 0, 1, \ldots n$ are the coefficients determined to infinite accuracy by some design method. Let c_i be the coefficients obtained after quantization of c_i' to a desired accuracy of $b + 1$ bits, i.e.

$$c_i = c_{io} \sum_{j=1}^{b} c_{ij} 2^{-j}, \quad c_{ij} \in \{0, 1\} \tag{15.5}$$

where c_{io} denotes the sign of c_i. The resulting deteriorated function is $F(c_i, \omega)$. The following question now arises: Is it possible to modify the last bit of c_i such that the maximum of the error $|F(c_i, \omega) - D(\omega)|$, where $D(\omega)$ is the desired function, is minimized? The answer turns out to be in the affirmative; however, the existing techniques for this purpose use nonlinear optimization methods, in which a global optimum is not guaranteed. Further, most of these methods are not computationally attractive. We, therefore, looked for an entirely different approach, and an exhaustive search finally led us to conclude that pseudo-Boolean (PB) techniques [44], based on a combination of dynamic programming and Boolean techniques, should be eminently suitable for the problem at hand. We studied this technique in-depth, including some of its highly successful applications in transportation and network flow problems. We found that the problem of quantized coefficient design of digital filters essentially reduces to that of modifying c_i to $c_i + {}_{\text{a}}c_i$, where ${}_{\text{a}}c_i = c_{io} S_{ci} q y_{ci}$; $q = 2^{-b}$ (quantization step); $y_{ci} \in \{0, 1\}$ are bivalent variables whose values, to be determined; and $S_{ci} \in \{-1, 0, 1\}$ define the direction of modification of c_i. To start with, we set $S_{ci} =$ sign of $(c_i' - c_i) - c_i)$. If ω_r is the frequency at which the extrema of the error $|F(c_i, \omega) - D(\omega)| = \delta_r$ is maximum, we can write

$$|F(c_i + a_{ci}, w_r) - D(w_r)| \leq \delta_r. \tag{15.6}$$

This is a PB inequality in the variables y_{ci} and has to be solved. From the possible solutions, we select the one giving the largest improvement in the error function, by evaluating $F(c_i + {}_{\text{a}}c_i, \omega)$ at a grid of frequency points. Corresponding to the best solution, we prepare for the next iteration by modifying c_i again as in the previous step, i.e. S_{ci} is replaced by $S_{ci}(2y_{ci} - 1)$, while δ_r is replaced by its new lower value, and inequality in Eq. (15.6) is reformulated. The procedure is repeated till no further improvement in δ_r results.

The nature of PB inequalities in Eq. (15.6) is different for finite-impulse response (FIR) and infinite-impulse-response (IIR) digital filters. For the former case, Eq. (15.6) becomes a linear PB inequality, the solution of which, we showed,

can be made non-iterative [13, 16]. The method is simple to use, takes much less computer time as compared to the existing techniques, and yet has much better chances of converging to a global optimum. We applied the method to a number of sample designs, including that of a DHT, and obtained excellent results.

For the IIR case, $F(c_i, \omega)$ is a rational function of the form $P(a_i, \omega)/Q(b_i, \omega)$, where P and Q are nonlinear functions of a_i and b_i, the maximum order of nonlinearity being equal to the order of the filter. In this case, Eq. (15.6) becomes

$$P(a_i + \alpha_{ai}, \omega_r) - \delta Q(b_i + \alpha_{bi}, \omega_r) \leq 0 \tag{15.7}$$

where $\delta_r = D(\omega_r) \pm \delta_r$. This set of PB inequalities is nonlinear and the usual method of solution by linearizing each nonlinear variable is computationally expensive. To overcome this difficulty, we formulated an alternative method for solving nonlinear PB inequalities [7, 16] and applied it successfully to a number of useful IIR designs.

15.7 Variable IIR Digital Filters

In many applications, such as speech processing and spectrum analysis, it is often required to change the cutoff frequency of a digital low-pass filter or the bandwidth or centre frequency of a digital bandpass filter, by controlling, preferably, a single parameter in the filter structure. In view of the importance of such variable digital filters, and in view of the fact that the available techniques are either not wholly satisfactory or not adequate for handling a number of important situations, including bandpass filters, I suggested to S. S. Ahuja, a Ph.D. student, to concentrate on this topic.

The usual approach for achieving variable digital filtering is to use a frequency transformation. For example, if each delay element in the prototype filter is replaced by an all-pass network, a variable filter results, the control being affected by the all-pass parameter(s) [45]. If the prototype filter is IIR, then this approach leads to delay-free loops in the structure, which cannot be implemented digitally. Schussler, therefore, suggested the use of an FIR prototype; later, Johnson [46] indicated a general approach for removal of delay-free loops by computation of modified coefficients through a filtering process. We first established the desirability of IIR prototypes by showing that variable digital filters with FIR prototypes are highly inefficient with regard to the number of multiplications and have very high roundoff noise variance at the output. We next obtained a very simple method for eliminating delay-free loops. We argued that since the usual realization of IIR will be either cascade or parallel combinations of first- and second-order sections, all we have to do is to rewrite the transfer function of each such section, after transformation, in such a way that it can be realized without delay-free loops [18]. As compared to Johnson's method, our procedure offers, besides simplicity, a reduction in computations by a factor of 3 for first order and 2.5 for second order; this can be of great advantage in reducing the settling time in adaptive IIR filters.

15.8 Variable FIR Digital Filters

For variable cutoff linear-phase FIR filters, Oppenheim et al. [47] suggested using the transformation

$$\cos\omega = \sum_{k=0}^{P} A_k (\cos\Omega)^k \tag{15.8}$$

where ω and Ω refer, respectively, to the prototype and transformed filters. This transformation preserves the frequency response but distorts the frequency axis. The first order (1°) form of Eq. (15.8), where $P = 1$ and single parameter control is achieved by constraining the DC gain to unity, has been found particularly useful, one reason being that it results in a canonic realization of the transformed filter. We investigated this approach to variable digital filters in-depth, and obtained the following new and useful results:

(i) In the 1° case, we showed that the variable parameter A_o ($A_1 = 1 - A_o$, for invariant DC gain) should lie in the range 0 to $\cos^2(\omega_c/2)$, where ω_c is the cutoff frequency of the prototype filter (corresponding to $A_o = 0$). We also showed that in the transformed filter, the cutoff frequency Ω_c increases, while the cutoff slope decreases, monotonically, with increasing A_o. To overcome these two drawbacks, we formulated the following optimum 2° transformation:

$$\cos\omega = A_o + \cos\Omega - A_o\cos^2\Omega \tag{15.9}$$

This retains single parameter control and achieves a two-way variation of Ω_c as well as cutoff slope. The most important feature of Eq. (15.9) is that the computational load of the transformed filter is approximately half of that of a double-length prototype with 1° transformation, the performance in the two cases being comparable [12].

(ii) We extended the 1° transformation to design a bandpass filter with fixed centre frequency but variable bandwidth, again with single parameter control. The transformation cannot, however, be used to obtain constant bandwidth and variable centre frequency. We showed, however, that if the transformed bandwidth $\Delta\Omega$ is larger than the prototype bandwidth $\Delta\omega$, then it is possible to maintain $\Delta\Omega$ fixed and vary Ω_o (transformed centre frequency) above or below ω_o (prototype centre frequency). One, therefore, has to choose $\Delta\omega$ appropriately [10]; this increases the filter order.

(iii) To overcome the difficulty of increased bandwidth of the 1° transformed variable bandwidth, constant centre frequency digital bandpass filter, we formulated a new 3° transformation, given by

$$\cos\omega = A_o + A_1\cos\Omega - A_3\cos^3\Omega \tag{15.10}$$

where A_1 and A_0 are related to A_3 through $\omega_{1,2}$ and $\Omega_{1,2}$ (the subscripts 1 and 2 refer to the lower and upper cutoff frequencies, of the prototype and transformed filters). We showed that by appropriately constraining the parameters and prescribing A_3 as 2^{-n}, n being an integer (so that multiplication by A_3 amounts to a simple shifting operation only), it is possible to achieve constant $\Delta\Omega = \Delta\omega$, and single element controlled variable centre frequency without increasing the computational load, as compared to the 1° transformation. The number of delay elements has, of course, to be increased three times [14].

(iv) We extended the 1° transformation to two-dimensional variable cutoff digital low-pass filters. Also, the drawbacks of the 1° transformation, viz. higher cutoff frequencies and wider transition bands, have been shown to be overcome by an appropriately formulated 2° transformation. As in the one-dimensional case, the order of the filter is doubled thereby, but the computational load remains unchanged [8].

15.9 Nested Structures for FIR Filters

It has already been mentioned that the finite word length effects are intimately related to the structure of the digital filter. While a large number of such structures have been reported for the IIR filters, much less is known about the FIR case. I, therefore, suggested to a Ph.D. student, A. Mahanta, to study the various existing structures and to search for new ones which have the potential of reducing the finite wordlength effects. We first investigated a special FIR filter viz. the discrete Hilbert transformer (DHT). For a symmetric DHT of length $N = 2M + 1$ (M odd), the transfer function is

$$H(z) = z^{-M}[h_1(z - z^{-1}) + h_3(z^3 - z^{-3}) + \ldots + h_M(z^M - z^{-M})] \tag{15.11}$$

For the optimal DHT (ODHT) [48], we observed that the impulse response coefficients $\{h_n\}$ fall more or less inversely with n. This encouraged us to rewrite Eq. (15.11) as

$$H(z) = z^{-M} g_1(z - z^{-1} + g_3(z^3 - z^{-3} + \ldots . + g_M(z^M - z^{-M})) \tag{15.12}$$

where $g_k = h_k/h_{k-2}$, k odd, and $g_1 = h_1$. In general, g_k's are all distinct, but beyond a certain k, all g_k's are approximately equal. Thus if we let $g_k = , k = K$ to M, then

$$\begin{aligned} H(z). = z^{-M} g_1(z - z^{-1} + g_3(z^3 - z^{-3} + \ldots + \alpha(z^K - z^{-K} \\ + \alpha(z^{K+2} - z^{-(K+2)} + \ldots + \alpha(z^M - z^{-M}) \ldots) \end{aligned} \tag{15.13}$$

When Eq. (15.13) is implemented in the direct form, it still requires, the same processing time, but the coefficient storage requirement is now reduced to $(K + 1)/2$,

instead of $(N + 1)/4$. We investigated this modified implementation (NDHT) for various lengths and found that its performance is only slightly inferior to that of the ODHT in terms of deviations, and that in terms of coefficient quantization error, the NDHT is as good as, if not better, than the ODHT, especially for small wordlengths. Also, the multiplication roundoff noise of the NDHT is much less than that in the direct form implementation [11].

Encouraged by these results, we next investigated the general linear-phase FIR filter, having the transfer function (length $N = 2M + 1$):

$$H(z) = z^{-M}[a_o + a_1(z + z^{-1}) + \ldots + a_M(z^M + z^{-M})] \tag{15.14}$$

We observed that, in most cases of filter design, a_n decreases (in magnitude) with n, as in DHT, or oscillates with decreasing amplitude, as in bandpass filters, and that there are usually several orders of magnitude difference between the largest and the smallest a_n. Consequently, when the coefficients are quantized to fixed number of bits, all a_n's cannot be represented with the same degree of relative accuracy, and this results in error in the filter response. As in the DHT case, we, therefore, rewrite Eq. (15.14) as

$$H(z) = z^{-M}b_o[1 + b_1(z + z^{-1}) + \ldots + b_M(z^M + z^{-M})\ldots] \tag{15.15}$$

where $b_n = a_n/a_{n-1}$, $n = 1$ to M and $b_o = a_o$, and argued that b_n's should all be comparable in magnitude. As a result, effects of quantization should be less pronounced than in the case in which Eq. (15.14) is directly implemented. We carried out a deterministic as well as a statistical analysis of the filter response error for the direct and the alternate forms, and found that in majority of cases of practical importance, the new structure, which we called the 'nested structure', gives rise to less error; further, we found that multiplication round off error in the nested form is much lower than that in the direct form [15].

Later, we made a comprehensive investigation of the generalized nested structures, in which the coefficients to be implemented had been made to lie in the range $\frac{1}{2}$ to 1, by suitable decomposition and scaling. We have shown that using fixed-point arithmetic, these structures achieve the same accuracy and about the same roundoff noise as those obtained in the floating-point implementation. The optimum structure to achieve the minimum roundoff noise can be found; however, the search in that direction was not as involved as in the cascade form, and in most cases, near-optimum results are easily obtained by mere inspection of the impulse response coefficients [19].

15.10 Fourier Transform at Arbitrary Frequencies

A major development in digital signal processing is the discovery of the fast Fourier transform (FFT), which now plays a central role in many signal processing applications, particularly in spectral analysis. As is well known, the FFT achieves its efficiency by exploiting the uniformity of its samples in the time domain; however, it finds the discrete FT (DFT) at uniformly spaced frequencies only. While this is indeed the requirement in a variety of applications, there do exist a few important spectral analysis situations where one has to estimate the spectrum at nonuniform frequencies. Typical examples are: constant-Q spectral analysis [48], speech processing for enhancement of formant resonances; correction of nonlinear distortion in frequency [50]; and detection and estimation of a narrow band spectral peak [49]. Available techniques for this purpose are either computationally inefficient, or applicable to some restricted situations only, and it appeared worthwhile to find alternative efficient techniques, which are of general applicability. This is what R. C. Agarwal, a colleague, and I suggested to a Ph.D. student, R. Sudhakar, to persue. We indeed came up with a new general method which retains the order of computation at FFT levels, but allows the flexibility of choosing arbitrary frequencies for a uniformly sampled signal. The method is based on the fact that the FT of a finite length sequence at any arbitrary frequency can be expressed as a weighted sum of its DFT coefficients. These weights are suitably approximated such that the FT is very nearly equal to the sum total of a few dominant terms of the weighted sum, which are computed directly, and the DFT of a derived sequence formed by multiplying the original sequence with a sawtooth function. The number of directly computed terms is so chosen that the error does not exceed the desired limits [20].

15.11 Inefficiency Compensation in CTD Delays and Signal Processors

The charge transfer device (CTD) is a monolithic IC analog sampled data delay line, first disclosed in 1970. CTDs are of two types, viz. bucket brigade device (BBD) and charge coupled device (CCD), and though only a decade old at that time, they proved to be uniquely suitable for many analog signal processing situations. They were successfully applied in MTI radar, TV ghost suppression, spectrum analysis, matched filtering ,etc. Being analog, CTDs can, at least theoretically, achieve all the advantages of digital signal processors, without the need for A/D conversion and without suffering from quantization errors [25].

CTDs, however, have their own limitations, the most important being the charge transfer inefficiency (CTI), which causes magnitude as well as delay distortions. CTDs operate by storing charge, proportional to the sampled value of a signal, in MOS capacitors, and transferring it from one stage to the next under the influence of external clock signals. The charge transfer, unfortunately, is not perfect; the fraction

of the charge which fails to get transferred is a measure of the CTI. If all the stages of the CTD are identical, then each delay stage can be modelled by the z-transfer function

$$z_{\varepsilon}^{-1} = (1 - \varepsilon)z^{-1}/(1 - \varepsilon z^{-1}) \tag{15.16}$$

instead of z^{-1} to which it reduces when $= 0$.

I was attracted to CTD's while in England, mainly because of the similarity between digital signal processing and CTD signal processing. After returning to IIT in 1974, I started thinking about the possibility of compensating for CTI by using a technique similar to predistortion in analog filter design. Along with a graduate student, R. K. Arora, I evolved the following simple procedure. Given a transfer function $H(z)$ to be realized, replace z by its equivalent expression in terms of z, viz.

$$z = (1-\varepsilon)z_e + \varepsilon \tag{15.17}$$

to yield a transfer function $G(z)$, which is, of course, identical to $H(z)$. $G(z)$ can now be realized by the nonideal CTD; the resulting structure, which is necessarily recursive, is an exact realization of $H(z)$. We applied this technique for realizing an ideal delay and a second-order resonator [21].

The preceding method was based on the assumption that all CTD stages had the same CTI and that there were no transfer losses. Later, V. G. Das, a Ph.D. student and I worked on the more general case where these assumptions are not valid, and obtained expressions for the feedback coefficients for exact compensation of a CTD delay. We also developed a tuning algorithm for the critical adjustment of all the feedback coefficients. One of the criticisms of this compensation technique was that a large number of feedbacks necessary for compensation would be difficult to fabricate on the chip. We, therefore, looked into the extent of compensation obtained by using a limited number of feedbacks, theoretically [28] as well as experimentally [23]. For the latter, we used a simple experimental set up with a highly inefficient CTD (obtained as a gift from Dr. H. Wallinga of Twente University, Netherlands) delay comprising of dissimilar stages, and to our pleasant surprise, we found that good compensation could be achieved using only the first two feedbacks.

Towards the end of 1977, Professor A. B. Bhattacharyya of IIT Delhi, and his group fabricated the first CCD in India, and by 1978, we got a sizeable number of CCDs to play with. The initial successes obtained in compensation by using a small number of feedbacks encouraged us to carry out more elaborate investigations. We argued that the whole compensating structure, consisting of one or two feedbacks, could be fabricated as a separate block and used as a pre-or post-filter. As a pre-filter, the compensation block simply serves to pre-distort the input signal in a manner, which is opposite to that introduced by the CTD. By extension of this concept, we evolved a simple, but accurate method for measurement of CTI. We fed a signal $K_1u(n) - K_2u(n - 1)$, where $u(n)$ stands for the unit discrete step, and found, from theoretical analysis, that the first two consecutive output samples will be equal when $K_2/K_1 = \sum\varepsilon_i$. In a simple experimental set up, we obtained a low-frequency square wave signal by

frequency division of the clock signal, and used a shift register and a summer to obtain the required signal, K_1 and K_2 being implemented by potentiometers [24]. As compared to other methods, which involve measurement of the difference between two pulse heights, our method has all the advantages of a null measurement.

A method of compensating for CTI in a CTD delay was proposed by Caywood and Buss [51] in which q CTD's each having M/q stages, are operated in parallel through an input steering circuit and an output multiplex circuit. Such a multiplexed CTD (MCTD) achieves a delay of M stages, but with reduced CTI, as shown by an approximate analysis by Caywood and Buss. We carried out a vigorous analysis of the MCTD and showed that the magnitude and delay responses are an equal ripple. Further, the magnitude ripple decreases with increasing q, but the delay ripple is independent of q and is determined by only the CTI [22].

I went to the United States in 1978, to work for a year at the Iowa State University, as a Visiting Fellow. At Iowa, I interacted with an outstanding researcher, Professor A. N. Michel. I visited several US Universities and one Canadian University for invited seminar lectures, and also gave an invited talk at a Workshop on Algebraic Complexity of Computations. I also interacted with an industry, Rockwell International, through a graduate course offered on TV, with frequent personal visits. All these gave me excellent opportunities for an in-depth exploration of state of the art in research and development in the fields of digital and CTD signal processing. In particular, I found that application of CTD in the industry mostly uses transversal structures, and that major efforts are being directed to the use of this structure in spectrum analysis. On my return to India, in 1979, the first problem I tackled, concerned the compensation of CTI in such structures. One of the very early suggestions for this purpose consists of predistorting the tap weights [52]. If the desired transfer function is

$$H_i(z) = \sum_{k=0}^{N} h_k z^{-k}, \tag{15.18}$$

and $h_k's$ are predistorted to h_k', then the realized transfer function will be

$$\begin{aligned} H_a(z) &= \sum_{k=0}^{N} h_k'(1-\varepsilon)^k z^{-k}/(1-\varepsilon z^{-1})^k \\ &\triangleq \sum_{k=0}^{\infty} h_k'' z^{-k} \end{aligned} \tag{15.19}$$

One way of choosing h_k' is to set $h_k'' = h_k$, $k = 0$ to N; following this, Buss et al. [52] established an iterative formula for h_k', which has been widely cited in the literature and used in CTD signal processor design. We found that h_k' can be expressed explicitly in terms of h_k and that this formula is rather simple. We also derived an explicit formula for h_k'', $k > N$, and used it for estimating the error in this method of compensation [26].

15.12 CTD Linear Phase Versus Minimum Phase Design

CTD transversal filters are usually designed for linear-phase characteristic; the resulting time delays are acceptable in most situations but not so in some telecommunication applications, e.g. in voice lines, due to the problems of echo and singing. To avoid this, Foxall [53] investigated an example of minimum phase design, and observed that besides reducing group delay, the minimum phase design is less sensitive to CTI and tap weight error. As a starting problem, I suggested to a Ph.D. student, Navin Kapur, to make a detailed analysis of the sensitivity of general linear and minimum phase designs to CTI in order to find out the validity or otherwise of Foxall et al.'*s* conclusions in general. We found that the answer is in the negative and that the sensitivity to CTI is a sensitive function of the zero locations. We also derived conditions under which the minimum phase design would be superior to linear-phase design [30].

15.13 Coefficient Sensitivity in CTD Filters

CCD transversal filters are most commonly implemented, using the split-electrode technique and the resulting tap weight errors constitute a major factor limiting the stopband performance. This error increases with decreasing tap weight, and unfortunately, in typical frequency-selective filters, there is a large number of small tap weights, the ratio of the smallest to the largest coefficient being of the order of a few thousands! For low-pass filters, however, we observed that the difference between the adjacent coefficients has a much smaller spread than the coefficients themselves. We exploited this fact to formulate a new implementation technique for reducing coefficient sensitivity. Instead of realizing a given

$$H(z) = \sum_{n=0}^{N-1} h(n) z^{-n} \tag{15.20}$$

directly, we formed a new sequence

$$h_1(n) = h(n) - h(n-1) \tag{15.21}$$

with $h_1(0) = h(0)$ and $h_1(N) = -h(N-1)$. Implementing a filter with the weights $h_1(n)$ gives a transfer function $H_1(z)$, which, upon multiplication by $H_2(z) = 1/(1 - z^{-1})$, gives $H(z)$. $H_2(z)$, we argued, can be realized by a switched capacitor differential integrator, which can be fabricated along with the CCD, by the same technology. The resulting structure does not use any extra hardware as compared to the usual realization of $H(z)$ because the differential integrator is a replacement of the differential amplifier, required in the latter. We also showed that the following additional advantages are obtained in the modified implementation: (i) Since $\sum h_1(n) = 0$,

the parallel channel present in earlier designs to cancel the output DC offset due to fat zero and clock transients, is not required. This implies a saving of chip area and simplified design layout and (ii) Passband gain G is higher, implying a smaller insertion loss and hence, a larger dynamic range [29].

15.14 Spectrum Analysis by CTD

A large part of Navin Kapur's Ph.D. thesis work related to the use of CCD for spectrum analysis. One published work relates to the implementation of prime transform algorithm (PTA) using CCD programmable transversal filters (PTF) [26]. The PTA converts the problem of computation of DFT into a circular convolution, and when the number of points N is large, the chip area increases considerably and the CTI has a significant effect on the performance. Previously reported systems employing a PTA used a $(2N - 3)$ stage fixed tap weight CCD filter for each convolution. Using a CCD programmable transversal filter based on a new concept viz. storing the signal to be analysed, and passing the basis vector sequence through the CCD, we showed that it is possible to implement an N-point DFT using the PTA with an $(N - 1)/2$ stage PTF, for each convolution. This represents a factor of four reductions in CCD stages over previously reported fixed tap weight systems.

15.15 Looking Ahead

In this chapter, I have briefly described our efforts and some new results arising from them in the areas of digital and CTD signal processing. After about a decade of interaction with these areas, we felt that we had solved only a few problems of contemporary interest at that time and that much remained to be done. One major problem which should warrant the attention of serious workers in digital signal processing concerns the heavy (and sometimes undue) reliance on analog filter techniques and computer-aided methods. A direct approach in the z-plane, independently of the s-plane, leading to analytical, rather than computer-aided methods, was yet to be developed for this young field to mature. This, I felt, required dedicated work by individuals who would accept potential failures as the key to future success, and to whom a negative result means one step forward to the ultimate goal.

Since the analysis and design of CTD and digital signal processors were very much alike, any development in the latter field would have definitely benefited the former. In addition, much efforts were needed to overcome the limitations of the device, imposed by fabrication technology and the physics of operation, while, concurrently, new applications of the device and improvement of performance in known applications by new design methods will continue. Also, a realistic assessment of what is better in a particular situation, digital or CTD signal processing, is long overdue, and deserves serious consideration, ignoring the claims of overenthusiastic workers in either field.

Digital and analog signal. processing techniques, as I commented in an earlier paper [53] are 'two arrows in the quiver', which are always 'better than one', A judicious combination of the two will, perhaps, lead the field of signal processing beyond presently known boundaries, particularly because of the capability of the CTD to perform analog as well as digital signal processing.

Acknowledgements I would like to take this opportunity to express my grateful appreciation to the IETE for bestowing a great honour upon me through the award of the Fifth Ram Lal Wadhwa Gold Medal, and for inviting me to deliver a memorial lecture on which the chapter is based. I have been very privileged to have studied under great teachers like (Late) Professor H. P. De, (Late) Professor J. N. Bhar, Professor A. K. Choudhury, Professor S. Deb, Professor M. K. Dasgupta, Professor A N. Daw and Professor B. R. Nag, to all of whom I express my gratitude. I have also been lucky to have worked in wonderful academic environments provided by the IIT Delhi, University of Minnesota, USA, University of Leeds, England, and the Iowa State University, USA; to all those responsible for creating and maintaining this environment, I owe deep indebtedness. Finally, I have been exceptionally lucky in working with a band of very brilliant and dedicated students, some of whom I have specifically named during the course of this lecture, and a group of outstanding colleagues; to all of them, I owe a special word of appreciation and thanks.

References[1]

A.1. On Digital Signal Processing

1. S.N. Hazra, S.C. Dutta Roy, A comparison of digital tan and sine filters with the generating analog filter. J. Franklin Inst. **292**, 225–230 (1971)
2. S.N. Hazra, S.C. Dutta Roy, Maximally flat low-pass digital filters with sharp cutoff characteristics. Proc. IEE. **119**, 1575–1576 (1972)
3. S.N. Hazra, S.C. Dutta Roy, A simple design of linear-phase, low-pass recursive digital filters. Int. J. Circuit Theor. Appl. **1**, 11–16 (1973)
4. S.C. Dutta Roy, A state of the art survey on linear phase digital filter design. J. IETE. **22**, 1–4 (1976)
5. S.C. Dutta Roy, Alternative matrix formulation of the discrete Hilbert transform. Proc. IEEE. **64**, 1435 (1976)
6. S.C. Dutta Roy, R.K. Patney, Roundoff noise properties of second order digital notch filters. Int. J. Circuit Theor. Appl. **6**, 183–202 (1978)
7. R.K. Patney, S.C. Dutta Roy, Design of IIR filters using pseudo-boolean methods. Proc. ICASSP. (1978)
8. S.S. Ahuja, S.C. Dutta Roy, Variable cutoff two dimensional low-Pass FIR digital filters. Electron. Lett. **14**, 422–423 (1978)
9. S.C. Dutta Roy, A. Agrawal, Digital low-pass filtering using the discrete Hilbert transform. IEEE Trans ASSP- **26**, 465–467 (1978)
10. S.S. Ahuja, S.C. Dutta Roy, Linear phase variable digital bandpass filters. Proc. IEEE. **67**, 173–174 (1979)
11. A. Mahanta, S.C. Dutta Roy, A new approximation of the discrete Hilbert transformer. Proc. IEEE. **67**, 174–175 (1979)

[1]Publications of the author and his co-workers.

12. S.C. Dutta Roy, S.S. Ahuja, Frequency transformation for linear phase variable cutoff digital filters. IEEE Trans. CAS **26**, 73–75 (1979)
13. R.K. Patney, S.C. Dutta Roy, Design of linear phase FIR filters using the pseudo-boolean methods. IEEE Trans. CAS **26**, 255–260 (1979)
14. S.S. Ahuja, S.C. Dutta Roy, Variable centre-frequency constant-bandwidth linear-phase digital bandpass filters. Electron. Lett. **15**, 245–246 (1979)
15. A. Mahanta, S.C. Dutta Roy, An alternate direct form realization of FIR filters and its noise analysis. Proc. IEEE **67**, 1077–1078 (1979)
16. R.K. Patney, S.C. Dutta Roy, A review of pseudo-boolean methods with applications to digital filter design. J. Inst. Electron. Telecommun. Eng. **26**, 1–8 (Invited Paper) (1980)
17. R.K. Patney, S.C. Dutta Roy, A different look at the roundoff noise in digital filters. IEEE Trans. CAS **27**, 59–62 (1980)
18. S.S. Ahuja, S.C. Dutta Roy, Variable digital filters. IEEE Trans. CAS-**27**, 9, 836–838 (1980), Corrections in CAS-**28**, 855 (1981)
19. A. Mahanta, S.C. Dutta Roy, R.C. Agarwal, FIR filter structures having low sensitivity and roundoff noise. Proc. ICASSP. 1205–1208 (1981)
20. R. Sudhakar, R.C. Aganral, S.C. Dutta Roy, Fast computation of fourier transform at arbitrary frequencies. IEEE Trans., CAS-**28**, 972–980 (1981)

A.2. On CTD Signal Processing

21. S.C. Dutta Roy, R.K. Arora, Exact circuit design using inefficient CTD's. Electron. Lett. **12**, 18–19 (1976)
22. S.C. Dutta Roy, On the frequency response of a multiplexed charge transfer delay line. Electron. Lett. **12**, 280–281. Corrections in **12**, 976, 391
23. S.C. Dutta Roy, V.G. Das, On exact compensation of transfer inefficiency in a charge transfer delay line. Electron. Lett. **14**, 115–116 (1978)
24. S.C. Dutta Roy, V.G. Das, Predistortion technique for the measurement of transfer inefficiency of charge transfer devices. Proc. IEEE **66**, 601–602 (1978)
25. S.C. Dutta Roy, A.B. Bhattacharyya, J.M. Vasi, V.G. Das, L. Shankar, N. Kapur, Signal processing applications of charge transfer devices. J. Inst. Electron. Telecommun. Eng. **24**, 400–418 (1976)
26. N. Kapur, S.C. Dutta Roy, Implementation of prime transform algorithm using CCD programmable transversal filter. Electron. Lett. **16**, 809–811 (1980)
27. S.C. Dutta Roy, Compensation for inefficiency in charge transfer device transversal filter. Microelectron. J. **12**, 5–7 (1981)
28. S.C. Dutta Roy, V.G. Das, Compensation of inefficiency in charge transfer delays and signal processors. Proc. ISCAS, 54–57 (1981)
29. N. Kapur, S.C. Dutta Roy, Methods for reducing coefficient sensitivity in lowpass CCD transversal filters. Electron. Lett. **17**, 435–437 (1981)
30. S.C. Dutta Roy, N. Kapur, Linear-phase versus Minimum-phase CCD transversal filters. J. Inst. Electron. Telecommun. Eng. **27**, 622–624 (1981)

B. Other References

31. C.M. Rader, B. Gold, Digital filter design techniques in the frequency domain. Proc. IEEE **55**, 149–471 (1967)

32. R. K. Otnes, An elementary design procedure for digital filters. IEEE Trans. AU-**16**, 330–335 (1968)
33. A. Budak, P. Aronhime, Maximally flat low-pass filter with steeper slope at cutoff. IEEE Trans. AU-**18**, 63–66 (1970)
34. S. C. Dutta Roy, On maximally flat sharp cutoff low-pass filters. IEEE Trans. AU-**19**, 58–63 (1971)
35. R.M. Lerner, Bandpass filters with linear phase. Proc. IEEE **52**, 249–268 (1964)
36. D.S. Humpherys, *The Analysis, Design and Synthesis of Electric Fibers* (Prentice-Hall, New Jersey, 1970), pp. 422–427
37. J.P. Thiran, Equal ripple delay recursive digital filters. IEEE Trans. CT-**18**, 665–669 (1971)
38. A.G. Deczky, Recursive digital filters having equiripple group delay. IEEE Trans. Circuits Syst. CAS-**21**, 131–134 (1974)
39. M.S. Sabri, W. Steenaart, A new approach to digital filter realization using the discrete Hilbert transform. Proc. IEEE Canadian Commun. Power Conf. 89–90 (1974)
40. F.E. Burris, Matrix formulation of the discrete Hilbert transform. IEEE Trans. CAS-**22**, 836–838 (1975)
41. N.K. Bose, K. A. Prabhu, Two dimensional discrete Hilbert Transform and computational complexity aspects in its implementaiion. IEEE Trans. ASSP-**27**, 356–360 (1979)
42. S.K. Mitra, K. Hirano, H. Sakguchi, A simple method of computing the input quantization and multiplication roundoff errors in digital filters. IEEE Trans. ASSP-**22**, 326–329 (1974)
43. M.G. Bellanger, G. Bonnerot, M. Coudreuse, Digital filtering by polyphase network: application to sample rate alteration and filter banks. IEEE Trans. ASSP-**24**, 109–114 (1976)
44. P.L. Iranescu, S. Rudeanu, *Pseudo-Boolean Methods for Bivalent Programming* (Springer, Berlin, 1966)
45. W. Schüssler, W. Winkelkemper, Variable digital filters. AEU **24**, 524–525 (1970)
46. D.H. Johnson, Variable digital filters having a recursive structure. IEEE Trans. ASSP-**27**, 98–99 (1979)
47. A.V. Oppenheim, W.G. Mecklenbrauker, R.N. Mersereau, Variable cutoff linear phase filters. IEEE Trans. CAS-**23**, 199–203 (1976)
48. L.R. Rabiner, R.W. Schafer, On the behaviour of minimax FIR digital Hilbert transformers. Bell Syst. Tech. J. **53**, 363–390 (1974)
49. C. Braccini, A.V. Oppenheim, Unequal bandwidth spectral analysis using digital frequency warping. IEEE Trans. ASSP-**22**, 236–244 (1974)
50. L.R. Rabiner, R.W. Schafer, *Digital Processing of Speech Signals* (Prentice-Hall, New Jersey., 1978)
51. J.M. Caywood, D.D. Buss, Frequency response of a multiplexed charge transfer delay line. IEEE J. Solid State Circuits. SC-**9**, 310–311 (1974)
52. D.D. Buss, D.R. Collins, W.H. Bailey, C.R. Reeves, Transversal filtering using charge transfer devices. IEEE J. Solid State Circuits. SC-**8**, 138–146 (1973)
53. T.G. Fotall, A.A. Ibrahim, G.J. Hupe, Minimum-phase CCD transversal filters. IEEE J. Solid-State Circuits. SC-**12**, 6, 638–642 (1977)
54. S.C. Dutta Roy, The ABCD's of digital signal processing. Students' J. Inst. Electron. Telecommun. Eng. **21**, 3–12 and 60–70 (1980)

Chapter 16
Signal Processing Applications of Charge-Coupled Devices

Abstract The charge-coupled device, first disclosed in the Bell System Technical Journal in 1970, is an analog, sampled data delay line and has made a significant impact in the field of signal processing. In this chapter, a brief introduction to the device is followed by the considerations, which make it so important and convenient in signal processing applications. A number of specific applications are outlined and their advantages and limitations are briefly discussed. In particular, the effect of the most important limiting factor, namely, charge-transfer inefficiency is discussed in some detail and several methods of compensation are outlined. The chapter concludes with an indication of the work being done at IIT Delhi on the various aspects relating to theory, fabrication and signal processing applications of the device.

Keywords Charge-coupled devices · Signal processing · Charge-transfer inefficiency · Transversal filters · Spectral analysis

16.1 Introduction

A charge-transfer device (CTD) can store a charge proportional to the sampled value of a signal and transfer it from one stage to the next under the control of external clock signals. These devices can store and transfer both analog and digital signals and provide a convenient, electrically variable delay for signal processing applications. Two types of monolithic IC realizations of these devices are available; these are known as 'bucket-brigade devices' (BBD) and 'charge-coupled devices' (CCD).

Bucket-brigade devices, which can be implemented in both discrete and IC versions, consist of a series of capacitors coupled by FET or bipolar switches, which can be switched ON or OFF in a controlled fashion by external clock pulses. Although the underlying principle of these devices was known for a long time [1], the first implementation in discrete version, using bipolar transistors as switches, was reported in 1967 [2]. The first integrated analog-sampled delay using MOS capacitors and MOSFET switches was developed in 1970 [3].

Source: S.C. Dutta Roy, A.B. Bhattacharyya, J.M. Vasi, L. Sankaranarayanan and Navin Kapur, "Signal Processing Application of Charge Coupled Devices," *JIETE*, vol 24, pp 400–418, October–November 1978.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_16

The ‘charge-coupled device’ was developed in the year 1970 by Boyle and Smith [4] at the Bell Telephone Laboratories. This device, for which no discrete equivalent exists, consists of closely spaced metal or poly-silicon electrodes on a silicon surface separated by a thin oxide layer. A signal can be stored in these MOS capacitors, as minority carrier charge, in the potential wells created by applying certain voltage pulses to these electrodes. By applying these pulses in proper sequence, the signal charge packet can be transferred from one electrode to the next. Due to their operational simplicity and novelty, CCDs stimulated a large number of workers throughout the world, to work on the development and applications of these devices. This resulted, in a very short time, in a large number of variations of the original 3-phase structure [4] and made a revolutionary impact in the fields of image sensing, signal processing and digital memories.

In the field of imaging, CCDs provide a compact and convenient device as it can store a charge proportional to the intensity of light falling on each of its cells and the charge can be transferred and obtained serially at the output in a natural way. This has led to the development of all solid-state TV cameras and facsimile systems [5–7]. Due to their low inherent noise characteristics, they are sensitive even at extremely low-light levels.

As digital memories, CCDs fill a gap between the low-cost discs and the fast-access-time semiconductor RAMs and are already available as single chips with as much as 64 K bit capacity. However, because of their dark current, the memories are volatile, and the stored information must be regenerated periodically. Access times of the order of tens of microseconds have been achieved by an intelligent organization of the storage elements and the peripheral circuitry.

Charge-transfer devices, and in particular, charge-coupled devices, are uniquely suitable for many signal processing functions and have found many applications in radar, TV, sonar and other areas of communications. They provide a simple and compact way of obtaining an accurate clock-controlled time delay of analog signals covering a range of a few 100 ns to several seconds. The analog signal is processed in its natural domain as there is no need to quantize the sampled signal. With split-electrode structures, compact transversal filters are realized, which have very sharp cutoff characteristics, and low in-band ripple. Devices suitable for signal correlation, matched filtering and spectrum analysis have been developed. These devices hold an advantage over other comparable techniques in terms of size, cost and power requirements.

As with all other devices, CTDs have their own limitations. Significant among these limitations is the dark current and the charge-transfer inefficiency, which are responsible for the degradation of the analog signal during processing. Dark current is the leakage current in CCDs produced due to thermally generated minority carriers. This adds charge at a constant rate, and eventually fills up the potential well, thus causing the signal to be swamped. To avoid this, the signal has to be swept out at a fast enough rate, which therefore limits the maximum delay achievable. Charge-transfer inefficiency is a more serious limitation of a CTD and results due to the incomplete transfer of a charge packet from one well to the next. This causes frequency-dependent attenuation and consequent distortion of signals. Efforts have been made to reduce

inefficiency by optimizing the device geometry and by improvements in the technology, e.g. by using buried channel techniques in CCDs. By these improvements, low values of inefficiencies are obtained even at high clock frequencies, and figures of the order of 5×10^{-5} at 135 MHz have been quoted [8]. Nevertheless, for a specified limit of distortion, the maximum number of stages that can be used are limited. Several ingenious circuit-theoretic techniques have also been developed for the compensation of the effect of transfer inefficiency on the propagated signal. The effect of transfer inefficiency and techniques for its compensation will be discussed in greater detail later in this chapter.

This chapter reviews the progress made in the field of signal processing with CCDs. As a prelude to this review, the basic operation and the physics of CCDs are given in the next section. The following aspects are discussed: the input and output schemes; a study of the effect of transfer inefficiency, and the various techniques for the measurement and compensation of its effect; a study of the transversal filters and various functions realizable with such filters; spectral analysis using CCDs; and examples of some of the numerous applications in radar and TV.

16.2 Physics of Charge-Coupled Devices

A charge-coupled device is a special kind of integrated circuit shift register, whose structure is similar to an MOS transistor with the difference that instead of one gate electrode, it has a number of them, between the source and the drain, as shown in Fig. 16.1. By alternately pulsing the series of gate electrodes, charge can be transported from the source to the drain. The transfer of charge is greatly influenced by the clocking waveforms. Originally, a three-level stepped clocking scheme was suggested for operating a three-phase CCD. Presently, a simpler two-level clocking scheme is used in practice, as described below. Let us consider a p-channel CCD. Initially, let the charge be under the electrode ϕ_2 at the instant t_1 when ϕ_1 and ϕ_3 are at the resting potential, say -5 V and ϕ_2 are at, say, -15 V (Fig. 16.2). At the instant t_2, when ϕ_3 turns on, charge is transported to the ϕ_3 well. For the duration in which the voltages on ϕ_2 and ϕ_3 are simultaneously on, charge will be shared equally between them (t_3). When ϕ_2 is being withdrawn, further charge transport

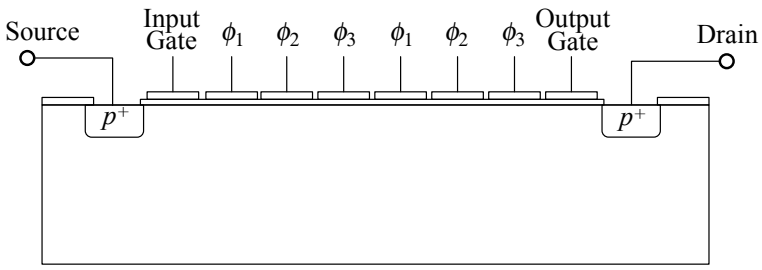


Fig. 16.1 Cross section of a 3-phase CCD showing the basic features

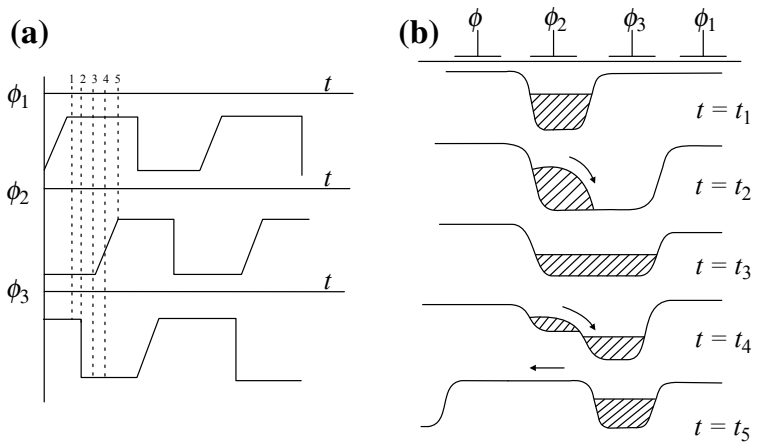


Fig. 16.2 Two-level clocking waveform for use in a 3-phase *CCD* and the potential diagram showing flow of charge, including possible 'back flow' at $t = t_5$

will take place from ϕ_2 to ϕ_3 (t_4). At the end of turn-off (t_5), a small fraction of charge remains under ϕ_2. What is important in this process is that a certain amount of charge flows backward. If the turn-off times are small and comparable to transfer time of the charge packet, the total charge will be significantly shared between the collecting electrodes (t_6). The 'Back Flow' of charge results in a loss of charge from the original packet [9].

A charge packet is not completely transferred to an adjacent electrode in the direction of transfer for many other reasons. First, due to the finite mobility of carriers, they require a certain time to travel the distance of an electrode and reach the next electrode. If the clocking frequency is high, all the charges are not in a position to be transferred, resulting in transfer loss [10]. Smaller electrode length and higher mobility of carriers (n-channels) will minimize such losses.

Surface state trapping is also a dominant mechanism of loss of charge from the travelling packet. This is so because surface traps are almost instantaneously filled but emptying time constants are dependent on energy, and there is a large dispersion in the value of these time constants. When the charge packet is transferred to the adjacent site, the trapped carriers in the surface states will gradually get released, but not all will be able to join the parent packet and those which cannot, get lost to the original signal [11, 12]. Such an inefficiency can be countered by using a non-empty zero known as 'Fat Zero'. The Fat Zero will attempt to keep the traps filled, thus making them inoperative during the presence of the signal.

Various models have been proposed for the discharging mechanism of the MOS capacitor containing a charge packet. The continuity equation describing the charge dynamics is given by

$$q\frac{\partial p(x,t)}{\partial t} = -\frac{\partial j(x,t)}{\partial x}. \tag{16.1}$$

The current consists of three components; (x is the direction of charge travel):

(a) Thermal diffusion,

$$j_D = qD\frac{\partial p(x,t)}{\partial x} \tag{16.2}$$

(b) Self-induced drift,

$$j_s = qp(x,t)\mu_p E_s(x,t) \tag{16.3}$$

(c) Fringing field drift,

$$j_f = qp(x_2 t)\mu_p E_f(x,t) \tag{16.4}$$

where q is the electronic charge, p is the number of charge carriers per unit volume, E is the electric field, j is the current density, D is the diffusion coefficient and μ is the carrier mobility.

In the absence of fringing fields, the thermal diffusion term dominates the long transfer times [10]. The fringing field becomes effective for short gate length [13]. The self-induced term manifests itself due to the charge gradient and is given by

$$E_s(x,t) = \frac{q}{C_{ex}}\frac{\partial p(x,t)}{\partial x} \tag{16.5}$$

The continuity equation may be solved numerically [14] or by charge-control approach [15, 16] the latter giving a closed-form solution. The charge-control method, though convenient to handle, suffers from the limitations that the space-charge distribution is assumed and the effect of transferred profile is not considered.

16.3 The Input and Output Schemes

16.3.1 Input Circuit

Input circuit is meant to introduce a charge packet proportional to the input signal into the device. The input scheme consists of a diode followed by two gates as shown in Fig. 16.3. The basic objective is to charge the surface capacitance below G_2, by a signal charge which subsequently is to be transferred. The function of the gate G_1 is to switch the diode on and off from the gate G_2. This provides the built-in sampling capability in the device.

There are two basic techniques of injecting/controlling the input charge.

(i) *Potential equilibration*: The surface potential below G_1 and G_2 are made equal to the source (diode) potential. The mobile charge under G_2 is given by

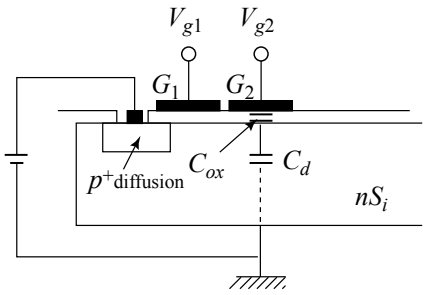


Fig. 16.3 Basic input scheme of a CCD

$$Q_{sig} = C_{o^x}(V_{g2} - V_T) \tag{16.6}$$

Equation (16.6) shows that the amount of charge can be controlled by (a) the applied gate voltage V_{o2}, and (b) the threshold voltage V_T which in turn depends on V_s.

(ii) *Current charging*: A certain amount of current is allowed to charge the surface capacitance below G_2 for a controlled duration. The diode (input), gates G_1 and G_2 may be regarded as constituting an MOS transistor as shown in Fig. 16.4. Charging takes place during the overlap time between the pulses on G_1 and G_2. The collecting potential under G_2 is the field-induced drain. Since the surface potential under G_2, simulating the MOST drain is a function of charge, the MOST may be characterized by a drain with time-varying drain voltage.

Amount of charge below G_2, in this scheme, is given by

$$Q_{sig} = \int_{t_1}^{t_2} I \, \mathrm{d}t \tag{16.7}$$

The operating conditions are such that I is saturated, making it time independent.

Here again, Q_{sig} may be controlled in two ways: (a) by controlling the duration of current charging out (b) by controlling the magnitude of current which depends on V_{g_1} and V_T, the latter being a function of the source potential V_s.

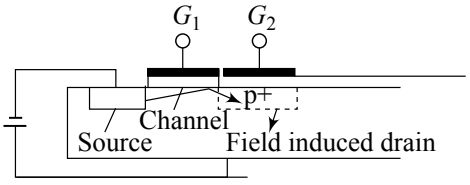


Fig. 16.4 MOST action between source and field-induced drain

16.3.2 Output Circuit

The simplest output scheme consists of dumping the charge out of the CCD through a drain diffusion, and measuring the current that flows at the output (Fig. 16.5). An output gate, usually maintained at a constant dc potential, is interspersed between the phase electrodes and the drain, to minimize clock pickup at the output.

On-chip amplification, and consequent improvement in noise characteristics, are provided by a scheme, by using reset and output MOS transistors [17–20], as shown in Fig. 16.6. In this scheme, the output charge is stored momentarily in the floating p^+ diffusion, which is connected to the gate of an on-chip output MOS. The output of this transistor, is, therefore proportional to the charge stored. To prepare the system for accepting the next charge packet, the reset gate is turned on, which resets the floating diffusion potential to a predetermined value.

More complex schemes, such as the use of correlated double sampling (Broderson and Emmons [19]) to reduce noise further, and the use of a floating gate, instead of the diffusion (Wen [20]), are also possible.

An important output method in which the outputs are tapped after every bit, and which is important in signal processing devices, is subsequently discussed in detail.

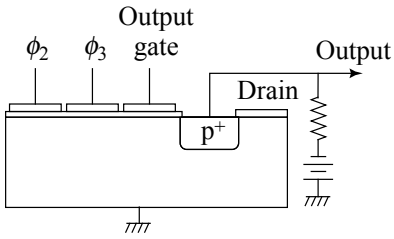


Fig. 16.5 Basic output scheme of a CCD

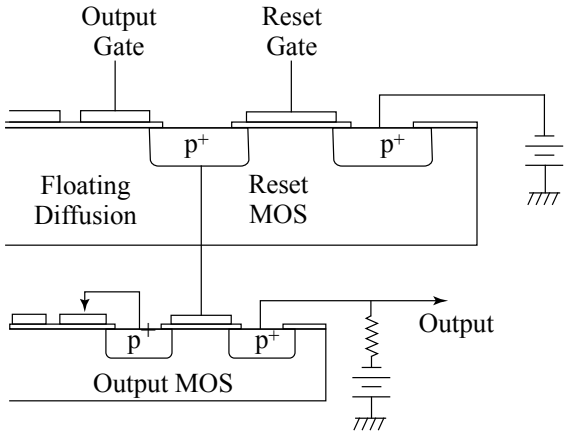


Fig. 16.6 Output scheme using a floating diffusion, reset MOS and output MOS to provide on-chip amplification

16.4 Characterization, Measurement and Compensation of Charge-Transfer Inefficiency

The basic function of a CTD is to give a time delay to the sampled analog signals. For obtaining a distortionless delay of signals, the signal charge packet integrity must be maintained, as it transfers from stage to stage. It is not, however, possible to obtain a complete transfer of the charge packet. A small proportion α of it is left untransferred in each transfer. Thus, in a p-phase device, $p\alpha$ times the charge packet remains back, and gets added to the next sampled signal charge packet. This proportion is called transfer inefficiency (ε) and the proportion of charge that gets transferred is known as transfer efficiency (η), their sum being equal to unity in most of the devices. In certain devices, a small proportion of the charge is lost irrevocably in which case their sum becomes less than unity. Thus,

$$\varepsilon + \eta \leq 1 \tag{16.8}$$

The transfer process is generally characterized by means of a simple block diagram of the CTD stage, as shown in Fig. 16.7. Here, z^{-1} represents an ideal delay of one clock period (T seconds), in the Z-transform notation. For sinusoidal excitation, we can write

$$z^{-1} = \exp(-j\omega T) \tag{16.9}$$

The transfer function of this stage (z_1^{-1}) is given by

$$z_1^{-1} = \eta z^{-1} / \left(1 - \varepsilon z^{-1}\right) \tag{16.10}$$

and the transfer function for the N-stage CTD is given by

$$H(z) = \left(z_1^{-1}\right)^N = \eta^N z^{-N} / \left(1 - \varepsilon z^{-1}\right)^N \tag{16.11}$$

The transfer function $H(z)$ of the N-stage CTD delay, given in Eq. (16.11), differs from that of an N-stage ideal delay, given by z^{-N}, due to the term $\eta^N(1-\varepsilon z^{-1})^{-N}$ which is responsible for attenuation and distortion. This distortion causes a train of spurious pulses in the unit sample response of the device. The frequency response deviates in magnitude as well as phase from that of an ideal delay. The unit sample response of the device can be obtained by taking the inverse Z-transform of Eq. (16.11), the result is

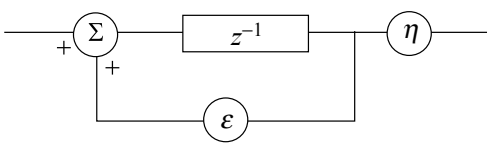


Fig. 16.7 Block diagram of a CTD stage

$$h(n) = \eta^N \Bigg[\delta(n-N) + \binom{N}{1} \in \delta(n-N-1) + \binom{N}{2} \in^2 \delta(n-N-2) + \cdots \Bigg] \tag{16.12}$$

where $\delta(n)$ is the unit sample function defined by

$$\delta(n) = \begin{cases} 1 \text{ for } n = 0 \\ 0 \text{ for } n \neq 0 \end{cases} \tag{16.13}$$

The delayed sample output, $\delta(n-N)$, which is the desired output, is attenuated by the factor η^N, and is followed by an infinite train of spurious pulses.

By replacing Z by exp. $(j\omega T)$ in Eq. (16.11), the frequency response of the inefficient CTD can be obtained as

$$H\left(e^{j\omega T}\right) = \eta^N e^{-j\omega T}\left(1 - \in e^{-j\omega NT}\right)^{-N} \tag{16.14}$$

For small values of $\in$, Eq. (16.14) may be approximated as

$$H(e^{j\omega T}) \cong \eta^N e^{-jN\omega T} e^{N\in(\cos\omega T - j\sin\omega T)} \tag{16.15}$$

The magnitude and phase components of Eq. (16.15) can be separated as

$$\left|H(e^{j\omega T})\right| \cong \eta^N e^{N\in\cos\omega T} \tag{16.16}$$

and

$$H(e^{j\omega T}) = \phi(\omega) \cong -N\omega T - N \in \sin\omega T \tag{16.17}$$

where ϕ stands for phase. The magnitude and phase responses of an ideal delay are given by unity and $-N_\omega T$, respectively. The inefficient CTD delay has a non-uniform magnitude response; its magnitude is unity at $\omega = 0$ and gradually reduces as frequency is increased, becoming equal to $e^{-2N\in}$ at half the sampling frequency (Nyquist limit). The phase response also deviates from the ideal by—$N\in\sin\omega T$. Both magnitude and phase responses violate the conditions for distortionless transmission.

Techniques for the measurement of inefficiency normally constitute measurement of this deviation either in time domain or in frequency domain, and by correlating this distortion with the expressions given in Eqs. (16.12), (16.16) and (16.17), the value of $\in$ is found out. In the time domain, strength of the spurious pulses in impulse response is compared with the strength of the desired impulse output [21–23]. From Eq. (16.12), one can see that the ratio of the second and the first output pulses is equal to $N\in$. Similarly, with a unit step input, expressions for $N\in$ can be obtained in terms of the ratio of the difference of pulse heights to the actual pulse heights at the output. These techniques normally require measurement of small pulse differences,

which is difficult and prone to several inaccuracies. To increase the magnitude of these differences in pulse heights, Levine [24] suggested a technique of multiple looping the signal, thus increasing the effective number of stages. In this method, the unit sample input fed to the CTD is made to pass through the device several times, by means of a feedback loop which feeds the output of the CTD to its input. This increases the effective number of transfers, thus making the trailing charge packet large enough for convenient measurement.

A simple method known as 'Pre-distortion technique' has been proposed by Dutta Roy and Das [25] for the measurement of the average transfer inefficiency of an N-stage CTD. This technique consists of introducing a known distortion in the square wave input to the CTD, so as to equalize two consecutive pulses at the output. The amount of distortion introduced gives a measure of the sum of inefficiencies of all the stages.

A variety of techniques have been suggested for reducing the effect of transfer inefficiency on the signal propagated through a CTD delay. By demultiplexing the sampled signal components into several parallel channels and combining them properly after the required delay, it has been shown that the effect of transfer inefficiency can be reduced [26, 27]. This is possible because the individual charge packet goes through less number of transfers. Another technique, several variations of which were proposed, is to store the signal in alternate stages only and to store zeros in the remaining stages [28–30]. The effective inefficiency, in this case, reduces to the order of $\in^2$.

Boonstra and Sangster [31] presented a simple technique for a first-order compensation of the output of a CTD for the effect of incomplete transfer. The technique employs a single feedback of coefficient $N\in$. The feedback, which is between two adjacent storage sites (in contrast with two adjacent stages), requires a negative amount of charge to be held at a storage site. Such a charge storage is possible with BBDs only, and is not physically possible with CCDs even with 'fat' zeros.

Thornber [32] presented a technique for compensation by linear filtering by means of a sampled data filter at the output of the CTD. The filter is matched to cancel out the distortion introduced by the transfer inefficiency. Two types of realizations of such filters have been proposed. The direct realization given in Fig. 16.5a uses an FIR filter and the feedback realization, given in Fig. 16.5b makes use of an IIR filter. In both the schemes, an infinite number of stages are required for achieving ideal compensation. It has been shown, however, that a good approximation can be achieved with only two stages in both cases for CTDs with $N\in \ll 1$ (Fig. 16.8).

A scheme for perfect compensation of the effect of transfer inefficiency has been proposed by Dutta Roy and Arora [33]. The scheme employs negative feedback, from the output of each stage to the input. It has been shown that any given transfer function can be realized exactly by this method using inefficient CTDs. The scheme consists of replacing z^{-1} in the given transfer function $H(z)$, by an expression consisting of terms of z_1^{-1}, which can be obtained from Eq. (16.10). This yields a new transfer function $G(z_1)$ which can be realized using inefficient CTD's. It may be recalled that z_1^{-1} represents the transfer function of the inefficient CTD. When used to realize an N-stage ideal delay, this requires negative feedbacks from the output of each of the

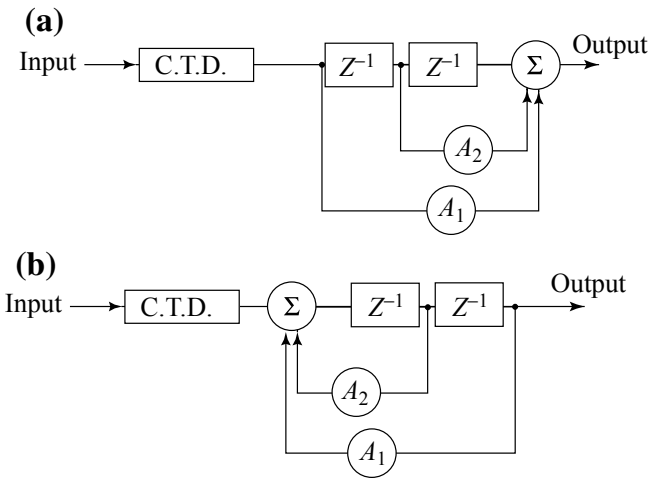


Fig. 16.8 Linear filtering techniques for the compensation of transfer inefficiency **a** direct filter technique **b** feedback filter technique

N stages to the input with coefficients $(h_1, -h_2, \ldots, -h_N)$ which are all functions of $\in$ and an amplifier of gain $A = \eta^{-N}$. Figure 16.9 gives the block diagram for this realization.

The scheme has been extended to cover the case of a CTD with non-identical stages by Dutta Roy and Das [34]. An 18-stage CTD with $N\in = 0.8$ and with non-identical stages, consisting of 2-stages of a discrete BBD and a 16-stage CCD, has been compensated by limiting the feedbacks to the first two stages only. This results in considerable improvement in the frequency response of the device.

Compensation schemes mentioned above require external circuitry for their implementation. Transfer inefficiency being a function of several parameters like clock frequency, wave shape, temperature, the feedback coefficients, which are functions of $\in$ require resetting, which is difficult to implement in practice. Due to these reasons, the tendency seems to be to live with it and till now compensation has been rarely adopted.

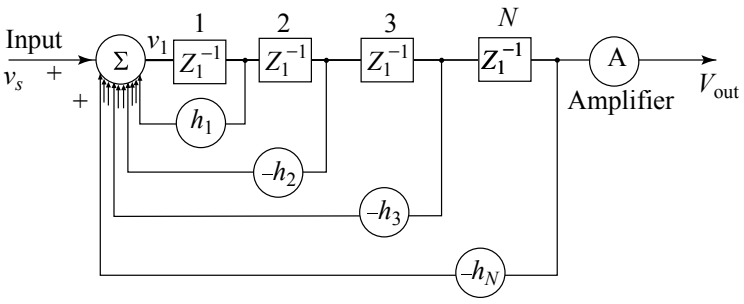


Fig. 16.9 Feedback compensation scheme for the realization of N-stage ideal delay using inefficient charge-transfer devices

16.5 CCD Transversal Filters

Transversal filters as a general class use time-delayed waveforms of a given input signal, weighted and summed with the proper signs to generate a wide range of filter characteristics [35].

Charge-coupled delay lines, when provided with taps along the channel, provide a versatile implementation of transversal filters; this is the building block for most systems using CCDs for signal processing [36, 37]. The output of an N-stage transversal filter is given in (Fig. 16.10).

$$V_0(nT_0) = \sum_{K=0}^{N-1} h_K V_K = \sum_{K=0}^{N-1} h_K V_i(nT_c - KT_c)$$
$$= \sum_{K=0}^{N-1} h_K V_i\{(n-K)T_c\} \tag{16.18}$$

where $h_K =$ Kth tap weighting coefficient, $V_K =$ Voltage on the Kth tap at the instant of time, nTc = value of the input voltage K sample periods earlier, $V_0 =$ output signal (sampled) at the instant of time and $T_c =$ delay per stage $= 1/f_c$ and $f_c =$ clock frequency.

This is a non-recursive technique, i.e. no feedback elements are used. In general, a transversal filter cannot yield as sharp a response characteristic as a recursive filter for a given number of delay stages, However, it has the following advantages which make it attractive [38]: (i) It is always stable, (ii) It can be designed to give an exactly linear phase characteristics and (iii) Response characteristics depend only on the relative values of the coefficients and not on their absolute values, thereby making the frequency response less sensitive to coefficient errors.

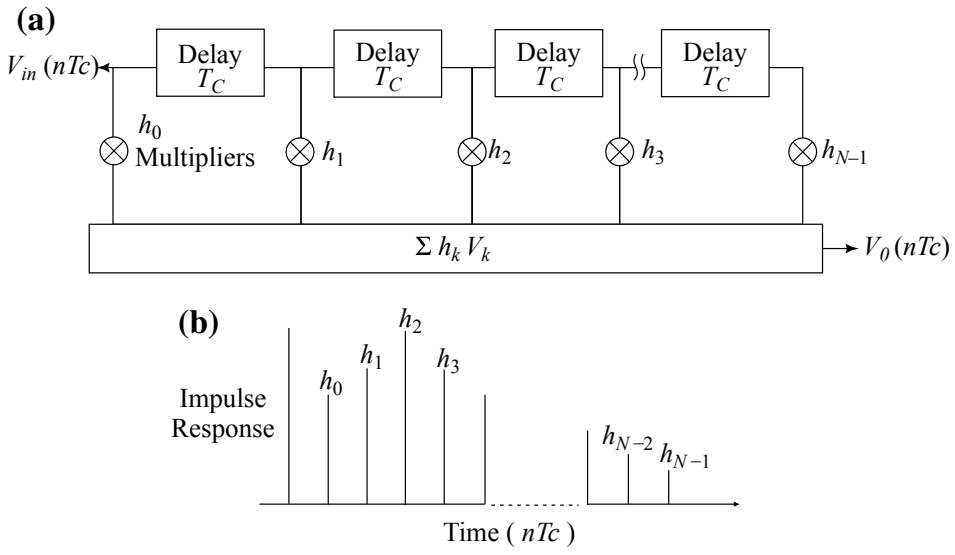


Fig. 16.10 Depicting an N-stage transversal filter **a** block diagram **b** impulse response

The most common implementation of a transversal filter employing CCDs is the split-electrode technique [39–44]. This structure provides fixed tap weights, and these are obtained by sensing the image charge of the signal in the transfer channel, by using sensing electrodes placed over the channel. To give the appropriate weight at each tap (h_K), these electrodes are split into two parts, which are connected independently (Fig. 16.11a) to two sense lines. The position of the cut defines the tap weight at this point which varies from + 1 to −1 as the slot position is varied from one edge of the charge-transfer channel to the other, i.e. the lengths of the two sides of the split electrode are chosen in the ratio $(1 + h_K)$: $(1 - h_K)$, so as to produce the net weight h_K at each node. The weighted sum of all the image charges of the signal packet in the register can then be sensed on the two sense bus lines, either by using a differential current metre, or by sensing the change in voltage when signal charges flow under the floating-sense electrodes.

The accuracy of the split-electrode technique is limited by several factors [45], viz. (i) Voltage dependence of the depletion capacitance; in this scheme, it is assumed that the oxide capacitance is much larger than the depletion capacitance and the latter can be considered constant, (ii) Variations in oxide thickness, (iii) Accuracy of locating electrode gaps, (iv) Transfer inefficiency and (v) DCI (differential current integrator) noise, due to (a) KTC noise [13] on the split-electrode clock lines, and (b) voltage amplifier noise. Of these, (b) is predominant in the devices tested to date.

The above sources of error will place a practical limit on the accuracy of the coefficients which can be realized. In frequency-selective filters, any coefficient errors will tend to have a more adverse effect on the stopband, rather than in the passband

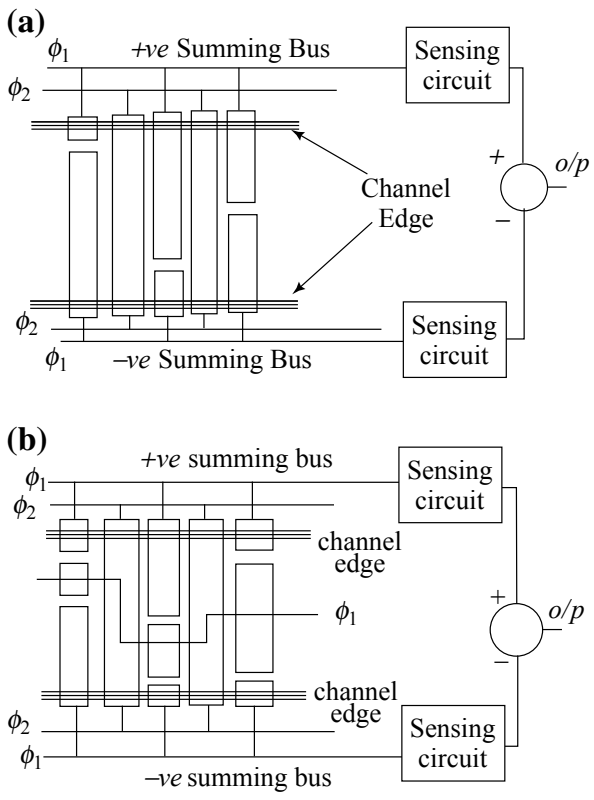


Fig. 16.11 Split-electrode CCD transversal filter
a single split-electrode weighting structure
b improved, double-split electrode weighting structure

regions of the filter. Thus, the technique is limited to some practical value of attenuation which can be realized. However, one could cascade several filters to obtain the desired attenuation. In particular to reduce the effect of the DCI noise, a new 'double-split electrode' structure has been fabricated [46] which provides improved filter performance over the signal-split approach by reducing the common mode signal and noise pick up (Fig. 16.11b).

The other tapping technique is the floating-gate reset (FGR) method [47]. This is also a non-destructive sensing scheme and is shown in Fig. 16.12. The relevant equations are

$$\Delta V_{\text{out}} = \frac{Q_{sig} C_{ox}}{C_{ox} C_L + C_D C_L + C_D C_{ox}} \tag{16.19a}$$

where

C_{ox} Total oxide capacitance under the tap electrode
C_D Total depletion capacitance under the tap electrode
C_L Total external load capacitance on each tap

The above expression can be approximated as

$$\Delta V_{out} \cong Q_{sig}/C_L \tag{16.19b}$$

for $C_{ox} \gg C_D$ and $C_L \gg C_D$. This is a very good approximation when the substrate resistivity is large (greater than 10 Ω cm).

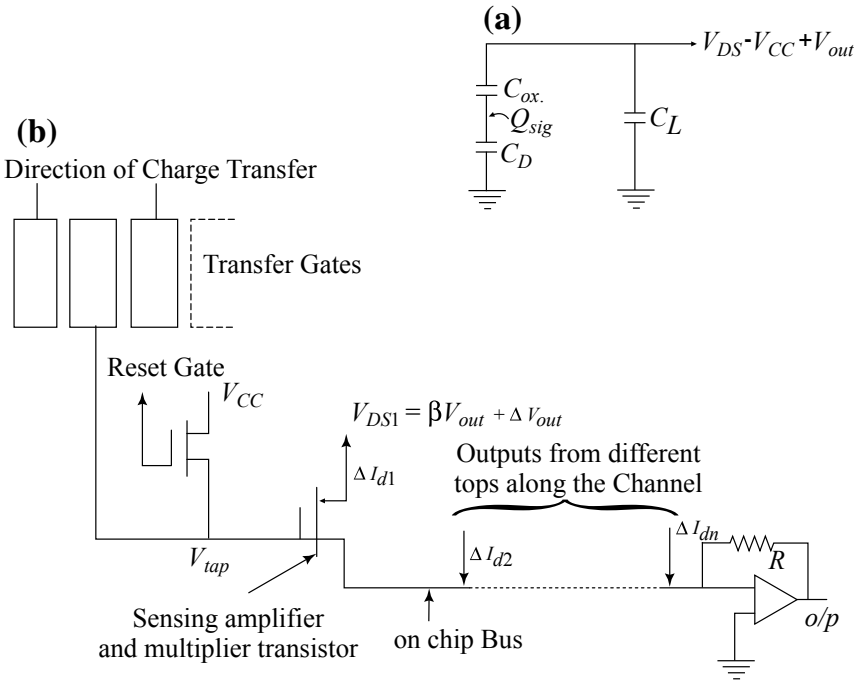


Fig. 16.12 **a** Capacitive model of floating-gate tap. **b** Circuit implementation of floating-gate tapping four-quadrant multiplication and convolution summing

A conventional clocked electrode in the CCD structure is replaced by a floating-gate reset-sense electrode. The electrode is first reset to a bias voltage whilst there is no signal charge underneath it. It is then isolated and any charge subsequently injected under the electrode (as a result of CCD clocking action) causes a related change in the electrode potential. The signal may be sensed via an MOS transistor amplifier to provide a voltage output at a low impedance. Now, for general-purpose signal processing, four-quadrant multiplication of the tapped signal and the tap weights are required at each convolution point. An economical multiplication technique is based upon the essentially linear transconductance of an MOS transistor operating in the non-saturation region. A first-order expression for the drain current of such a transistor is given by

$$I_D = \beta\{(V_{GS} - V_T)V_{DS} - \tfrac{1}{2}V_{DS^2}\} \tag{16.20a}$$

from which the change in drain current obtained upon disturbing the gate voltage by an amount ΔV_{GS} is ($\Delta V_{GS} = \Delta V_{\text{out}}$),

$$\Delta I_D = \beta V_{DS}\Delta V_{GS}. \tag{16.20b}$$

This provides an accurate four-quadrant multiplication. β is a process-dependent parameter and V_{DS} can be adjusted externally to provide the correct tap weight. Changes in drain current of the multiplier transistors at each tap can then be summed via a sense bus bar feeding an operational amplifier to provide the output V_0 (KT_c). Since the voltages V_{DS} are programmable, this structure offers a promising realization of a powerful signal processing element. Using this technique, fabrication of a completely monolithic 64-point programmable transversal filter has been completed [48]. Initial results confirm operation of the device in both frequency and matched-filtering configurations over a range of frequencies suitable for sonar and other sub-megahertz applications.

16.5.1 Limitations of This Technique

From the above, we see that the programmable filter consists of a signal register, a reference register, a multiplier and a summer suitably interfaced to provide the convolution sum. The sources of error in each of these blocks are the following.

Signal Register

(a) transfer inefficiency and
(b) dark current

Reference Register

(a) information decay due to charge leakage from the sample-and-hold capacitor (external to the CCD chip) and

(b) capacitive breakthrough of the sample waveform into the hold capacitor via gate overlap capacitance.

Multiplier

(a) changes in drain current may modulate the reference voltage due to the finite output admittance of the reference amplifiers and
(b) variations in the value of β of the multiplier transistor due to processing non-idealities.

Summer

The summing amplifier is a limiting factor in terms of speed in that its output must settle within the time period for which the signal is true. Although the intrinsic CCD/MOS circuitry may be designed to operate at clock frequencies up to 10 MHz, it is difficult to achieve sampling rates at the summing amplifier much above I MHz without resorting to bulky, high-power amplifiers.

Effects of these error sources at the output of the convolver are as follows. Charge-transfer inefficiency ($\in$) in the CCD register degrades the signal information at high frequencies, this being a cumulative effect with increasing $N\in$ products. In terms of frequency filtering applications, the net effect at low frequencies is a shift in the transition edges between passband and stopband of a factor $(1 - \in)$, regardless of the number of filter stages. The effect upon correlator performance is signal-dependent but $\in$ is known to cause a slump in the correlation peak and an increase in the sidelobe levels. Random gain errors in the multiplication process are analogous to tap weight errors in split-gate CCD filters and as such may be expected to impair stopband suppression in frequency filtering applications. Correlation applications are more tolerant to these errors due to the processing gain of the filter.

The signal distortion term imposed by the finite output admittance of the multiplier stages causes dc and second harmonic distortion for a pure sinusoidal signal. In correlation applications, additional peaks are to be expected whenever a match is detected between the generated harmonic signal component and the reference waveform. Refreshing and subsequent decay of reference values results in a modulation noise of the output waveform, since it is not generally synchronized with the incoming data.

Finally, although several potential error sources have been identified, not all will be operative for a single application and quantitative effects depend on the signal.

Having gone over the chip implementations of the CCD transversal filters, let us now consider some of the important signal processing applications of CCDs.

The transversal filter performs real-time convolution by non-destructively sampling the charge packet at each stage of delay, multiplying each sample by a weighting coefficient and summing the resulting products at each sampling instant. Hence by proper selection of the weighting coefficients, one can achieve a very wide variety of filtering or correlation functions such as low-pass, bandpass, band-stop, matched filters and cross-correlators. Note that the tap weights $h_K = h\ (KT_c)$ arc nothing but sequential elements of the impulse response of the tapped delay line. Also, all of these functions are tunable, i.e. the cutoff frequency, centre frequency and bandwidth

can be varied by simply changing the clock frequency. There exists a vast literature on the techniques of designing digital nonrecursive (FIR) filters, which are directly applicable to the design of CCD weighting coefficients [49]. We shall now take up some specific filter types [41, 42, 50].

16.5.2 CCD Low-Pass Filters

By choosing the h_K as equispaced discrete values of the sin x/x function, a low-pass filter can be realized which will have repetitive passbands spaced f_e apart in the frequency domain. Applications of CCD low-pass filters include a channel filter for a frequency division multiplex communication system, an anti-alias filter for a time division multiplex communication system, an anti-alias filter for a digital monitoring system and an interface between two asynchronous sampled data systems, e.g. a video communications system [42, 43].

Low-pass filters, as such, are required in a large number of systems. In some cases, the required filter characteristics are easily obtained with simple passive networks but in many cases system performance can be improved, if higher performance filters are used. Until now, it may not have been economically attractive to use higher quality filters in these applications, but with the availability of low-cost CCD filters having very sharp cutoff characteristics, and low in-band ripple, many opportunities now exist for upgrading the quality of low-pass filters.

Sampled data analog systems, for example, all require a low-pass filter for the elimination of alias responses. This filter must have a passband that is adequate for the signal bandwidth, but for typical applications, it must be down by 40–50 dB at the Nyquist frequency. In the past, this has been accomplished by oversampling the data by a factor of 3 or 4, so as to move the Nyquist frequency sufficiently far above the signal passband, so that a relatively simple filter can suffice. Now, although this practice simplifies the filter, it requires that three or four times as many samples be stored, processed or transmitted, as are actually needed. More storage means a larger memory, more processing means increased CPU loading and more samples to be transmitted means increased channel bandwidth, all adding up to increased cost. In such cases, a CCD low-pass filter with a sharper cutoff, and a reasonable stopband attenuation provides a convenient solution.

16.5.3 Bandpass Filters

Bandpass filters have numerous applications in speech and data processing. We discuss briefly the design of a bandpass filter and illustrate the application of window functions for apodization to control sidelobe response (out-of-band rejection). Truncation of the impulse response of an ideal bandpass filter (in order to permit an economical realization) inherently leads to lower rejection in the stopband, i.e. higher

sidelobe levels. Windowing techniques modify the coefficients of the ideal filter so as to minimize the effects of termination of the response at a finite number of samples. Since weighting coefficients can be easily modified by relocating position of the slot in the split-electrode structure, high-quality bandpass filters are obtainable using CCDs.

The weighting coefficients for a bandpass filter with a narrow bandpass centred at f_0 can be written in the form

$$h_K = \omega_K \sin(2\pi K f_0/f_c) \tag{16.21}$$

where ω_K's are samples of a window function. The particular selection of a window function depends on the desired passband shape and sidelobe levels. The price one pays for higher sidelobe rejection is a greater width of the passband. Practical results show that a 101-stage filter gives a sidelobe level of −13.2 dB and a 3-dB bandwidth of 3.5% of the centre frequency; however, the apodized filter gives a maximum sidelobe of −42.8 dB and a 3-dB bandwidth of 5.1% of the centre frequency [51].

With CCD transversal filters, it is difficult to achieve filters of high Q because of the finite time duration of the impulse response, and because tap weight error limits the out-of-band rejection. Recursive filters are potentially useful for achieving high Q as has been demonstrated using BBDs with off-chip feedback (Buss et al. [36]). CCD transversal filters are most useful when the magnitude and phase of the frequency response must be accurately determined and can be used for signal levels down to 50 V together with a large dynamic range [37].

16.5.4 Matched Filtering

Consider a filter whose weighting coefficients are chosen to represent an impulse response $h(t)$, which is matched to a particular signal waveform $V_s(t)$

$$h(t) = (1/A)V_s = (T_d - t), 0 < t < T_d \tag{16.22}$$

where A is the amplitude of the signal waveform and $h(t)$ is normalized so that its maximum amplitude is unity. When V_s is applied to an ideal filter ($\in = 0$), the output is given by

$$V_{out}(nT_c) = A \sum_{K=1}^{M} h_K h_{M-n+K} \tag{16.23}$$

and the correlation peak occurring at $t_M = MT_c$ is

$$V_{out}(t_M) = A \sum_{K=1}^{M} h_{K^2}. \tag{16.24}$$

Charge-transfer loss, however, introduces a mismatch between the filter and the signalling waveform and results in loss of sensitivity, i.e. degradation of peak-to-sidelobe ratio [42, 52].

16.5.5 Hilbert Transformer

A Hilbert Transform consists of a convolution with t^{-1} and can be implemented with a CTD transversal filter [36] having weighting coefficients,

$$h_k = \left\{ K - \left(\frac{M+1}{2} \right) \right\}^{-1}, \quad K = 1, \ M; \ M \text{ even} \sqrt{a^2 + b^2} \tag{16.25}$$

Such a filter can be used to generate single sideband signals and has a wide potential application in communication equipment [5, 53]. It may be pointed out that, so far, the complexity of modulation equipment was the major inhibiting factor towards greater use of SSB transmission.

16.5.6 Adaptive Filtering

Due to the changing characteristics of a telephone channel, adaptive equalization is required in a MODEMs, as to permit error-free detection of data [42]. This requires a variable tap weight convolution filter (VTWCF) which is most easily implemented using CCDs with FGR sensing and four-quadrant multiplication, as described earlier. Such a filter would be useful not only in MODEMs, but in other adaptive filtering applications such as voice recognition, adaptive beam forming, remote intrusion detection [54].

16.5.7 Complex Coding

CTD filters can be used to generate as well as to detect arbitrary waveforms (Buss and Bailey [41]). This presents the possibility of choosing M waveforms to represent $K = \log_2 M$ bits of information. This type of coding employed in MODEMs, but the choice of waveform is limited by the available equipment and the waveforms actually used are not optimized to the transmission medium. CTDs offers a great deal of flexibility in the selection of waveforms. In addition, the relative case with which CTD filters can be matched to arbitrary waveforms may make possible the use of 1-ary (instead of binary) communication, where it has not been previously practical.

16.5.8 Interference Rejection Filter Using CCDs

A multiple notch filter has been realized, which can be used to track and remove a periodic interfering signal and its harmonics [55]. It has also been shown that to a first-order filter, notch depths are not dependent on transfer inefficiency and tap weight errors, provided the inefficiency value is spatially uniform, i.e. is constant throughout the CCD. Calculations show that even for inefficiency products as high as 0.2, notch depths of 40 dB are possible.

16.5.9 Linear Regression Analysis Using CCDs

A common requirement in signal processing is that of detecting a slowly varying signal against a background of noise. In a typical situation, if a system of low-pass filtering and threshold detection is used, there is the problem of matching the threshold to the noise level, and a slowly changing baseline can therefore cause considerable errors. If, however, several successive samples of the signal are available over the region of interest, statistical techniques of straight-line fitting can provide an estimate of the continuous waveform [56]. Equations for straight-line fitting reduce to the manipulation of weighted sums of sampled data, which can be easily implemented using transversal filters.

16.5.10 Radar Display System Using CCDs for Time Expansion

A system using the Fairchild CCD 321 has been built, which allows real-time signal processing, and display of radar data; the video data is read in parallel into eight CCDs, and read out sequentially at a much slower rate (time expanded), to match the subsequent display circuit capabilities [57].

16.5.11 Waveform Generation Using CCDs

The impulse response of a transversal filter has the same form as the tap weights of the CCD. It is realized by propagating a single charge pulse along the CCD. In this way, arbitrary waveforms of finite length arc easily generated at the summing point of the tap weights. Amplitude is controlled by amplitude of the impulse, and the rate may be controlled by the clock frequency.

16.5.12 Cost and Performance Trade-Offs in Various Approaches to Signal Processing

Both analog and digital methods have been used for analog signal processing. In the analog approach, a desired response (specified in the frequency domain) is approximated by a number of poles and zeros in the complex frequency plane, and these are implemented with passive discrete components, and possibly some operational amplifiers. This approach has three important advantages [39, 40, 50]: (a) negligible power is consumed, (b) involves low cost and (c) the filter realization is direct.

There are also some disadvantages: (a) separate filter hardware must be dedicated to each channel, i.e. a single filter cannot be multiplexed between several channels, since the processing is 'all analog', and not on the basis of Nyquist sampling, (b) the only designs that can be implemented are those which require a modest number of poles and zeros and (c) values of the discrete components must be precisely determined. Often they must be individually adjusted and temperature drift and ageing can cause problems, (d) impedance matching has to be achieved and (e) very low-frequency realizations are a problem because the R, L, C elements become expensive and accuracy is impaired.

In the digital approach, the signal is sampled and digitized and is processed numerically. This approach solves the problems cited above, but has two major drawbacks: (a) A/D conversion is required which adds considerably to cost and complexity and (b) the multiplication is performed in a digital multiplier, which apart from its cost consumes a large amount of power and time. However, the high accuracy of this technique is a very compelling reason for adopting it.

The importance of charge-transfer technology for analog signal processing is that it combines the best features of both approaches—namely, it permits an arbitrary impulse response to be implemented easily, it provides high precision, without requiring individual adjustment of components, and A–D conversion is inherent and exact. Also, the additional circuit complexity required to convert a CCD delay line into a transversal filter is marginal and the computational power of a CCD transversal filter, when measured in number of multiplications per second is very large. Moreover, time delays are stable with time and temperature depending only on the master clock (as for digital filters).

However, the CCD filter retains to some extent the problem of realizing very low-frequency (Hertz range) filter functions, because of storage time (dark current) limitations, which arise mainly from material imperfections and process-induced defects. A significant development effort is in progress to alleviate this problem. At this point, it is pertinent to mention another charge-transfer device similar to the CCD—the BBD [45, 58–67]. This is a sequence of MOS transistors coupled together by diffusion enhanced Miller capacitance. Although these devices do not operate at frequencies as high as CCDs, they have better low-frequency performance since, they include active devices. Transfer efficiencies are, however, poorer for BBDs.

Further, since the impulse response of a CTD filter can be chosen arbitrarily, CTD filters can be matched to arbitrary signalling waveforms in much the same

way that surface wave devices are used to perform matched filtering, CTDs do not compete directly with SWDs for this application, however, because SWDs have higher bandwidth (GHz) and limited time delay (tens of microseconds). CTDs, on the other hand, are capable of processing signals up to 1 s. in time duration but are more restricted in bandwidth.

Compared with digital implementations, CTD filters are less flexible and are limited to time-bandwidth products of 10^3 or less. On the other hand, when CTDs can be used for a particular matched-filtering function, overwhelming advantages in cost, power consumption, size, weight and reliability can be realized. Let us take an illustrative example:

It is required to implement a 100-point convolution filter with a dynamic range which requires 8-bit digital logic. An analog signal is to be sampled at 100 kHz, filtered and presented at the output in analog form. When this filtering operation is implemented using TTL, the cost and power consumption are determined primarily by the 8-bit A/D converter ($2 000, 5 W) and the digital filter consisting of approximately 400 TTL networks costing $5 per piece and dissipating 100 mW each. In evaluating the cost of the CTD implementation, the important factor of part volume enters. The digital filter can be implemented with standard catalogue items, whereas, the CTD is a custom IC and as such the cost is dependent on the volume sold since the development cost is high up to $80 000. Assuming a sale of 100 000 devices, the following cost comparison emerges

	Digital	CCD
Cost:	$4 000	$56
Power	45 W	600 mW

The cost of the CCD implementation includes the cost of clock drivers, output amplifiers and output sample-and-hold circuits—this comes to about $50. This comparison is an oversimplification of real systems and it is heavily weighted in favour of CTDs, because it is based upon a simple convolution filtering operation for which CTDs are ideally suited. However, it illustrates the saving in cost and power, which is possible with CCDs—and this cost and power advantage has been responsible for widespread use of CTDs in military communication systems.

The advent of the microprocessor and LSI memory means that powerful and versatile filtering systems can be programmed and controlled (Mavor and Denyer [48]). Such an arrangement is extremely important in the signal processing context as any desired filtering function can be realized by simply programming its impulse response. Interfaced with a microprocessor, this system could provide intelligent adaption of the filter response, to optimize performance. Potentially, new areas include sonar, instrumentation and mobile communication equipment.

Single-chip (monolithic) programmable transversal filters have an enormous potential in applications where space and power arc crucial and high computing precision is not essential.

16.6 Spectral Analysis Using CCDs

Analysing a signal for its different frequency components is a vital area of signal processing. The bulk of spectrum analyzers, use a sweeping local oscillator to cover the frequency interval of interest and so require continuous input waveforms. Furthermore, the time required to scan the frequency interval with a narrow band filter can be substantial, being proportional to the ratio of the swept bandwidth to the square of the filter bandwidth. There are a number of situations where the Fourier transform of a finite number of time samples is required—for example, in Sonar and Radar target signature analysis. The conventional spectrum analyzer is of limited use in these areas, and other techniques are required. The most popular alternative is to use a digital processor, which evaluates the discrete Fourier transform (DFT) of a set of samples, by using a fast Fourier transform (FFT) algorithm [68]. However, such processors may be too slow to allow the data to be processed in real time, or consume excessive power and cost so much so that liberal use of them cannot be made. The FFT algorithm is most efficient when the number of samples N is a power of 2—a situation which can always be achieved by augmenting N with zeros.

An alternative approach [69] restructures the DFT in such a way that the bulk of the computation is in the form of a complex convolution of the input data with a set of fixed weights. Such a convolution is very conveniently performed in a CCD transversal filter, fabricated on a single substrate, giving substantial savings in size, weight and power consumption over a comparable digital solution. This modified approach known as the 'Chirp-Z transform' algorithm has been implemented practically using CCDs with 100 taps, which is capable of working up to 1 MHz clocking frequency [70]. The input data samples, taken at time intervals of T_c, propagate along a set of storage locations in discrete steps under the control of an external clock. At each storage location, the signal is non-destructively tapped—that is, the tapping process leaves the propagating signal unchanged. The signal from each tap is then multiplied by a weighting coefficient before the final summation, which forms the output. It is the simultaneous execution of these multiplications, which makes the transversal filter so powerful. The CCD Discrete Fourier Transform Processor can also be used to compute the power density spectrum. It may be pertinent to point out that the Chirp-Z-transform (CZT) is one of the classes of fast Fourier transforms [71] that has had limited use in the past, because it does not provide any significant advantages over the more common FFTs in a digital implementation; however, it is the preferred algorithm for CCD implementation [44].

Further, a significant reduction in hardware (over that of the CZT) are possible with the CCD implementation of the Prime Transform (Jack et al. [72]) for purely real input data, the prime transform uses two CCD filters as compared to four for the CZT filters. CCD prime transform processors, which give an output accuracy of 2% rms have been reported. Although the accuracy of CCD transforms is inferior to the digital FFT, they offer low-power compact analog processors which can operate at

bandwidths up to 5 MHz in real time. It is therefore in systems applications which require such engineering advantages at the expense of transform accuracy, e.g. sonar and speech processing (analysis) that the analog CCD transforms will find application [51, 70, 72–76].

16.7 MTI Radar and Television Ghost Suppression Applications

Charge-transfer devices provide a cheap serial memory, replacing complex digital circuits in MTI (moving target indicator) and ECM (electronic countermeasure) radar applications. The MTI radar system accepts the reflected signals from moving targets only, and rejects all signals from stationary objects, by making use of simple cancelling filters. Digital signal processing, with its high internal stability, has become an integral part of new radar systems, which need sharp out-off Notch filters (cancelling filters). The major disadvantage of high-performance digital systems, as discussed earlier, has been the expense and complexity of A/D converters, whenever a large dynamic range is required in conjunction with high bandwidth (5 MHz or more). The solution to this problem is to use charge-coupled devices (CCDs) as low-cost analog-sampled data processors [77]. In CCDs used as analog data processors, the input signal is Nyquist-sampled, and the samples are stored and processed as discrete analog packets of charge. This eliminates the need for A/D conversion entirely or in the case of a CCD preprocessor, reduces the cost of subsequent A/D conversion to the more economical realm of low-bit high-speed technology. The following device parameters, furthermore, favour the application of CCD delay lines in MTI filters (Bounden et al. 1974): (a) Signal storage and delay times up to ≅100 ms, (b) Clock rates up to 10 MHz, (c) Dynamic range >40 dB, (d) Serial access to data is simple, (e) Simpler, faster, smaller and cheaper than digital systems and (f) Delay is variable, and can be accurately controlled.

16.7.1 Simple Cancelling Filter

Figure 16.13a shows a simple cancelling filter incorporating a time delay element (τ_c) introduced by a CCD analog delay line. It has a frequency response as shown in Fig. 16.13b in which cancellation occurs at frequencies, which are multiples of pulse repetition frequency (f_r) where $f_r = 1/\tau_c$ Such a filter can be used as a coherent MTI for a pulse radar system, if τ is the pulse repetition interval (PRI) of the radar transmitter. Thus in such a radar set, the reflected return signals at radio frequency are mixed with a reference signal at the transmitter frequency. If the target giving the reflection is moving, there will be a change in phase of the returned signal between consecutive transmitter pulses, and hence, the mixer output shows a beat at the

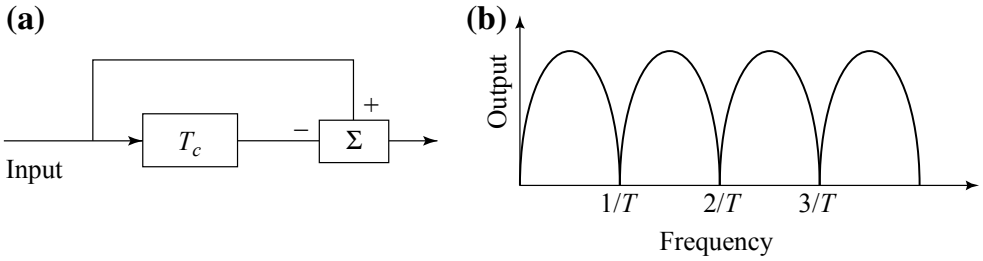


Fig. 16.13 **a** Block diagram of a simple cancelling filter. **b** Frequency response of the cancelling filter

Doppler frequency. If this signal is applied to the cancelling filter, the output will be obtained only for moving targets, because the subtraction cancels the 'clutter' signals, which are stationary from pulse to pulse. Targets with speeds corresponding to Doppler frequencies, which are multiples of f_r (PRF) are also cancelled. The extent to which suppression of signals from stationary objects and of Doppler frequencies, which are multiples of f_r, can be achieved over the full video bandwidth of the system is determined by the characteristics of the delay line. The total clutter power reduction can be improved by having a broad clutter rejection notch, obtained by placing two cancellers in series. This configuration processes three pulses simultaneously, and the magnitude response will be that of a squared sine wave. In the simple cancelling filter, the input–output transfer function is written as

$$H(j\omega) = 1 - e^{-j\omega T} \tag{16.26}$$

The magnitude response of this is

$$|H(j\omega)| = 2|\sin(\pi f T)| \tag{16.27}$$

This is shown in Fig. 16.13b. Preliminary experimental circuits with CCD delay lines demonstrated notch depths of up to 40 dB in the output frequency response characteristics [78]. A simple CCD canceller using an N-element CCD takes N samples of the incoming waveform every pulse repetition interval, and can cover the acquisition of targets in N range elements (range gates).

16.7.2 Application to Practical MTI Filters

In practice, MTI filters usually operate with several hundred range gates and with a PRI of the radar of 1 ms, so that the use of a serial–parallel–serial or linear CCD delay line is appropriate. The total required delay T could be several milliseconds, depending on system requirements, which is within the limits imposed by the dark current, which adds a charge to the charge packets as they travel through the CCD. To minimize target smearing, the charge-transfer inefficiency must be kept small.

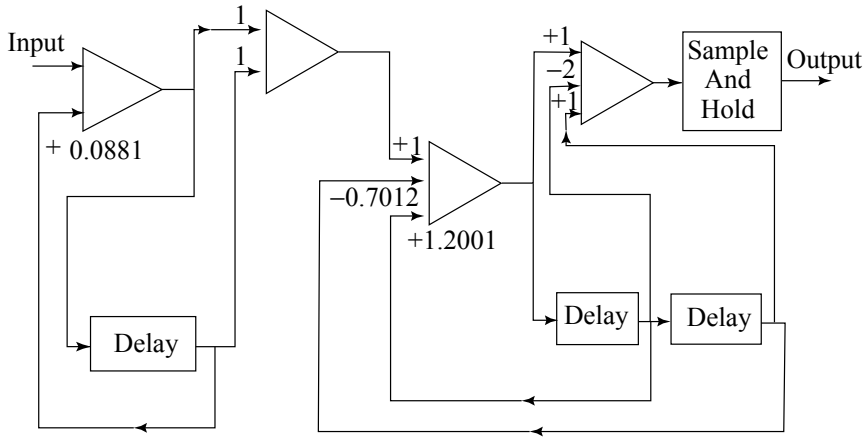


Fig. 16.14 3-pole Chebyshev analog filter for Radar MTI applications using CCD delay lines

Its allowed value depends on the number of range gates and the practical frequency response required. Also the smearing effect is found to increase when the Doppler frequency is near the edge of the filter passband. The linearity of the CCD output as a function of input is important as it determines the cancellation ratio attainable. In a realizable MTI filter using operational amplifiers, the bandwidth and settling time parameters must be carefully chosen when clocking speeds in excess of 1 MHz are contemplated (Fig. 16.14).

A three-pole Chebyshev analog filter [79] using CCD delay lines for MTI cancellation has been implemented. The advantage of the Chebyshev filter over the simple canceller is that the frequency response is better matched to real clutter which has a finite spectral width, and also shows a more uniform passband response. Figure 16.15 shows the practical circuit of a 3-pole Chebyshev filter, with the signal weightings fed into the summing amplifiers (15-MHz bandwidth, IC operational amplifiers) indicated. The CCDs used were 8-element, 3-phase surface *p*-channel devices operated at a clock frequency of 330 kHz [80]. The mode of clocking used was interrupted

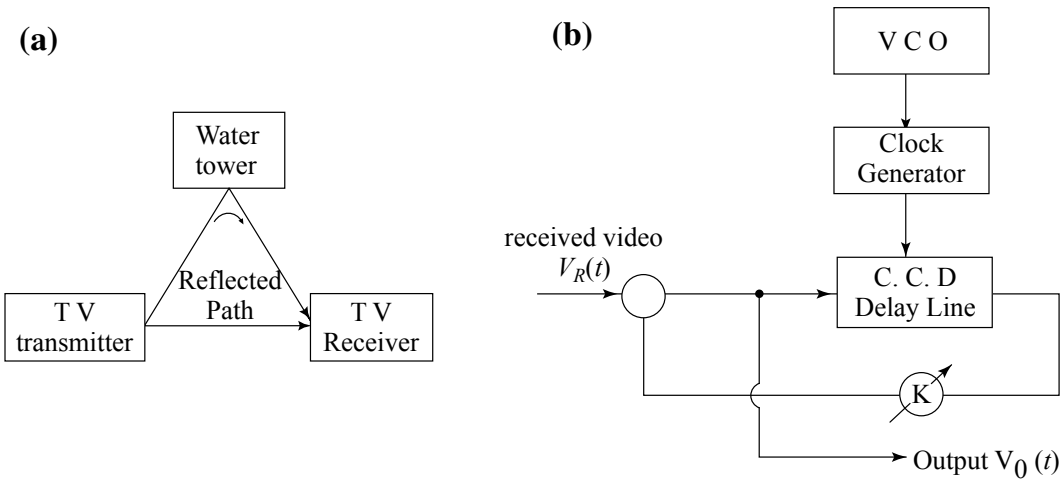


Fig. 16.15 **a** Depicting the reflection problem. **b** Block diagram of ghost-suppressor circuit

clocking, in which the clocks are stopped for a rest period t_r during which ϕ_2, one of the phases remains on. So the analog information in the CCD remains stored under the ϕ_3 electrode. In this case, T is given by

$$T = 1/f_r = t_x + N/f_c \tag{16.28}$$

where f_c is the clocking frequency. If we are operating with t_r= 2 μs, T will be 26 μs and f_r will be 38.4 kHz.

The response with the above filter parameters showed the occurrence of minima at frequencies 0, 38.4 and 76.8 kHz (i.e. PRF of MTI Radar is 38.4 kHz).

16.7.3 Processor for Pulse-Doppler Radar

CCD and surface acoustic wave (SAW) techniques [81] have been combined to advantage for processing coherent radar signals, in order to extract target range and velocity informations. This will replace the complex and costly bank of Doppler filters for each range gate, which might otherwise be needed to extract the information.

The essential functions of the processor are to segregate echoes from different ranges and to separately detect the presence of the resolvable frequency components present in each range cell. The spectral resolution required might typically be 100 Hz, so that a processor capable of storing data over a period of milliseconds is needed, with the facility of sorting it into 100 or more (range) channels for spectral analysis in real time. CCDs now achieve the bandwidths needed for radar signals and can store sampled analog data for tens of milliseconds. The external control of time scale by means of the clock rate means that the same CCD data store can be used with radars of different bandwidth. The flexibility of timing also allows the data input and output to be at different rates, which means that the duration of the stored data can be compressed by a large factor.

The spectrum analyzer for this can conveniently be implemented using SAW 'chirp' filters, the inflexiblity of the filter parameters being unimportant, because processing time and bandwidth can be matched to the requirement by the CCD store.

16.7.4 Charge-Transfer Memory for ECM

The objective of the Radar ECM memory system is to generate synthetic or false targets. This is achieved by acquiring the Radar pulse of interest, storing it for a maximum of one or more PRIs, and recalling and retransmitting it in command at any time within the storage period. But the requirement of transfer efficiency of the CCD used for ECM is higher than MTI requirement due to the demand of reasonably large time delay-bandwidth products.

16.7.5 Television Ghost Suppression Using CCD

TV ghosts are commonly generated by reflections from objects such as water towers, buildings etc. A system was experimented upon (Butler [78]) at General Electric Company for suppression of ghost images in TV employing charge-transfer devices. Geometry of a simple single-reflection problem is illustrated in Fig. 16.15a. The net signal received at the receiver is the sum of desired signal collected over a direct path and a reflected signal, that is attenuated and delayed with respect to the desired signal.

The amplitude of the detected video signal can be approximated by

$$V_R(t) = V(t) + \alpha V(t - \tau) \cos \omega_c \tau \tag{16.29}$$

where α is the attenuation characteristic of the reflected path, ω_c is the RF carrier frequency, and τ is the differential delay between paths. The $\cos \omega_c \tau$ term varies with τ, changing from positive to negative, as the relative propagation path lengths of the desired and ghost signals change by $\lambda_c/2$ (λ_c = carrier wavelength). The ghost signal can therefore be negative or positive, depending on the phase relationships of the main and the reflected RF carriers, and the ghost suppression circuit must be capable of phase reversal as well as gain control, in order to cancel all ghosts. Also, note that the ghost amplitude and polarity is a function of the relative motion of the transmitting and receiving antennas and of the ghost reflecting surface. This is especially so at UHF, where the wavelength is very short (34–64 cm). Automatic adjustment of a ghost suppression circuit would be a highly desirable feature at these frequencies during a windstorm, for example.

The basic principle of the ghost-suppressor circuit is to use the desired signal to synthesize a ghost of the correct magnitude polarity and differential delay so that when the 'synthetic' ghost is summed with the required signal, the synthetic ghost cancels the real ghost and yields a ghost-free picture. A block diagram of a circuit that accomplishes this is shown in Fig. 16.15b. The received signal is fed to a summing node, where it is differentially summed with the weighted output of the CCD analog delay line. The output of the summing node is fed to the TV receiver and entered into the delay line. The parameters of the circuit, i.e. K and throughput-time delay of the delay line T_d are adjusted so that $T_d = \tau$ and $K = \alpha \cos \omega_c \tau$. The adjustment is accomplished manually, while viewing the TV picture. Control of T_d is achieved by varying the clock frequency K is simply a gain variable. The extent to which one can achieve complete cancellation of the ghost is thus dependent only on the capability of the CCD to accommodate the full bandwidth video signal.

In a serial CCD structure capable of storing n signal samples, the clock frequency f_c is equal to the sampling frequency f_s and $T_d = n/f_c$. Since the maximum bandwidth of a sampled data device is $f_s/2$, it follows that the maximum ωT_d (ω= bandwidth) product of the CCD is $n/2$. In practice, ωT_d will be less than $n/2$ because of incomplete charge transfer. For TV video signal processing applications, bandwidths of the order of 4–5 MHz are required, and depending on this specific application delays, ranging

from a few μs up to a full TV frame of delay (33 ms) and more, may be necessary. Thus ωT_d products, from as low as 10 s to as high as several 100 s or 1000 s may be required. Charge-transfer inefficiency, which ultimately limits the maximum ωT_d product of a CCD is therefore one of the most important operating parameters of these devices. Experimental results from high-speed bulk channel CCDs show ωT_d products in excess of 100. The most objectionable ghosts are often those with short time delays (<10 μs); charge-transfer technology thus appears to be well suited to this application.

16.8 Concluding Remarks

It is usual in a concluding section to refer to prospects for the future. However, such comments have been well documented elsewhere, for example, in some papers [82–84]. We would rather conclude this paper by describing some of the work on CCDs that were undertaken at lIT Delhi.

The two major efforts at lIT Delhi in CCDs have been in device characterization and fabrication and in circuit techniques for compensating ∈. In device fabrication, two CCD structures were successfully made. One is a 3-bit 3-phase CCD, which achieves small inter-electrode spacings (less than 1 micron) by a process known as shadow masking [71, 85]. These small inter-electrodes gaps, which cannot be achieved by using standard photolithography, are necessary for efficient transfer of charge from one electrode to the next.

This CCD can be clocked at frequencies from 500 Hz to 50 kHz, thereby providing an electronically variable analog delay of 60 μs–6 ms. The upper frequency range is limited by the large (60μ) electrode length of this experimental CCD. Within this frequency range, the inefficiency ∈ is less than 10^{-2}. It also has output taps at each bit for possible use in compensating circuits and transversal filters.

The other CCD structure which has been fabricated is a novel 2-phase 4-bit double-dielectric device which achieves narrow inter-electrode spacings by the anodic oxidation of aluminium to provide a thin separation of Al_2O_3 between electrodes. This has the advantage over the shadow masked CCD in that it is more stable and reliable. It can be clocked from 1 to 60 kHz with an ∈ less than 10^{-2}. This CCD includes on-clip amplification in the form of reset and output transistors.

Work was also undertaken at IIT Delhi to grow oxides which gave rise to large storage times, or in other words, low dark currents. This is specifically useful in CCDs to be used as low-frequency signal processing devices and imaging devices. Using controlled amounts of trichloroethylene in the oxygen, used to grow the oxides, has resulted in devices with storage times as high as 500 s, so that very low-frequency operation of such devices is possible.

The other major thrust at IIT had been the development of special techniques and circuits for the compensation of ∈ in lossy CCDs. Several techniques for the compensation of transfer inefficiency had been developed. A method using external feedback was proposed for exact realization of any given transfer function. An 18-

stage CTD delay using both BBD and CCD stages had been compensated with suitable modifications taking practical limitations into account. Several FIR and IIR filter structures were developed which compensate for the effect of $\in$ to a good approximation, when connected in series with the device.

References

1. J.M.L. Janssen, Discontinuous low-frequency delay line with continuously-variable delay. Nature **169**, 148–149 (1952)
2. G. Kraus, Analog Speicherketle: Eine neuartige Schaltung zum Speichem und Verzoegern von Signalen. Electron. Lett. **3**, 544–546 (1967)
3. F.L.J. Sangster, The bucket brigade delay line, a shift register for analog signals. Philips Tech. Rev. **31**, 97–110 (1970)
4. W.S. Boyle, G.E. Smith, Charge-coupled semiconductor devices. Bell Syst. Tech. J. **49**, 587–593 (1970)
5. M.F. Tompsett, G.F. Amelio, G.E. Smith, Charge coupled 8-bit shift register. Appl. Phys. Lett. **17**, 111–115 (1970)
6. M.F. Tompsett, W.J. Bertram, D.A. Sealer, C.H. Sequin, Charge-coupling improves its image, challenging video camera tubes. Electronics **46**, 162–168 (1973)
7. M.F. Tompsett, G.F. Amelio, W.J. Bertram, R.R. Buckley, W.J. Mc Namara, J.C. Mickelsen, D.A. Sealer, Charge-coupled imaging devices: experimental results. IEEE Trans. ED-**18**, 992–996 (1971)
8. L.J.M. Esser, The peristaltic charge-coupled device for high speed charge transfer. ISSCC Dig. Tech. Papers 28–29 (1974)
9. M.P. Singh, D.R. Lamb, P.C.T. Roberts, Effect of signal and fat zero size on the performance of 3-phase CCD shift registers. Proc. IEE. **122**, 693–696 (1975)
10. C.K. Kim, M. Lenzlinger, Charge-transfer in charge-coupled devices. J. Appl. Phys. **42**, 3586–3594 (1971)
11. J.E. Carnes, W.F. Kosonocky, Fast interface state losses in charge-coupled devices. Appl. Phys. Lett. **20**, 261–263 (1972)
12. M.F. Tompsett, The quantitative effects of interface states on the performance of charge-coupled devices. IEEE Trans. ED-**20**, 45–55 (1973)
13. J.E. Carnes, W.F. Kosonocky, E.G. Ramberg, Drift-aiding fringing fields in charge-coupled devices. IEEE J. Solid State Circuits SC-**6**, 322–326 (1971)
14. R.J. Strain, N.L. Shryer, A nonlinear diffusion analysis of charge-coupled device transfer. Bell Syst. Tech. J. **50**, 1721–1740 (1971)
15. L.G. Heller, W.H. Chang, A.W. Lo, A model of charge transfer in bucket-brigade and charge-coupled devices. IBM J. Res. Dev. **16**, 184–187 (1972)
16. H.S. Lee, L.G. Heller, Charge-control method of charge-coupled device-transfer analysis. IEEE Trans. ED-**19**, 1270–1279 (1972)
17. W.F. Kosononocky, J.E. Carses, Two-phase charge coupled devices with overlapping polysilicon and aluminium gates. RCA Rev. **34**, 164–202 (1973)
18. M.A. White, D.R. Lampe, F.C. Blaha, I.A. Mack, Characterisation of surface channel *CCD* image arrays at low light levels. IEEE J. Solid-State Circuits SC-**9**, 1-D (See also Ref. 18 for correlated double sampling.) (1974)
19. R.W. Broderson, S.P. Emmons, Noise in buried channel charge-coupled devices. IEEE J. Solid State Circuits SC-**11**, 147–155 (1976)
20. D.D. Wen, Design and operation of a floating-gate amplifier. IEEE J. Solid State Circuits SC-**9**, 410–414 (1974)
21. G.F. Vanstosf, J.B.G. Roberts, A.E. Long, The measurement of the charge residual for C. C. D. transfer using frequency and impulse responses. Solid-State Electron. **17**, 889–895 (1974)

22. M.F. Tompsett, B.B. Kosicki, D. Kahg, Measurements of transfer inefficiency of 250-element undercut isolated charge-coupled devices. Bell Syst. Tech. J. **52**, 1–7 (1973)
23. R.W. Broderson, D.D. Buss, A.F. Tasch Jr., Experimental characterisation of transfer efficiency in charge-coupled devices. IEEE Trans. ED-**22**, 40–46 (1975)
24. P.A. Levine, Measurement of transfer efficiency by use of feedback to increase the effective number of transfers. IEEE J. Solid-State Circuits SC-**8**, 104–108 (1973)
25. S.C. Dutta Roy, V.G. Das, Pre-distortion technique for the measurement of transfer inefficiency of charge transfer devices. Proc. IEEE **66**, 601–602 (1978)
26. J.M. Caywood, D.D. Buss, Frequency response of a multiplexed charge-transfer delay line. IEEE J. Solid-State Circuits SC-**9**, 310–311 (1974)
27. S.C. Dutta Roy, On the frequency response of a multiplexed charge-transfer delay line. Electron. Lett. **12**(7), 280–281 (1976). Erratum in **12**, 391 (1976)
28. R.C. Tozer, G.S. Honson, Reduction of high-level nonlinear smearing in C.C.D.s. Electron. Lett. **12**, 355–356 (1976)
29. D.C. Cooper, E.H. Darlington, S.M. Pettord, J.B.G. Roberts, Reducing the effect of charge-transfer inefficiency in a C.C.D. video integrator. Electron. Lett. **11**, 384–385 (1975)
30. J. Mavor, M.C. Davie, P.B. Denyer, Techniques for increasing the effective charge-transfer efficiency of tapped C. C. D. registers. Electron. Lett. **13**, 31–33 (1977)
31. L. Boonstra, F.I.J. Sangster, Progress on bucket brigade charge-transfer devices. ISSCC Dig. Tech. Papers **15**, 140–144 (1972)
32. K.K. Thorbner, Optimum linear filtering for charge transfer devices. IEEE J. Solid-State Circuits SC-**9**, 285–291 (1974)
33. S.C. Dutta Roy, R.K. Arora, Exact circuit design using inefficient C.T.D.s. Electron. Lett. **12**, 18–19 (1976)
34. S.C. Durra Roy, V.G. Das, On exact compensation of transfer inefficiency in a charge-transfer delay line. Electron. Lett. **14**, 115–116 (1978)
35. H.E. Kallman, Transversal filters. Proc. IRE **7**, 302–310 (1940)
36. D.D. Buss, D.R. Collins, W.H. Bailey, C.R. Reeves, Transversal filtering using charge-transfer devices. IEEE J. Solid-State Circuits SC-**8**, 138–146 (1973)
37. R.D. Baertsch, W.E. Engeler, H.S. Goldberg, C.M. Puckbite, J.J. Tiemnn, The design and operation of practical charge-transfer transversal filters. IEEE Trans. ED-**23**, 133–141 (1976)
38. A.A. Ibrahim, G.J. Hupe, L.P. Sellars, Multiple filter characteristics using a single CCD structure, in *Proceedings of the 1975 International Conference on the Applications of Charge Coupled Devices*, October (1975), pp. 245–249
39. D.D. Buss, W.H. Bailey, J.D. Holmes, L.R. Hitb, Charge transfer device transversal filters for communication systems. Microelectronics **7**, 46–53 (1975)
40. D.D. Buss, R.W. Brodbrsen, C.R. Hewes, A.F. Tasch Jr., Communication applications of CCD transversal filters. IEEE Natl. Telecom. Conf. Rec. **1**, 1–5 (1975)
41. D.D. Buss, C.R. Reeves, W.H. Bailey, D.R. Collins, Charge-transfer devices in frequency filtering, in *Proceedings of the 26th Annual Symposium on Frequency Control*, June (1972), pp. 171–179
42. D.D. Buss, W.H. Bailey, Applications of charge transfer devices to communications, in *Proceedings of the CCD Applications Conference*, San Diego, September (1973), pp. 83–93
43. G.P. Weckler, A tapped analog-delay for sampled-data signal processing-PETICON Application Note No. 105 (1975)
44. R.W. Brodersen, C.R. Hewes, D.D. Buss, Spectral filtering and Fourier analysis using CCD's, in *Presented at IEEE International Symposium on Circuits and Systems*, Massachusetts, April (1975)
45. A.A. Ibrahim, L. Sellars, CCD's for transversal filter applications, in *Technical Digest on IEEE International Electron Devices Meeting*, Washington, December (1974), pp. 240–243
46. T. Foxall, A.A. Ibrahim, G.J. Hupe, Double-split electrode transversal filter. Electron. Lett. **13**, 323–324 (1977)
47. P.B. Denyer, J. Mavor, Design of CCD delay lines with floating-gate taps. IEEE J. Solid-State Circuit **1**, 121–129 (1977)

48. J. Mavor, P.B. Denyer, Design and development of CCD programmable transversal filters. IEE J. Electron. Circuits Syst. **2**, 1–8 (1978)
49. A.O. Oppenhcim, R.W. Schafer, *Digital Signal Processing* (Prentice Hall, Inc., Englewood Cliffs, NJ, 1975)
50. C.R. Hewes, R.W. Brodbrsen, D.D. Buss, Frequency filtering using change-coupled devices, in *Proceedings of 29th Annual Frequency Control Symposium*, May (1975)
51. R.W. Brodersen, C.R. Hewes, D.D. Buss, A 500-stage CCD transversal filter for spectral analysis. IEEE J. Solid-State Circuits SC-**11**, 75–83 (1976)
52. D.R. Collins, W.H. Bailey, W.M. Gosney, D.D. Buss, Charge-coupled device analog matched filters. Electron. Lett. **8**, 328–329 (1972)
53. A. Chowaniec, G.S. Hobson, A wide-band quadrature phasing system using charge-transfer devices. Solid-State Electron. **19**, 201–207 (1976)
54. R.D. Melen, J.D. Schott, J.T. Walker, J.D. Meindl, CCD dynamically focussed lenses for ultrasonic imaging systems, in *Proceedings of CCD* (1975), pp. 165–171
55. J. Carver, G.S. Hobson, Interference rejection filter using charge-coupled devices. Electron. Lett. **13**, 732–733 (1977)
56. J.W. Arthur, J. Mavor, C.F.N. Cowan, Novel linear regression analyser using analog CCD. Electron. Lett. **13**, 751–753 (1977)
57. T.P. Pulford, A.H. James, Radar display system using CCD for time expansion. Electron. Lett. **13**, 700–701 (1977)
58. F.L.J. Sangster, K. Teer, Bucket-brigade electronics-new possibilities for delay, time-axis conversion and scanning. IEEE J. Solid-State Circuits SC-**4**, 131–136 (1969)
59. T.A. Zimmerman, C.S. Miller, The application of charge-coupled devices to digital-signal processing, in *International Communications Conference*, San Franscisco, June (1975)
60. D.A. Scaler, C.H. Sequin, A.M. Mohsen, M.F. Tompsett, Design and characterisation of charge coupled devices for analog signal processing, in *International Communications Conference*, San Franscisco, June (1975)
61. A Programmable Binary-Analog Correlator-Reticon Application Note No. 106
62. G.S. Hobson, Charge-coupled devices. Proc. IEE **124**, 925–945 (1977)
63. D.D. Buss, W.H. Bailey, A.F. Tash Jr., Signal processing applications of charge-coupled devices, in *Proceedings of CCD 1974 Conference* (University of Edinburgh, 1974)
64. D.D. Buss, W.H. Bailey, D.R. Collins, Analysis and applications of analog CCD circuits, in *Proceedings of the International Symposium on Circuit Theory*, Toronto, April (1973), pp. 3–7
65. D.D. Buss, W.H. Bailey, R.W. Broderson, C.R. Hewes, Signal processing applications of charge coupled devices. 1974 WESCON Professional Program, Session 2 (1974)
66. J.A. Sekula, P.R. Prince, C.S. Wank, Non-recursive matched filters using charge-coupled devices, in *Technical Digest on IEEE International Electron Devices Meeting*, Washington, December (1974), pp. 244–247
67. D.D. Buss, W.H. Bailey, L.R. Hite, Spread-spectrum communications using charge transfer devices, in *Proceedings of 1973 Symposium on Spread Spectrum Communications*, San Diego, California, March (1973), pp. 83–92
68. J.E. Cooley, J.W. Tukey, An algorithm for the machine computation of complex fourier series. Math. Comput. **19**, 297–301 (1965)
69. L.I. Bluestein, A linear filtering approach to the computation of the discrete fourier transform. NEREM Rec. 218–219 (1968)
70. B. Wardrop, E. Bull, A discrete fourier transform processor using charge-coupled devices. Marconi Rev. **40**, 204 (First Quarter) (1977)
71. IEEE J. Solid-State Circuits, Special Issue on Analog Circuits, SC-12 (1977)
72. M.A. Jack, D.G. Park, P.M. Grant, CCD spectrum analyser using prime transform algorithm. Electron. Lett. **13**, 431–432 (1977)
73. L.R. Rabiner, R.W. Schaffer, C.M. Radar, The chirp-Z-transform algorithm and its application. Bell Syst. Tech. J. **48**, 1249–1292 (1969)
74. D.D. Buss, R.L. Yeenkant, R.W. Brodersen, C.R. Hewes, Comparison between the CCD CZT and the digital FFT, in *Proceedings of CCD, 1975*, San Diego, October (1975), pp. 267–281

75. J.S. Seeks, J.D. Jackson, G.R. Adams, Spectrum analysis with a CCD delay line. Electron. Lett. **13**, 738–739 (1977)
76. J.H. Mcclellan, T.W. Parks, L.R. Rabiner, A computer program for designing optimum FIR linear phase digital filters. IEEE Trans. AU-**21**, 506–526 (1973)
77. W.J. Butler, W.E. Engeler, H.S. Goldberg, C.M. Pucketie, H. Lobenstein, Charge transfer analog memories for radar and ECM systems. IEEE Trans. ED-**23**, 161–168 (1976)
78. W.J. Butler, C.M. Puckett, N.C. Giltinger, An experimental TV ghost suppressor circuit using charge transfer devices. IEEE J. Solid-State Circuits SC-**10**, 247–249 (1975)
79. J.B.G. Roberts, E. James, K.A. Roche, Moving target indicator recursive radar filter using bucket brigade devices. Electron. Lett. **9**, 89–90 (1973)
80. J.E. Bounden, M.J. Tomlinson, CCD Chebyshev filter for radar MTI applications. Electron. Lett. **10**, 89–90 (1974)
81. J.B.G. Roberts, R. James, D.V. Mccaughan, R.F. Simons, A processor for pulse-doppler radar. IEEE J. Solid-State Circuits SC-**11**, 100–104 (1976)
82. D.D. Buss, W.H. Bailey, M.M. Whatley, R.G. Brodersen, Application of CCDs to radar signal processing, in *Northeast Electronics Research and Engineering Meeting*, Part 4 (1974), pp. 83–98
83. B. Wardrop, E. Bull, Application of CCD to MTI filters. Marconi Rev. (Fourth Quarter) (1976)
84. D.F. Barbe, W.D. Baker, K.L. Davis, Signal processing with charge-coupled devices. IEEE J. Solid-State Circuits SC-**13**, 34–51 (1978)
85. R.A. Haken, J.D.E. Beynon, I.M. Baker, P.C.T. Roberts, Charge-coupled structures with self-aligned sub-micron gaps. IEEE Trans. ED-**22**, 289–293 (1975)

Chapter 17
Linear-Phase Versus Minimum-Phase CCD Transversal Filters

Abstract It is shown that the effect of charge-transfer inefficiency on the frequency response of CCD transversal filters is a sensitive function of the actual zero locations and in contrast to the observation made in the paper by Foxall et al. (1977), the minimum phase design is not always less sensitive than its linear-phase prototype.

Keywords Linear phase · Minimum phase · Transversal filters

17.1 Introduction

In order to avoid the long time delays associated with the linear-phase CCD transversal filters, Foxall, Ibrahim and Hupe investigated, in [1], an example of the minimum-phase design. Besides reducing the group delay, the latter design was observed to be less sensitive to transfer inefficiency and tap-weight error. In this chapter, a detailed analysis is carried out of the sensitivity of general linear and minimum-phase designs to transfer inefficiency and it is shown that the observation made by Foxall et al. [1] is not always valid. The conditions for the superiority of minimum-phase design are given, along with examples in which the linear-phase design shows less sensitivity to transfer inefficiency.

17.2 Analysis

Consider a linear-phase CCD FIR filter having the quad of zeros at $z = re^{\pm j\theta}$ and $(1/r)re^{\pm j\theta}$, where $r > 1$. The transfer function will be given by

$$H_{I_o}(Z) = \left(Z^{-1} - \frac{1}{r}e^{j\theta}\right)\left(Z^{-1} - \frac{1}{r}e^{-j\theta}\right)(Z^{-1} - re^{j\theta})(Z^{-1} - re^{-j\theta}) \quad (17.1)$$

Source: S.C. Dutta Roy and N. Kapur, "Linear-Phase Versus Minimum-Phase CCD Transversal Filters," *JIETE*, vol 27, pp 622–624, December 1981.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_17

A minimum-phase FIR filter having the same magnitude response as of Eq. (17.1) will have a transfer function

$$H_{mo}(Z) = \left(1 - \frac{1}{r}e^{j\theta}Z^{-1}\right)\left(1 - \frac{1}{r}e^{-j\theta}Z^{-1}\right)(Z^{-1} - re^{j\theta})(Z^{-1} - re^{-j\theta}) \tag{17.2}$$

Let the first factors in Eqs. (17.1) and (17.2) be denoted by $L_o(Z)$ and $M_o(Z)$, respectively. When the charge transfer is not perfect, the term Z^{-1} changes to $Z_\varepsilon^{-1} = (1-\varepsilon)Z^{-1}/(1-\varepsilon Z^{-1})$, where ε denotes the transfer inefficiency. Let

$$L_\varepsilon(Z) \underline{\underline{\Delta}}\, L_O(Z_\varepsilon^{-1}) \text{ and } M_\varepsilon(Z) \underline{\underline{\Delta}}\, M_o(Z_\varepsilon^{-1}) \tag{17.3}$$

The frequency response of these factors can be obtained by putting $Z = e^{j\omega}$, where the clock period has been normalized to unity. After some simplifications, the magnitude responses can be shown to be given by

$$|L_o|^2 = \left[r^2 + 1 - 2r\cos(\omega + \theta)\right]/r^2 \tag{17.4}$$

$$|L_\varepsilon|^2 \cong \left[1 + r^2(l - 2\varepsilon) + 2\varepsilon r\cos\theta - 2r(l - \varepsilon)\cos(\omega + \theta) - 2\varepsilon\cos\omega\right]/\left[r^2(l - 2\varepsilon\cos\omega)\right] \tag{17.5}$$

$$|M_\varepsilon|^2 \cong \left[r^2 + (1 - 2\varepsilon) + 2\varepsilon r\cos\theta - 2r(1 - \varepsilon)\cos(\omega + \theta) - 2r^2\varepsilon\cos\omega\right]/\left[r^2(1 - 2\varepsilon\cos\omega)\right], \tag{17.6}$$

where, in Eqs. (17.5) and (17.6), terms containing ε^2 and higher powers of ε are neglected. It may be easily verified that Eqs. (17.5) and (17.6) reduce to Eq. (17.4) when $\varepsilon = 0$, as should be the case. Define the fractional errors in the magnitude-squared functions, $|L_\varepsilon|^2$ and $|M_\varepsilon|^2$ due to non-zero ε, as

$$\alpha_1 = \left(|L_\varepsilon|^2 - |L_o|^2\right)/|L_o|^2 \text{ and } \beta_1 = \left(|M_\varepsilon|^2 - |M_o|^2/|M_o|^2\right) \tag{17.7}$$

then it can be shown that

$$\alpha_1 = f(\varepsilon, \omega, r, \theta)\left\{r^2(\cos\omega - 1) + r[\cos(\omega + \theta) - \cos(2\omega + \theta)]\right\} \tag{17.8}$$

$$\beta_1 = f(\varepsilon, \omega, r, \theta)\{(\cos\omega - 1) + r[\cos(\omega + \theta) - \cos(2\omega + \theta)]\} \tag{17.9}$$

where,

$$f(\varepsilon, \omega, r, \theta) = 2\varepsilon/\left\{(1 - 2\varepsilon\cos\omega)\left[r^2 + 1 - 2r\cos(\omega + \theta)\right]\right\} \tag{17.10}$$

Now consider the term $(Z^{-1} - re^{j\theta}) \underline{\underline{\Delta}}\, N_o(Z)$ which is common to Eqs. (17.1) and (17.2). Let

$$N_\varepsilon(Z) = N_o(Z_\varepsilon^{-1}) \tag{17.11}$$

Then proceeding in an exactly similar manner, it can be shown that the fractional error

$$\gamma_1 = \left(|N\varepsilon|^2 - |N_o|^2\right)/|N_o|^2 \tag{17.12}$$

is the same as β_1 given by Eqs. (17.9) and (17.10).

Let the fractional errors due to the second and fourth factors in Eq. (17.1) be denoted by α_2 and γ_2 respectively; these can be obtained from Eqs. (17.8) and (17.9), respectively, along with Eq. (17.10), by replacing θ by $-\theta$. Also let

$$H_{1\varepsilon}(Z) \underline{\underline{\Delta}} H_{1_0}(Z_\varepsilon^{-1}) \tag{17.13}$$

Then it is obvious that

$$\left|H_{1\varepsilon}(e^{j\omega})\right| / \left|H_{1_0}(e^{j\omega})\right| = [(1+\alpha_1)(1+\alpha_2)(1+\gamma_1)(1+\gamma_2)]^{\frac{1}{2}} \cong 1 + (\alpha_1+\alpha_2+\gamma_1+\gamma_2)/2 \tag{17.14}$$

where, the approximate expression results from the fact that $\alpha_{1,2}$ and $\gamma_{1,2}$ are each proportional to ε and their magnitudes are, in practice, small compared to unity. Proceeding in a similar fashion with the minimum-phase transfer function, given by Eq. (17.2), and defining

$$H_{m\varepsilon}(Z) \underline{\underline{\Delta}} H_{mo}(Z_\varepsilon^{-1}) \tag{17.15}$$

it follows that

$$\frac{\left|H_{me}(e^{j\omega})\right|}{\left|H_{mo}(e^{j\omega})\right|} \cong 1 + \frac{1}{2}(\beta_1+\beta_2+\gamma_1+\gamma_2) = 1 + (\gamma_1+\gamma_2) \tag{17.16}$$

In general, let there be n quads of zeros in the linear-phase transfer function. Then it is obvious that the percentage error in the linear-phase magnitude response will be

$$\delta_1 = \frac{1}{2}\sum_{i=1}^{n}(\alpha_{1i}+\alpha_{2i}+\gamma_{1i}+\gamma_{2i}) \times 100\% \tag{17.17}$$

Similarly, the percentage error in the corresponding minimum phase magnitude response will be

$$\delta_m = \sum_{i=1}^{n}(\gamma_{1i}+\gamma_{2i}) \times 100\% \tag{17.18}$$

From Eqs. (17.8), (17.9) and (17.10), it should be apparent that α's and γ's can be positive or negative depending on the values of r, θ and ω. A detailed investigation is

therefore necessary to establish the superiority or *otherwise* of the minimum-phase transfer function over the linear-phase one. From the preceding results, it should be clear that further analytical investigation in this direction would be difficult, if not impossible. Extensive numerical computations were, therefore, undertaken to obtain the nature of dependence of errors on r, θ and ω. These and other results are given in the next section.

17.3 Results of Numerical Computations

Numerical investigations reveal that for a particular quad and fixed value of θ, there exists a critical value of r, above which the magnitude error is less in the minimum phase case, for the entire frequency range $0 \leq \omega \leq \pi$. A plot of this critical value as a function of θ is shown in Fig. 17.1. If r is less than this critical value, the linear phase design is superior to the minimum phase case, but only over a part of the frequency $0 \leq \omega \leq \pi$. In other words, it is not possible to have a linear phase design, which is superior to the minimum-phase case over the entire frequency range of interest. The performance of a general n-quad filter will depend on the contribution due to the individual quads.

Two illustrative examples will now be presented.

Example 1 Consider a multiple band-pass linear phase filter and its corresponding minimum-phase case characterized by the sets[1] of r and θ given in Table 17.1. The variations of percentage error in the magnitude response due to $\theta = 10^{-3}$ are shown

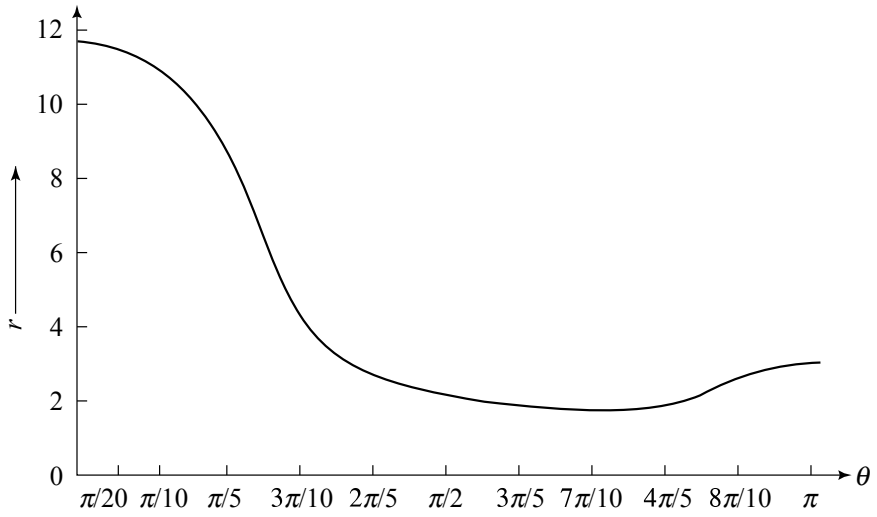


Fig. 17.1 Plot of the critical value of r versus θ. If r is greater than this critical value, the minimum-phase design is superior to the linear-phase design over the entire frequency range $0 \leq \omega \leq \pi$

[1]For a given r ($\neq 1$) and θ ($\neq 0$ or π), the linear-phase case has zeros at $re^{\pm j\theta}$ and $\frac{1}{r}e^{\pm j\theta}$, while the minimum-phase design has double zeros at each of the two positions $\frac{1}{r}e^{\pm j\theta}$. For $r = 1$ and $\theta \neq 0$ or π, each design has two zeros at $e^{\pm j\theta}$. For $r = 1$ and $\theta = 0$ (or π) each design has only one

Table 17.1 Zero locations* for example 1

r	θ	No. of Zeros
1.5833	0	2
1.2	2π/20	4
1.15	3π/20	4
1.15	4π/20	4
1	11π/20	2
1	14 π/20	2
1	17π/20	2
1	π	1

in Fig. 17.2. Clearly, the linear-phase design has less error in the frequency range 0.135 π to 0.395 π.

Example 2 This example concerns a wide-band band-pass filter characterized by the sets[2] of r and θ given in Table 17.2. The variations of percentage error in the magnitude response due to an $\varepsilon = 10^{-3}$ are shown in Fig. 17.3. The linear-phase design has less error in the frequency range 0.1 π to 0.5 π.

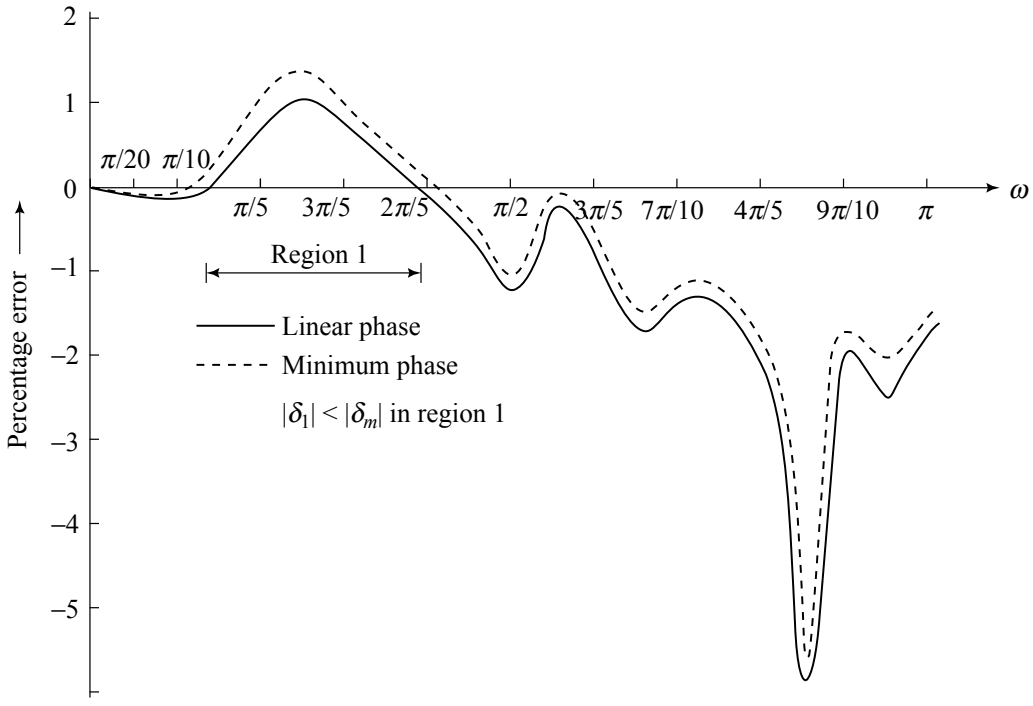


Fig. 17.2 Error plots for the filters of Example 1

zero at +1 (–1). For $r \neq 1$ and $\theta = 0$ the linear-phase design has two zeros at r and $1/r$, while the minimum-phase design has double zeros at $1/r$.

[2]Vide foot-note on the previous page.

Table 17.2 Zero locations* for example 2

r	θ	No. of Zeros
1.2	0	2
1.2	$2\pi/20$	4
1.15	$3\pi/20$	4
1.15	$4\pi/20$	4
1	π	1

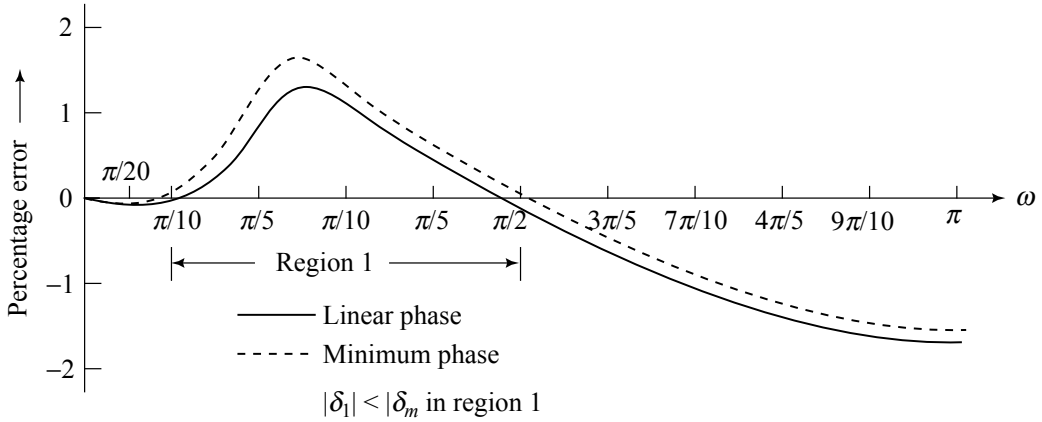


Fig. 17.3 Error plots for the filters of Example 2

17.4 Conclusion

It has been shown by analysis as well as numerical investigations that the minimum-phase CCD transversal filter is not always less sensitive to transfer ineffictency as compared to the linear-phase design. Sufficient conditions for the former to be superior are depicted in Fig. 17.1.

Acknowledgements The authors wish to thank Dr. J. Vasi for helpful discussions.

Reference

1. T.G. Foxall, A.A. Ibrabim, G.J. Hupe, Minimum-Phase CCD Transversal Filters. IEEE J. Solid-State Circuits. SC **12**(6), 638–642 (1977)

Part V
Digital Signal Processing

This part contains 11 chapters. We start with a survey of linear phase digital filters in Chap. 18, and then delve into generalized linear phase conditions in Chap. 19. Chapter 20 reviews the application of Pseudo-Boolean techniques in digital filter design. Maximally linear digital differentiators form the subject of Chap. 21, while Maximally Flat and Monotonic FIR filters using the Bernstein polynomial are discussed in Chap. 22. Design of FIR notch filters form the contents of Chap. 23. Chapter 24 deals with analytic inversion of Van der Monde and confluent Van der Monde matrices. Chapters 25–28 deal with the solution of an outstanding problem in DSP, *viz*. Canonical lattice realization. Chapters 25 and 26 deal with FIR lattice, while Chaps. 27 and 28 deal with IIR lattices.

Chapter 18
A State-of-the-Art Survey on Linear Phase Digital Filter Design

Abstract Following a brief review of digital filters and associated terminology, the problem of linear phase design is stated in precise form. The relative difficulty of IIR design as compared to FIR design is discussed and a state-of-the-art survey is presented for both cases.

Keywords Linear phase · Digital filter

18.1 Introduction

The subject of digital filters can be called a relatively recent development in filter theory and design. The upsurge of interest in digital filters was caused by many factors, the most important of which was the increased use of digital hardware, rather than analog ones, for processing of signals. This, in turn, was stimulated by advances in microelectronic technology, leading to the availability of low cost, high-speed digital circuits. Also, in many applications, the data to be processed is basically discrete, so that digital processing is the most natural thing to do. Digital filtering finds applications in many practical situations like radar and sonar signal processing, speech communication, geophysical data processing.

Digital filtering has many advantages over analog filtering, some of them being (i) accurate, drift-free operation with predictable performance and a high degree of precision, (ii) exact realization of linear phase transmission, (iii) absence of any matching problem, (iv) high degree of component utilization arising due to the use of multiplexing, (v) ability to realize a greater variety of filter functions and (vi) economy at very low frequencies when an aggregate of such filters is needed. On the other hand, a digital filter is more complex than an analog filter and its advantages are felt only for filters with certain minimum critical size.

There are many aspects into which recent developments in digital filters can be categorized. This chapter is specifically concerned with some recent research on linear phase design, with particular reference to the work done at the Indian

Source: S.C. Dutta Roy, "A State-of-the-Art Survey on Linear Phase Digital Filter Design," *JIETE*, vol 22, pp 1–4, January 1976.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_18

Institute of Technology, New Delhi. As a prelude to this presentation, it would be of advantage to review, very briefly, some basic concepts and definitions pertaining to digital filters. First, we note that a digital filter accepts a quantized time series $\{x(kT)\}$ and produces another quantized time series $\{y(kT)\}$ through the solution of a linear difference equation

$$y(kT) = \sum_{i=0}^{m} a_i x(kT - iT) - \sum_{i=1}^{n} b_i y(kT - iT) \tag{18.1}$$

It should be clear from this description that a general-purpose digital computer can be used for digital filtering. Alternatively, a special-purpose hardware, consisting of delays, multipliers and adders, can be built to perform the calculations prescribed by Eq. (18.1). Because of finite length registers used in either implementation, the consideration of multiplication errors or noise, as it is commonly called, becomes very important. If the signal to be processed is analog, we need additional hardware, viz., bandlimiter (analog), sampler and A/D converter. Because of finite bit capacity of A/D converters, an additional component of noise is thereby introduced. If the output signal is desired to be in analog form, we need a D/A converter, followed by a data reconstructor and a simple analog low-pass filter.

The frequency-domain description of the digital filter is obtained through the z-transform, defined by

$$X(z) \triangleq Z\{x(kT)\} = \sum_{k=0}^{\infty} x(kT) z^{-k} \tag{18.2}$$

where the z-plane is related to the s-plane by $z = e^{sT}$. Combining Eqs. (18.1) and (18.2), it is possible to define a z-domain transfer function as

$$H(z) = \frac{Y(z)}{X(z)} = \left(z^{n-m} \sum_{i=0}^{m} a_{m-i} z^i \right) \Bigg/ \left(\sum_{i=0}^{n} b_{n-i} z^i \right) \tag{18.3}$$

where $b_0 = 1$. The frequency response is obtained by putting $z = e^{j\omega T}$; it is clear that $H(e^{j\omega T})$ is periodic in T, with the baseband $-\pi \leq \omega T \leq \pi$. It should also be clear that for stability, poles of $H(z)$ must be located within the unit circle in the z-plane, but there are no such restrictions on zeros of $H(z)$. For real time realization, $H(z)$ must be rational, i.e. $H(z) = N(z)/D(z)$, where degree of $N(z)$ is less than or equal to degree of $D(z)$. If this is not the case, an appropriate number (p) of poles at $z = 0$ has to be introduced; then the magnitude, $M(\omega) = \left| H(e^{j\omega T}) \right|$ remains unaffected while $\phi(\omega) =$ angle of $H(e^{j\omega T})$ is decreased by the amount $p\omega\, T$.

The impulse response of a digital filter is defined as the response $\{y(kT)\}$ to a signal $\{x(kT)\} = \{1, 0, 0, \ldots\}$. Based on the duration of the impulse response, digital filters are classified into two categories ,viz., finite duration impulse response (FIR) and infinite duration impulse response (IIR). The former type is obtained when $H(z)$

has poles at $z = 0$ only; this is commonly known as 'nonrecursive'. If $H(z)$ poles are not so restricted, then we obtain an IIR filter, which is commonly known as 'recursive'. The terms 'recursive' and 'nonrecursive' used in the above sense, are misleading, because they refer to how a filter is realized and not to the duration of the impulse response. As shown by Gold and Jordan [1], FIR filters can also be realized recursively and IIR filters can also be realized nonrecursively.

For a given specification, IIR filters are computationally more efficient and require less storage capacity than FIR filters. On the other hand, as will be discussed shortly, a strictly linear phase can be achieved only by FIR filters.

18.2 The Problem of Linear Phase and FIR Design

For any transmission channel or network, which is required to transmit a band of frequencies, the group delay $\tau(\omega)$ should be invariant with frequency in the passband, i.e. the transfer function should be a linear phase one. It is a fact of life that analog filters cannot realize this exactly; however, various approximations are known, of which the Bessel type approximates constant delay in the maximally flat (MF) sense. In linear phase digital filter design, we aim for a low-pass $M(\omega)$, because, as will be seen shortly, by multiplying the resulting transfer function by a suitable polynomial, all other types of $M(\omega)$ can be approximated, without affecting the linear phase characteristic or its approximation. We can, therefore, state the problem as follows: to find a transfer function $H(z)$ such that for $z = e^{j\omega T}$,

$$M(\omega) = \begin{cases} 1 \text{ for } 0 \le \omega T \le \omega_c T < \pi \\ 0 \text{ for } \omega_c T < \omega T \le \pi \end{cases} \tag{18.4}$$

and

$$\tau(\omega) = -\frac{d\phi(\omega)}{d\omega} = \begin{cases} \tau_0 \text{ for } 0 \le \omega T \le \omega_0 T \\ \text{arbitrary for } \omega_c T < \omega T \le \pi \end{cases} \tag{18.5}$$

where τ_0 is a constant.

We first consider the FIR case. Consider the following transfer function[1]

$$H(z) = \left\{ \begin{array}{l} (z-1)^{n_1}(z+1)^{n_2} \prod_{i=1}^{n_3} \left[(z-q_i)(z-q_i^{-1})\right] \prod_{i=1}^{n_4} [(z+1)(z+u_i^{-1})] \cdot \prod_{i=1}^{n_5} [(z-e^{j\theta_i})(z-e^{-j\theta_i})] \\ \prod_{i=1}^{n_6} [(z-r_i e^{j\psi_i})(z-r_i e^{-j\psi i}) \cdot (z-r_i^{-1}e^{j\psi i})(z-r_i^{-1}e^{-j\psi_i})] \end{array} \right\} / z^n$$
$$= N(z)/D(z) \tag{18.6}$$

[1] This is a more generalized form of the transfer function considered by Nowak [2] in that the zeros on the positive as well as the negative real axis and those on the z-plane unit circle are shown explicitly.

where $n = n_1 + n_2 + 2n_3 + 2n_4 + 2n_5 + 4n_6$; all the constants are positive; and further q_i, u_i and r_i are not allowed to have a value of unity. For $z = e^{j\omega T}$, is easy to show that

$$\phi(\omega) = -n\omega T/2 + \pi[(\text{number of unit circle zeros between } \omega T = 0 \text{ and } \omega T) + (n_1/2) + n_3] \tag{18.7}$$

Thus, if there are no unit circle zeros in the range $0 \leq \omega T \leq \omega_0 T$, the phase response is strictly linear with no phase jumps.

In Eq. (18.6), the term $(z+1)^{n_2}$ and each term with the square brackets have the property that if the order of the polynomial is p, then the coefficients of z^i and z^{p-i} are equal. Such polynomials are called mirror image polynomials (MIP) [3]. If the coefficients of z^i and z^{p-i} are equal in magnitude but opposite in sign, the resulting polynomial is called a negative MIP (NMIP) [2]. The term $(z-1)^{n_1}$ is an MIP if n_1 is even and an NMIP if n_1 is odd. The product of MIPs is an MIP while the product of MIPs and NMIP is an NMIP; the $N(z)$ in Eq. (18.6) is either an MIP or $(z-1)$ time an MIP. For $z = e^{j\omega T}$, $M(\omega)$ will obviously be a Fourier series in cos (ωT) or sin (ωT) depending on whether $N(z)$ is MIP or NMIP. As an example when n_1 and n are even,

$$M(\omega) = \sum_{i=0}^{n} d_i \cos(i\omega T) \tag{18.8}$$

The problem of FIR linear phase design therefore effectively reduces to that of choosing d_i, $i = 0 \rightarrow$ such that Eq. (18.8) approximates Eq. (18.4) in a desired fashion. If $M(\omega)$ is constrained to be MF in the passband, it has been shown in [4] that $M(\omega)$ can be expressed in the following elegant form:

$$M(\omega) = (1-\beta)^m \sum_{i=0}^{n-m} {}^{m+i-1}C_i \beta^i \tag{18.9}$$

where

$$\beta = \sin^2(\omega T/2)$$

and m is chosen to satisfy a certain cutoff point at $\beta = \lambda$ according to the relation

$$m = n - (n\lambda + 1/2) \tag{18.11}$$

Except for this special case, analytical solution cannot, in general, be contained for d_i's in Eq. (18.8) to satisfy a prescribed $M(\omega)$. Iterative techniques, however, can be applied to obtained optimum results. Several such techniques have been proposed, of which the windowing, frequency sampling and equiripple designs are the most popular ones. These have been well documented in the literature [5].

18.3 Linear Phase IIR Design

Owing to stability constraints, IIR filters cannot achieve a strictly linear phase. However, in situations where the high degree of FIR filters is not acceptable, one is led to approximate a linear phase by an IIR filter. The most common method of IIR digital filter design is based on the design of the corresponding analog filter, followed by a suitable transformation from the s- to the z-plane. Examples are the impulse invariance, bilinear transformation and the matched z-transform techniques. Such methods provide good results for a filter having piecewise constant frequency response but fail to give satisfactory results for linear magnitude or phase characteristics. For example, bilinear transformation of the analog Bessel filter yields a digital filter which does not have an MF delay characteristic. The synthesis of linear phase IIR filters, therefore, necessitates direct approximation.

For MF delay, Thiran [6] showed that the required transfer function is of the form

$$H(z) = \frac{(2n)!/\left[n!\sum_{i=n+1}^{2n}(2\tau_0+i)\right]}{\sum_{k=0}^{n}(-1)^k \prod_{i=0}^{n}\{2\tau_0+i)/(2\tau_0+k+i)\}z^{-k}} \tag{18.12}$$

where the denominator can be shown to be the Gauss's hypergeometric function $F(-n, 2\tau_0; 2\tau_0 + n + 1; z^{-1})$, and that Eq. (18.12) is always stable. No such elegant solution could, however, be found for equal ripple delay. Starting with the transfer function

$$H(z) = K/D(z^{-1}) = K/\sum_{i=1}^{n}(z^{-1} - z_i^{-1}) \tag{18.13}$$

and putting equal ripple constraints on the group delay, Thiran [7] obtained a set of n simultaneous nonlinear algebraic equations for the pole locations. Explicit solutions can be obtained for the trivial case of $n = 1$ and for the nontrivial case $\omega_0 T = \pi$; for the latter case,

$$D(z^{-1}) = z^{-n} + (-1)^n \in / \tau_0 \tag{18.14}$$

where $\in$ is the ripple amplitude in delay; and, $\in$ and τ_0 are restricted to certain combinations for a stable $H(z)$. Note that for this special case, the poles are uniformly distributed over a circle of radius less than unity. Starting from this explicit solution, a generalized Newton–Raphson procedure was adopted by Thiran for solving the nonlinear equations for an arbitrary $\omega_c T < \pi$.

A conceptually simpler solution for equiripple delay was proposed by Deczky [8]. The method is based on the observation that the group delay of the transfer function

$$H_i(z) = K_i z/\{z - r_i \exp(j\theta_i)\} \tag{18.15}$$

when plotted versus $(\omega T - \theta_i)/\pi$, has a resonance type characteristics, with the sharpness of the peak depending only on r_i while its position is a function of θ_i. Clearly, the superposition of a number of such responses with appropriate r_i's and θ_i's, subject to complex poles occurring in conjugate pairs, should achieve an equal ripple delay. This ultimately leads again to n number of nonlinear equations, with the same explicit solution for the special case $\omega_c T = \pi$. With this as the initial approximation, an iterative solution can be obtained for $\omega_c T < \pi$ using numerical techniques.

Compared to the equiripple designs of Thiran [7] and Deczky [8], a remarkably simple design has been obtained by Hazra and Dutta Roy [9]. The method is based on the observation that the transfer function

$$H_0(z) = 1/(r^m + z^m) \tag{18.16}$$

having poles at $z_k = r\, e^{j\theta_k}$; $\theta_k = (2k+1)\pi/m$; $k = 0 \rightarrow m-1$, is stable when $r < 1$ and has a phase angle for $z = e^{j\omega T}$, given by

$$\begin{aligned} \phi_0(\omega) &= -m\omega T + \arctan\{[r^m \sin(m\omega T)]/[1 + r^m \cos(m\omega T)]\} \\ &\cong -m\omega T + r^m \sin(m\omega T) \end{aligned} \tag{18.17}$$

The approximation in Eq. (18.17) is valid for large m. $H_0(z)$, therefore, approximates a linear phase with a maximum periodic error of r^m radians, which can be made as small as desired by choosing r and m suitably. The magnitude of $H_0(e^{j\omega T})$, viz.

$$\begin{aligned} M_0(\omega) &= [1 + r^{2m} + 2r^m \cos(m\omega T)]^{-1/2} \\ &\cong 1 - r^m \cos(m\omega T) \end{aligned} \tag{18.18}$$

has an equiripple all-pass characteristic. To make the filter a low-pass one, the poles of $H_0(z)$ in the stopband are cancelled by introducing zeros at these points. The resulting transfer function, $H_0(z)\, U(z)$ say, will no longer be linear phase; to retain the linear phase approximation, an equal number of zeros are introduced on a circle of radius $1/r$ at the same angles as those of the zeros of $U(z)$. The new filter function, $H(z) = H_0(z)\, U(z)\, V(z)$ say, will now be linear phase, because $U(z)\, V(z)$ is an MIP having a strictly linear phase characteristics. To be specific, let m be even and let $2n < m$ be the number of passband poles. Then $U(z)\, V(z)$ is an MIP of order $2(m-2n)$, and the angle of $H(e^{j\omega T})$ becomes

$$\phi(\omega) = -2n\omega T + r^m \sin(m\omega T) \tag{18.19}$$

Finally, for real time realization and minimum delay, the numerator and denominator degrees of the transfer function are equalized, resulting in the new transfer function

$$H(z) = K\, H(z)/z^{m-4n} \tag{18.20}$$

The phase angle of $H(e^{j\omega T})$ is

$$\phi(\omega) = -(m - 2n)\omega T + r^m \sin(m\omega T) \tag{18.21}$$

Investigation of the magnitude response of $H(e^{j\omega T})$ shows that for $\theta_k \ll \omega T$ and $(1 - r)^2/r \ll [(\omega T - \theta_k)/\pi/]^2$,

$$M(\omega) \leq \text{constant}/\omega^{4n} \tag{18.22}$$

which leads to the conclusion that the asymptotic decibel loss is upper bounded by $24n$ dB/octave. Thus, in addition to linear phase approximation, the method also provides a certain degree of freedom in shaping the magnitude characteristic. For a prescribed delay τ_0, a maximum delay error $\in$, and an asymptotic slope S in dB/octave, the filter can be designed through the following simple relations

$$n \geq S/24;\ m = 2n + \tau_0/T \text{ and } r = [\in /(2n + \tau_0/T)]^2/(2n + r_0/T) \tag{18.23}$$

18.4 Conclusion

Explicit solutions to the problem of linear phase digital filter design have so far been obtained in a few special cases only. Specifically, these are the MF magnitude FIR filter and the MF delay IIR filter. For equal ripple magnitude or delay, numerical techniques have to be resorted to in general. The technique of Hazra and Dutta Roy for equiripple IIR delay is an exception to this. With a negligibly small deviation from the strictly equiripple condition, it achieves a remarkably simple solution without the need for solving a set of simultaneous nonlinear equations, as required in the other approaches available for the purpose. Besides $r < 1$, there are no additional stability constraints to be set or tested for. Also, it provides a certain degree of flexibility on the magnitude response.

References

1. B. Gold, K.L. Jordan, A note on digital filter synthesis. Proc. IEEE **56**, 1717–1718 (1968)
2. D.J. Nowak, Linear phase digital filter design. Proc. IEEE **57**, 850–851 (1969)
3. C.M. Rader, B. Gold, Digital filter design techniques in the frequency domain. Proc. IEEE **55**, 149–171 (1967)
4. O. Herrmann, On the approximation problem in non-recursive digital filter design. IEEE Trans. CT-**18**, 411–413 (1971)

5. L.M. Rabiner, Techniques for designing finite duration impulse response digital filters. IEEE Trans. COM-**19**, 188–195 (1971)
6. J.P. Thiran, Recursive digital filters with maximally flat group delay. IEEE Trans. CT-**18**, 659–664 (1971)
7. J.P. Thiran, Equal ripple delay recursive digital filters. IEEE Trans. CT-**18**, 6, 665–669 (1971)
8. A.G. Deczky, Recursive digital filters having equiripple group delay. IEEE Trans. CAS-**21**, 131–134 (1974)
9. S.N. Hazra, S.C.Dutta Roy, A simple design of linear phase low-pass recursive digital filters. Int. J. Circuit Theory Appl. **1**, 11–16 (1973)

Chapter 19
Generalized Linear Phase Conditions for Causal FIR Systems Revisited

Abstract It is well known that a causal FIR system has generalized linear phase when its impulse response is either symmetric or antisymmetric. Most textbooks demonstrate the sufficiency of these conditions; that they are also necessary has been dealt with by only a few authors. In this chapter, we give a critical review of these methods, including a recent one, proposed by the author, which is based on a different way of looking at the problem.

Keywords Digital signal processing · FIR systems · Generalized linear phase · Linear phase conditions

19.1 Introduction

In this section, we first review the basic concepts, including definitions of the terms linear phase, generalized linear phase, phase delay, and group delay, and the importance of linear phase. This will be followed by demonstrating the sufficiency of the linear phase conditions and listing the types of linear phase systems. The necessary and sufficient conditions for linear phase will then be specifically mentioned, along with the scope of the chapter.

19.1.1 Basic Concepts

Any system, analog or digital, is said to have generalized linear phase (GLP) if the phase $\varphi(\omega)$ is of the form

$$\varphi(\omega) = -\alpha\omega + \beta \tag{19.1}$$

Source: Suhash Chandra Dutta Roy, "Generalized Linear Phase Conditions for Casual FIR Systems Revisited," *IETE Technical Review,* vol 30, pp 398–403, September–October 2013.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_19

where α and β are constants. For analog systems, $\omega = 2\pi f$ is the angular frequency in radians/s, with f being the frequency in Hz, while for digital systems, $\omega = 2\pi f/f_s$ is the normalized digital frequency in radians, with f_s being the sampling frequency. The phase delay is defined as $-\varphi(\omega)/\omega$, while the group delay is defined as $-d\varphi(\omega)/d\omega$. Obviously, for $\beta = 0$, the phase and group delays are the same, each being equal to α, while for $\beta \neq 0$, the phase delay varies with frequency but the group delay is a constant, being equal to α. In signal processing, it is the constancy of group delay which is more important than that of phase delay. The underlying reason is that any information-carrying signal will contain a group or band of frequencies, which are together at the input of the processor. If the group delay depends on frequency, then different signals will arrive at the output at different times, causing what is known as delay distortion.

This chapter is concerned with finite impulse response (FIR) digital system, and the conditions for it to have GLP.

19.1.2 Conditions for GLP and Proof of Sufficiency

Consider a causal FIR system with impulse response $h(n)$, $n = 0$ to $N - 1$. The frequency response of the system, which is the Fourier transform (FT) of $h(n)$, is given by

$$H(e^{j\omega}) = \sum_{n=0}^{N-1} h(n)e^{-j\omega n} \tag{19.2}$$

If the impulse response is symmetric, i.e. $h(n) = h(N - n - 1)$, and the length of $h(n)$, N, is odd, then Eq. (19.2) can be rewritten as

$$H(e^{j\omega}) = h(0)[1 + e^{-j\omega(N-1)}] + h(1)[e^{-j\omega} + e^{-j\omega(N-2)}] + \ldots + h((N-1)/2)e^{-j\omega(N-1)/2} \tag{19.3}$$

By taking $e^{-j\omega(N-1)/2}$ common from each term, Eq. (19.3) can be simplified to the following:

$$H(e^{j\omega}) = e^{-j\omega(N-1)/2}\left[h((N-1)/2) + 2\sum_{n=1}^{(N-1)/2} h((N-1)/2 - n\cos(\omega n)\right] \tag{19.4}$$

The expression within the square brackets of Eq. (19.4) is real, but it can be positive or negative. Hence, it is called pseudo-magnitude, not magnitude. The phase of $H(e^{j\omega})$ is $-\omega(N-1)/2$ as long as the pseudo-magnitude does not change sign; every time it does, there will be an increment of π in phase. Hence, the phase is linear, piecewise linear to be precise, and the group delay is $(N-1)/2$ samples. The

significance of linear phase, as mentioned earlier, is that a signal passing through the system will undergo magnitude filtering, as needed, but will have no delay distortion. This feature is not realizable in either analog systems or digital systems having infinite impulse response (IIR), and is a distinct advantage of FIR digital systems.

The case we have just considered is given the name Type 1 linear phase in the literature. If $h(n)$ is symmetric but N is even, then the summation in Eq. (19.4) shall remain the same, except that the last term will be

$$e^{-j\omega(N-1)/2}h((N-2)/2)[e^{j\omega/2}+e^{-j\omega/2}] \tag{19.5}$$

Once again, taking $e^{-j\omega(N-1)/2}$ common from each term, and simplifying, the frequency response can be written as

$$H(e^{j\omega}) = 2e^{-j\omega(N-1)/2}\sum_{n=1}^{N/2} h((N/2)-n)\cos(\omega[n-(1/2)]) \tag{19.6}$$

This case is designated as Type 2 linear phase.

In a similar manner, one can show that if $h(n)$ is antisymmetric, i.e. $h(n) = -h(N-1-n)$, the frequency responses are given by

$$H(e^{j\omega}) = 2je^{-j\omega(N-1)/2}\sum_{n=1}^{(N-1)/2} h((N-1)/2)-n)\sin(\omega n) \tag{19.7}$$

for N odd, and

$$H(e^{j\omega}) = 2je^{-j\omega(N-1)/2}\sum_{n=1}^{N/2} h(N/2)-n)\sin[\omega(n-(1/2))] \tag{19.8}$$

for N even. These are designated as Type 3 and Type 4 linear phase FIR systems, respectively. Note that all the four types have GLP, because the pseudo-magnitude can be positive or negative, while for types 3 and 4, in addition, the factor j contributes an additional phase of $\pi/2$.

19.1.3 Necessary and Sufficient Conditions

We have just demonstrated that if the impulse response is symmetric or antisymmetric, then the system has GLP. These are sufficient conditions; are they also necessary? To show that they are necessary, we have to prove that if the system has GLP, then the impulse response must be symmetric or antisymmetric. Barring a few, most of the textbooks do not deal with this aspect. Various methods have been used in the literature to prove necessity; we shall review them critically in the next few sections,

including a recent method proposed by the author. This last method is based on a different approach to the problem, and is believed to be crisper and more straightforward than the other methods.

19.2 Method 1

This method appeared in Ref. [1] and is based on Fourier series properties. In this, it is argued that a linear phase FIR frequency response can be expressed as

$$H(e^{j\omega}) = A(\omega)e^{-j\omega\alpha+j\beta} \tag{19.9}$$

where $A(\omega)$ is real, α is the delay (in number of samples), and β is a constant. This expression is validated by the four transfer functions given in Eqs. (19.4) and (19.6)–(19.8), where for the first two, $\beta = 0$, and for the last two, $\beta = \pi/2$, and for all of them, $\alpha = (N-1)/2$. Note that writing the four transfer functions with a negative sign would not have affected the GLP character. Thus, what we have to prove is that if the transfer function has GLP, then we must have (1) $\alpha = (N-1)/2$ and either (2) $e^{j\beta} = \pm 1$ and $h(n) = h(N-1-n)$ or (3) $e^{j\beta} = \pm j$ and $h(n) = -h(N-1-n)$.

Equating Eqs. (19.2) and (19.9), we can write

$$A(\omega) = \sum_{n=0}^{N-1} h(n)e^{-j[\omega(n-\alpha)+\beta]} \tag{19.10}$$

Since $A(\omega)$ is real, the imaginary part of Eq. (19.10) must vanish, i.e.

$$\sum_{n=0}^{N-1} h(n)\sin[\omega(n-\alpha)+\beta] = 0 \tag{19.11}$$

This is a Fourier series type expression which implies that any solution for α and β must be unique. It is easily verified that Eq. (19.11) can be satisfied in two ways: (1) for $\beta = 0$, $h(n) = h(N-1-n)$ and (2) for $\beta = \pi/2$, $h(n) = -h(N-1-n)$, and in both cases, $n - \alpha = -(N-1-n-\alpha)$, i.e. $\alpha = (N-1)/2$.

Two questions arise at this point. First, Eq. (19.11) is not a Fourier series, even if truncated, because $\omega(n-\alpha)$ is not a harmonic frequency if α is not an integer. Secondly, it is not clear how these special values of β were arrived at. When these points were raised in a paper, an anonymous reviewer commented that 'Fourier methods are essentially based on orthogonality of certain sets of functions, the sine functions in this case, and as such, are far more general than just Fourier series, which is based upon orthogonality properties of particular countable subsets of these functions (again sine functions in this case)'. Hence, 'deductions from Eq. (19.11) by Fourier methods are entirely appropriate'.

Readers who are not comfortable with the line of reasoning followed in this method may look at method 6 given later in this chapter, where also summations with non-integer values of $n - \alpha$ occur, but the orthogonality argument does not have to be evoked, at least not explicitly.

19.3 Method 2

In this method, given in Refs. [2] and [3], the possible values of β are first established and then symmetry or antisymmetry conditions are proved. Since $h(n)$ is real, $H(e^{j\omega})$ must be Hermitian symmetric (both Refs. [2] and [3] mention conjugate symmetry; that Hermitian symmetry is the correct property was pointed out by the anonymous reviewer mentioned in the previous section), i.e.

$$H(e^{j\omega}) = H^*(e^{-j\omega}) \tag{19.12}$$

as can be easily verified from Eq. (19.2). In Eq. (19.12), the superscript * denotes complex conjugate. Combining Eq. (19.12) with (19.9), we get

$$A(\omega)e^{-j\omega\alpha+j\beta} = A(-\omega)e^{-j\omega\alpha-j\beta} \tag{19.13}$$

Since the magnitudes of $H(e^{j\omega})$ and $A(\omega)$ are equal, $A(\omega)$ must be either even or odd, i.e. $A(-\omega) = \pm A(\omega)$. If $A(\omega)$ is an even function, then the positive sign is applicable, and Eq. (19.13) gives

$$e^{j2\beta} = 1 \tag{19.14}$$

Thus, $\beta = 0$ or π, and from Eq. (19.10), we can write

$$A(\omega) = \pm \sum_{n=0}^{N-1} h(n)e^{-j\omega(n-\alpha)} \tag{19.15}$$

and

$$A(-\omega) = \pm \sum_{r=0}^{N-1} h(r)e^{-j\omega(r-\alpha)} \tag{19.16}$$

The positive signs in Eqs. (19.15) and (19.16) hold for $\beta = 0$ and the negative signs hold for $\beta = \pi$. Note that the variable in Eq. (19.16) has been changed from n to r, without any loss of generality, for convenience of carrying out the next step, which is to put $r = N - 1 - n$. The result is

$$A(-\omega) = \pm \sum_{n=0}^{N-1} h(N-1-n) e^{j\omega(N-1-n-\alpha)} \tag{19.17}$$

Equating (19.15) and (19.17), we arrive at the conditions

$$h(n) = h(N-1-n) \text{ and } \alpha = (N-1)/2 \tag{19.18}$$

On the other hand, if $A(\omega)$ is odd, i.e. $A(-\omega) = -A(\omega)$, then Eq. (19.13) gives

$$e^{j2\beta} = -1, \tag{19.19}$$

i.e. $\beta = \pm\pi/2$, and Eq. (19.10) gives

$$A(\omega) = \pm j \sum_{n=0}^{n-1} h(n) e^{-j\omega(n-\alpha)} \tag{19.20}$$

By following a similar procedure as in the previous case (i.e. change of variable), it is easy to show that

$$A(-\omega) = \pm j \sum_{n=0}^{N-1} h(N-1-n) e^{j\omega(N-1-n-\alpha)} \tag{19.21}$$

In the above two equations, the positive signs hold for $\beta = -\pi/2$ and the negative signs hold for $\beta = +\pi/2$. Now using $A(-\omega) = -A(\omega)$, we get from Eqs. (19.20) and (19.21), the conditions

$$h(n) = -h(N-1-n) \text{ and } \alpha = (N-1)/2 \tag{19.22}$$

This completes the proof.

19.4 Method 3

As compared to the other methods, method 3, given in Ref. [4], has the distinctive feature that it does not assume the system to be FIR to start with. Instead, it starts with the fact that $H(e^{j\omega})$ is periodic with period 2π. Thus from Eq. (19.9), we should have

$$A(\omega) e^{-j\omega\alpha + j\beta} = A(\omega + 2\pi) e^{-j(\omega+2\pi)\alpha + j\beta} \tag{19.23}$$

On simplification, this gives

$$A(\omega) = A(\omega + 2\pi)e^{-j2\pi\alpha} \tag{19.24}$$

Since $A(\omega)$ is real, 2α must be an integer. Note, in passing, that if α is an integer, then $A(\omega)$ is periodic with period 2π and that if α is an integer plus 1/2, then $A(\omega)$ is periodic with period 4π. If neither is the case, then $A(\omega)$ is not periodic.

Now, as in method 2, real $h(n)$ implies that $H(e^{j\omega})$ is Hermitian symmetric, i.e. Eqs. (19.12) and (19.13) are valid, and from the latter, in general, we can write

$$e^{j2\beta} = A(-\omega)/A(\omega) \tag{19.25}$$

Again, since $A(\omega)$ is real, it follows that $e^{j2\beta}$ is necessarily real. As before, even $A(\omega)$ implies $\beta = 0$ and odd $A(\omega)$ implies $\beta = \pi/2$. In retrospect, it is clear that Ref. [4] uses essentially the same procedure as in Refs. [2] and [3] to establish the possible values of β, but does not admit the values π and $-\pi/2$.

From this point onwards, the procedure deviates from that of Refs. [2] and [3]. In general, from Eq. (19.9), one can write

$$A(\omega) = H(e^{j\omega})e^{j\omega\alpha - j\beta} \tag{19.26}$$

The sample $h(2\alpha - n)$ of the impulse response can be obtained from the inverse FT as follows:

$$h(2\alpha - n) = [1/(2\pi)] \int_{-\pi}^{\pi} He^{j\omega} e^{j\omega(2\alpha - n)} d\omega \tag{19.27}$$

For $\beta = 0$, combining Eqs. (19.9) and (19.27) gives

$$h(2\alpha - n) = [1/(2\pi)] \int_{-\pi}^{\pi} A(\omega) e^{j\omega(\alpha - n)} d\omega \tag{19.28}$$

Since $h(2\alpha - n)$ is real, we can replace the right-hand side of Eq. (19.28) by its complex conjugate to get

$$h(2\alpha - n) = [1/(2\pi)] \int_{-\pi}^{\pi} A(\omega) e^{j\omega(n - \alpha)} d\omega \tag{19.29}$$

where we have used the fact that $A(\omega)$ is real. Now combine Eq. (19.29) with (19.26) for $\beta = 0$; the result is

$$h(2\alpha - n) = [1/(2\pi)] \int_{-\pi}^{\pi} H(e^{j\omega}) e^{j\omega n} d\omega = h(n) \tag{19.30}$$

Since the system is causal, $h(n) = 0$ for $n < 0$, and therefore, also for $n > 2\alpha$. Thus, the system is FIR, and if $h(n)$ exists for $n = 0$ to $N - 1$, 2α must be $N - 1$, i.e. $\alpha = (N - 1)/2$, and by Eq. (19.30), $h(n) = h(N - 1 - n)$. Note that this procedure involves going back and forth between $A(\omega)$ and $H(e^{j\omega})$.

In the second case of $\beta = \pi/2$, Eq. (19.9) becomes

$$H(e^{j\omega}) = jA(\omega)e^{-j\omega\alpha} \tag{19.31}$$

Putting this in Eq. (19.27) gives

$$h(2\alpha - n) = [j/(2\pi)] \int_{-\pi}^{\pi} A(\omega)e^{j\omega(\alpha-n)}d\omega \tag{19.32}$$

Again take the complex conjugate of both sides, and take account of the realness of $h(2\alpha - n)$ and $A(\omega)$ to get

$$h(2\alpha - n) = [j/(2\pi)] \int_{-\pi}^{\pi} A(\omega)e^{j\omega(n-\alpha)}d\omega \tag{19.33}$$

Finally, substitute for $A(\omega)$ from Eq. (19.31) to get

$$h(2\alpha - n) = -[1/(2\pi)] \int_{-\pi}^{\pi} H(e^{j\omega})e^{j\omega n}d\omega = -h(n) \tag{19.34}$$

By similar arguments as in the previous case, we conclude that $h(n)$ is of finite length N, with $\alpha = (N - 1)/2$, and by Eq. (19.34), $h(n) = -h(N - 1 - n)$.

19.5 Method 4

This method appeared in Ref. [5] and has the interesting feature that it does not assume $h(n)$ to be real to start with. In general, $h(n)$ is the inverse FT of $H(e^{j\omega})$, i.e.

$$h(n) = [1/(2\pi)] \int_{-\pi}^{\pi} H(e^{j\omega})e^{j\omega n}d\omega \tag{19.35}$$

Combining this with Eq. (19.9), we get

$$h(n) = e^{j\beta}([1/(2\pi)] \int_{-\pi}^{\pi} A(\omega) e^{j\omega(n-\alpha)} d\omega) \tag{19.36}$$

Recognize that the quantity which multiplies $e^{j\beta}$ in Eq. (19.36) is the sample $a(n - \alpha)$ in the inverse FT of $A(\omega)$. Hence,

$$h(n) = e^{j\beta} \alpha(n - \alpha) \tag{19.37}$$

which can also be written as

$$\alpha(n) = e^{-j\beta} h(n + \alpha) \tag{19.38}$$

Since $A(\omega)$ is real, $a(n)$ must be Hermitian symmetric (as can be easily verified from the inverse FT expression), i.e.

$$a(n) = a * (-n) \tag{19.39}$$

Combining this with Eq. (19.38) gives

$$h(n + \alpha) = e^{j2\beta} h * (-n + \alpha) \tag{19.40}$$

Putting $n = -\alpha$ in Eq. (19.40) gives

$$h(0) = e^{j2\beta} h * (2\alpha) \tag{19.41}$$

From this, it is argued in Ref. [5] that if $h(n)$ is causal and of finite duration, from $n = 0$ to $n = N - 1$, one must necessarily have

$$\alpha = (N - 1)/2 \tag{19.42}$$

With regard to this step, the author of Ref. [6] commented that 'there is a logic flaw in this derivation', and that 'in the next addition of the same text in Ref. [7], the authors side-step the mistake by providing a weaker version under the assumption that the group delay is a multiple of 1/2 sample, which is short of the most general conclusion'.

However, we shall 'side-step' this criticism and proceed further, as in Ref. [5]. Replacing n by $n - \alpha$ in Eq. (19.40) and using Eq. (19.42) gives

$$h(n) = e^{j2\beta} h * (N - 1 - n) \tag{19.43}$$

which is the general equation that the coefficients of a linear phase FIR system must satisfy. When $h(n)$ is real, $h(n) = h^*(n)$, and Eq. (19.43) implies that $\beta = k\pi/2$, where k is an integer, positive or negative; thus Eq. (19.43) gives

$$h(n) = (-1)^k h * (N - 1 - n) \tag{19.44}$$

i.e. the impulse response must be either symmetric or antisymmetric. Finally,

$$H(e^{j\omega}) = A(\omega)e^{-j\omega[(N-1)/2]+j(k\pi/2)} \tag{19.45}$$

For all practical purposes, it is sufficient to consider only $k = 0$ and 1 because all other values of k will be equivalent to one of these cases. Accordingly, $k = 0$ and N odd gives Type 1; $k = 0$ and N even gives Type 2; $k = 1$ and N odd gives Type 3; and $k = 1$ and N even gives Type 4 linear phase filters.

19.6 Method 5

This method appeared in Ref. [6], where, with regard to Eq. (19.11), the author commented that 'this result is inconvenient to use for further derivation', but did not explain why. In this method, $h(n)$ is assumed real, and it is first proved that the group delay α must be an integer or half integer, 'an issue which is overlooked or not carefully dealt with in most texts on digital signal processing' [6].

Note that Eq. (19.1) can be expressed as

$$H(e^{j\omega}) = H_R(e^{j\omega}) + jH_I(e_{j\omega}) \tag{19.46}$$

where the subscripts R and I stand for the real and imaginary parts, respectively, with

$$H_R(e^{j\omega}) = \sum_{n=0}^{N-1} h(n)\cos(\omega n) \tag{19.47}$$

and

$$H_I(e^{j\omega}) = \sum_{n=0}^{N-1} h(n)\sin(-\omega n) \tag{19.48}$$

Using Eqs. (19.46)–(19.48) in Eq. (19.26), we get after simplification,

$$A(\omega) = A_R(\omega) + A_I(\omega) \tag{19.49}$$

where

$$A_R(\omega) = H_R(e^{j\omega})\cos(\alpha\omega - \beta) - H_I(e^{j\omega})\sin(\alpha\omega - \beta) \tag{19.50}$$

and

$$A_I(\omega) = H_R(e^{j\omega}) \sin(\alpha\omega - \beta) + H_I(e^{j\omega}) \cos(\alpha\omega - \beta) \tag{19.51}$$

Since $A(\omega)$ is real, its imaginary part must be identically zero. As in method 3, the author then invokes the periodicity of $H(e^{j\omega})$ to conclude that $e^{j2\pi\alpha}$ must be a real number, i.e. α must be an integer or half integer. The next step is to use the same argument as in method 2, viz. that $H(e^{j\omega})$ must be Hermitian symmetric, which leads to the conclusion that $e^{j2\beta}$ must be either $+1$ or -1.

Instead of considering the specific possibilities for the value of β, Ref. [6] considers, as in method 3, the sample $h(2\alpha - n)$ of the impulse response. Replacing n by $2\alpha - n$ in Eq. (19.35), and using Eq. (19.9), we get after simplification,

$$h(2\alpha - n) = [1/(2\pi)] \int_{-\pi}^{\pi} A(\omega)^{j\omega\alpha + j\beta} e^{-j\omega n} d\omega \tag{19.52}$$

which can be rewritten as

$$\begin{aligned} h(2\alpha - n) &= e^{j2\beta} \left([1/(2\pi)] \int_{-\pi}^{\pi} A(\omega) e^{j\omega\alpha + j\beta} e^{-j\omega n} d\omega \right) * \\ &= e^{j2\beta} \left([1/(2\pi)] \int_{-\pi}^{\pi} H(e^{j\omega}) e^{j\omega n} d\omega \right) * \\ &= e^{j2\beta} h * (n) \end{aligned} \tag{19.53}$$

Here we have used the facts that $h(n)$ and $A(\omega)$ are real. Note that as in method 3, this method also involves going back and forth between $A(\omega)$ and $H(e^{j\omega})$. The proof is now complete, showing symmetry or antisymmetry of $h(n)$, depending on whether $e^{j2\beta}$ is $+1$ or -1.

It is clear that method 5 uses part of method 2 [2, 3], and has been significantly influenced by method 3 [4]. However, it is surprising that although the author comments on methods 1 and 4, he makes no reference to [2–4] in his paper in Ref. [6].

19.7 Method 6

We finally present here yet another method proposed by the author [8], which looks at the problem from a different perspective. We use the same method as in Refs. [2] and [3] to derive the possible values of β. Next, we rewrite Eq. (19.9) as

$$A(\omega) = e^{-j\beta} \sum_{n=0}^{N-1} h(n) e^{-j\omega(n-\alpha)} \tag{19.54}$$

Denoting the summation in Eq. (19.54) by $G(e^{j\omega})$ and letting $m = n - \alpha$ and $g(m) = h(m + \alpha)$, we get

$$G(e^{j\omega}) = \sum_{m=-\alpha}^{N-1-\alpha} g(m) e^{-j\omega n} \tag{19.55}$$

When α is an integer, $G(e^{j\omega})$ is clearly the FT of $g(n)$. When α is not an integer, $G(e^{j\omega})$ is like the FT of $g(n)$ with the difference that n takes non-integer values, but at intervals of unity. In a sense, this is time-domain dual of the case in Eq. (19.11), which is in the frequency domain, when α is not an integer.

If $\beta = 0$ or π, then since $A(\omega)$ is real, Eq. (19.54) dictates that $G(e^{j\omega})$ must necessarily be a real quantity, which is possible only if $g(n)$ is an even sequence starting at $n = -\alpha$ and ending at $N - 1 - \alpha$. This implies that

$$N - 1 - \alpha = \alpha \tag{19.56}$$

or

$$\alpha = (N - 1)/2 \tag{19.57}$$

Further, $g(n) = g(-n)$ leads to $g(n - \alpha) = g(\alpha - n)$, and therefore,

$$h(n) = h(2\alpha - n) = h(N - 1 - n) \tag{19.58}$$

i.e. $h(n)$ is symmetric.

Similarly, when $\beta = \pm\pi/2$, $e^{-j\beta} = \pm j$, and Eq. (19.54) dictates that $G(e^{j\omega})$ must be purely imaginary. This is possible only if $g(n)$ is an odd sequence. This implies Eq. (19.57) and $g(n) = g(-n)$, which, as in the previous case, translates to

$$h(n) = -h(2\alpha - n) = -h(N - 1 - n) \tag{19.59}$$

Thus $h(n)$ is antisymmetric. This completes the proof.

19.8 Conclusions

In this chapter, the importance of a linear phase system has been explained and the sufficient conditions for an FIR system to have GLP have been demonstrated by deriving the four types of such systems. For proving the necessity of these conditions,

six methods have been critically examined. These methods have varying degrees of ingenuity and innovation. Except for method 4, where a crucial step has not been adequately justified, all other methods appear to be sound, if one accepts the explanation offered by the anonymous reviewer for method 1. However, even with this explanation, method 6 appears to have an edge over method 1 because it is based on clear logic, and is believed to be crisper and more direct than the other four methods.

To be fair, it must be pointed out that method 6 is not as crisp as it appears, because it uses method 2 for finding the possible values of β, and then takes a different route to find α and the symmetry or antisymmetry of the impulse response.

This chapter supports the universal truth, stated in the Rig Veda as '*Ekam satyam vipra bahudha vadanti*', or in the English saying [9], 'Our paths may be different. Our destination is the same', or in what Thakur Ramakrishna Paramahamsa said in Bengali, '*Jato mat tato path*'.

Acknowledgements The author thanks the anonymous reviewer mentioned under method 1 for his comments, Professor P. S. R. Diniz for supplying the relevant pages of Refs. [5] and [7], and Sumantra Dutta Roy for his help in the preparation of this chapter.

References

1. L.R. Rabiner, B. Gold, *Theory and Applications of Digital Signal Processing*, (Englewood Cliffs: New Jersey, 1975)
2. T.J. Cavicchi, *Digital Signal Processing* (Wiley, New York, 2000)
3. S.K. Mitra, *Digital Signal Processing—A Computer Based Approach* (McGraw Hill, New York, 2011)
4. B. Porat, *A Course in Digital Signal Processing* (Wiley, New York, 1997)
5. P.S. Diniz, *Digital Signal Processing; System Analysis and Design* (Cambridge University Press, Cambridge, 2002)
6. S. Zhang, Note on generalised linear phase property of discrete time finite impulse response systems. Circ. Syst. Signal Process. **31**, 385–388 (2012)
7. P.S. Diniz, *Digital Signal Processing; System Analysis and Design* (Cambridge University Press, Cambridge, 2010)
8. S.C. Dutta Roy, A note on generalised linear phase conditions for causal FIR systems. Circ. Syst. Signal Process. **32**, 3131–3134 (2013)
9. A. Tripathi, *The Spirit of the Nagas* (Westland Ltd., Chennai, 2012)

Chapter 20
Review of Pseudo-Boolean Methods with Applications to Digital Filter Design

Abstract The technique of pseudo-Boolean methods which forms the basis for bivalent (0, 1) programming, has been used for many socio-economic and engineering problems in the past. In this chapter, we review these methods and discuss their applications to the quantized coefficient design of digital filters. Some other applications of these methods are also briefly described.

Keywords Pseudo-Boolean methods · Digital filters · Coefficient quantization

20.1 Introduction

There exists a number of socio-economic and engineering problems, which can be solved through integer programming formulation, since the variables in these problems take only integer values [1]. There is a subset of these problems in which the variables take only two values, viz. 0 and 1 [2, 3]. Although integer programming methods can be applied to this subset also, bivalent programming, based on a combination of dynamic programming with Boolean techniques, are particularly suited to them. These are called Pseudo-Boolean (PB) methods [4, 5].

Our attention was drawn to PB methods during the course of investigation of some coefficient sensitivity problems in digital filters [6], and a number of useful design procedures resulted from the application of these methods. In this chapter, we first review some important aspects of PB methods. In particular, we will describe methods for obtaining the solutions of PB equations and inequalities and also, for finding the minima of a PB function. This is followed by a discussion on the application of PB methods to the quantized coefficient design of digital filters. Such design procedures, necessitated by the finite word-length implementation in digital filters, yield the set of quantized filter coefficients in a discrete-parameter space, optimally. The chapter is concluded with a brief overview of some other applications of PB methods.

Source: R.K. Patney & S.C. Dutta Roy, "A Review of Pseudo –Boolean Methods with Applications to Digital Filter Design," *JIETE*, vol 26, pp 1–8, January 1980.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_20

20.2 Pseudo-Boolean Functions

We start by defining a Boolean algebra B, consisting of two elements 0 and 1, and the operations of disjunction (U), conjunction (.) and negation which can be performed on them.

A PB function is defined as a real-valued function $F(x_1, \ldots, x_n \ldots)$, where $x_i's$ are bivalent variables i.e. $x_i \in \{0, 1\}$, $i = 1, 2, \ldots n$. An important property of PB functions is that they can be written in terms of an arithmetic expression. This is possible because the basic operations of B can be expressed in arithmetic form i.e. for $x, y \in \{0, 1\}$, $xUy = x + y - xy$, $x \,.\, y = xy$ and $\bar{x} = 1 - x$, where bar denotes negation. Another useful property is that every PB function can be made linear in any one of its bivalent variables. In other words, we can write

$$F(x_1, \ldots, x_n) = x_i F_1(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) + F_2(x_1, \ldots, x_{i-1}, x_{i+1} \ldots, x_n) \tag{20.1}$$

which is linear in the variable x_i; in Eq. (20.1)

$$F_1(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = F(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) - F(x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_n) \tag{20.2}$$

and

$$F_2(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = F(x_1, x_{i-1}, 0, x_{i+1} \ldots, x_n) \tag{20.3}$$

As an example, consider $F = x_1 + x_2 + 2x_1x_2 + 4x_2x_3 + x_1x_2x_3$ for which $F_1 = 1 + 2x_2 + x_2x_3$ and $F_2 = x_2 + 4x_2x_3$, if F is made linear in x_1.

20.3 Linear PB Equations and Inequalities

A major task in PB methods is to find the solutions of a PB equation or inequality. We first consider a linear PB equation which is given in its arithmetic form as

$$a_1z_1 + a_2z_2 + \cdots + a_nz_n = t \tag{20.4}$$

where a_i, ($i = 1, 2, \ldots, n$) and t are real numbers and z_i ($i = 1, 2, \ldots n$) $\in \{0, 1\}$. In order to solve Eq. (20.4), the following substitution is made

$$x_i = \begin{cases} z_i & \text{if } a_i > 0 \\ \bar{z}_i = 1 - z_i & \text{if } a_i < 0 \end{cases} \tag{20.5}$$

Then Eq. (20.4) can be written in its canonic form

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n = d \tag{20.6}$$

where, $c_i = |a_i|, i = 1, 2, \ldots n, x_i\ (i = 1, 2, \ldots, n) \in \{0, 1\}$ and $d = \sum_r |a_r| + t$ for all r such that $a_r < 0$. Also the terms in Eq. (20.6) may be reindexed such that $c_1 \geq c_2 \geq c_n > 0$. Out of a total 2^n possible combinations of the set $\{x_i\}$, the solutions of Eq. (20.4) can be very easily tracked by following a simple set of rules as given in Table 20.1. For example, if in Eq. (20.6), $d > \sum_i c_i$, there exist no solutions, while if $d = 0$ then the unique solution is $x_i = 0$ for all i. After all the solutions of Eq. (20.6) are obtained, the inverse of the transformation (20.5) is applied to get all the solutions of Eq. (20.4).

If the equality sign in Eq. (20.4) is replaced by an inequality sign ($\geq$),[1] we get a linear PB inequality. This can also be converted to its canonic form

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n > d \tag{20.7}$$

by making use of the transformation given in Eq. (20.5). The solutions of Eq. (20.7) come grouped into families, a family of solutions being one in which some variables take fixed values and the remaining ones can take any arbitrary value from the set $\{0, 1\}$. These families of solutions of Eq. (20.7) can be tracked by using the set of

Table 20.1 Rule for tracking the solutions of PB equation (20.6)

No.	Case	Conclusions
1.	$d < 0$	No solutions
2.	$d = 0$	The unique solution is $x_1 = x_2 = \ldots = x_n = 0$
3.	$d > 0$ and $c_1 \geq \ldots \geq c_p \geq d \geq c_{p+1} \geq \ldots \geq c_n$	The solution, if any, satisfy $x_1 = \ldots = x_p = 0$ and $\sum_{j=p+1}^{n} c_j x_j = d$
4.	$d > 0$ and $c_1 = \ldots = c_p = d > c_{p+1} \geq \ldots \geq c_n$	There are $p + 1$ approaches: $a_k)\ x_k = 1, x_1 = \ldots = x_{k-1} = x_{k+1} = \ldots = x_n = 0$ for each $k = 1, 2, \ldots p$ and $a_{p+1})$ The other solutions, if any, satisfy $x_1 = \ldots = x_p = 0$ and $\sum_{j=p+1}^{n} c_j x_j = d$
5	$d > 0, c_i < d$ for all i and $\sum_{i=1}^{n} c_i < d$	No solutions
6.	$d > 0, c_i < d$ for all i and $\sum_{i=1}^{n} c_i = d$	The unique solution is $x_1 = x_2 = \ldots = x_n = 1$
7.	$d > 0, c_i < d$ for all i, $\sum_{i=1}^{n} c_i > d$ and $\sum_{j=2}^{n} c_j < d$	The solution, if any, satisfy $x_1 = 1$ and $\sum_{j=2}^{n} c_j x_j = d - c_1$
8.	$d > 0, c_i < d$ for all i, $\sum_{i=1}^{n} c_i > d$ and $\sum_{j=2}^{n} c_j \geq d$	There are two approaches: $\beta_1)\ x_1 = 1$ and solutions, if any, of $\sum_{j=2}^{n} c_j x_j = d - c_1$ $\beta_2)\ x_1 = 0$ and solutions, if any, of $\sum_{j=2}^{n} c_j x_j = d$

[1] Note that other forms of inequalities i.e. $<$, $\leq$ and $>$ can easily be converted to the form $\geq$.

rules given in Table 20.2, this being similar to the use of Table 20.1 for the case of a PB equation.

A system of PB equations and inequalities, whose members have been transformed into their canonic forms, is similarly solved by applying the set of rules in Table 20.1 or Table 20.2, to a choice member of the system depending upon whether this choice member is an equation, or an inequality. The information so obtained about the values of some of the variables is used to reduce the order of the whole system. This process is repeated until all the solutions of the system are obtained. The key to the efficiency of this method lies in determination of the choice member of the system, which is subjected to either Table 20.1 or Table 20.2; a stepwise method for doing this is available in the literature [3, 4].

Table 20.2 Rules for tracking the solutions of PB inequality given by (21.7)

No.	Case	Conclusions
1.	$\underline{d} \leq 0$	Redundant inequality i.e. $x_1 = x_2 \ldots = y_n =$ arbitrary
2.	$d > 0$ and $c_1 \geq \ldots > c_p \geq d \geq c_{p+1} \geq \ldots \geq c_n$	There are p + 1 approaches: a_k) For every $k = 1, 2, \ldots$ p $x_1 = \ldots = : k_{-1} = 0$; $x_k = 1$ and $\boxed{x_{k+1} = x_n = \text{arbitrary}}$ (The part within the rectangle is to be deleted for the case of a system of PB inequalities) a_{p+1}) $x_1 = \ldots = x_p = 0$ and solutions, if any, of $\sum_{j=p+1}^{n} c_j x_j \geq d$
3.	$d > 0$, $c_i < d$ for all and $\sum_{i=1}^{n} c_i < d$	No solutions
4.	$d > 0$, $c_i < d$ for all i and $\sum_{i=1}^{n} c_i = d$	The unique solution is $x_1 = \ldots = x_n = 1$
5	$d > 0$, $c_i < d$ for all i, $\sum_{i=1}^{n} c_i < d$ and $\sum_{j=2}^{n} c_j < d$	$x_1 = 1$ and solutions, if any, of the inequality $\sum_{j=2}^{n} c_j x_j \geq d - c_1$
6.	$d > 0$, $c_i < d$ for all i $\sum_{i=1}^{n} c_i = d$ and $\sum_{j=2}^{n} c_j \geq d$	There are two approaches: β_1) $x_1 = 1$ and solutions, if any, of the inequality $\sum_{j=2}^{n} c_j;\ x_j \geq d - c_1$ β_2) $x_1 = 0$ and solutions, if any, of the inequality $\sum_{j=2}^{n} c_j x_j \geq d$

20.4 Nonlinear PB Equations and Inequalities

In order to obtain the solutions of a nonlinear PB equation or inequality, it is first linearized. This is easily done for the case of PB functions by observing that if $y_1, y_2 \ldots y_m$ are bivalent variables, then so is $x_k = y_1, y_2 \ldots, y_k$, $k \leq m$. Thus, a nonlinear equation or inequality

$$F(y_1, y_2, \ldots, y_m) \square 0 \tag{20.8}$$

where '$\square$' is either '=' or '$\geq$', can be converted to a linear equation of the form (20.6) or a linear inequality of the form (20.7), which can be solved by using Table 20.1 or Table 20.2 respectively. Let $(x_1^*, x_2^* \ldots x_n^*)$ be a solution of the linearized form of Eq. (20.8). Then there is a Boolean function f^* such that $f^*(x_1^*, x_2^*, \ldots x_n^*) = 1$. This is called the characteristic equation, f^* being the characteristic function of the solution (*) [3]. The characteristic function f of all the solutions of the linearized form of Eq. (20.8) is then given by $f = f^1 \mathrm{U} f^s U \ldots U f^2$.

Now if each x_k in f is given its corresponding expansion in terms of $y_i's$ then f will represent the solutions of Eq. (20.8), which can be obtained after simplifying this expression, by inspection.

As an example, consider the equation

$$4y_1y_2y_3 + 3y_1y_2 + 2y_1 + y_3 = 5 \tag{20.9}$$

If we let $y_1y_2y_3 = x_1, y_1y_3 = x_2, y_1 = x_3$ and $y_3 = x_4$, we get a linear Eq. $4x_1 + 3x_2 + 2x_3 + x_4 = 5$ whose solutions, by using Table 20.1, are (a) $x_1 = x_4 = 1, x_2 = x_3 = 0$, and (b) $x_1 = x_4 = 0, x_2 = x_3 = 1$. Thus $f(x_1, x_2, x_3, x_4) = x_1\bar{x}_2\bar{x}_3x_4 \,\mathrm{U}\, \bar{x}_1x_2x_3\bar{x}_4$. Substituting for $x_i's$ in terms of $y_i's$, and simplifying gives $f(y_1, y_2, y_3) = y_1y_2\bar{y}_3$. Therefore, the only solution to Eq. (20.9) is $y_1 = y_2 = 1$ and $y_3 = 0$.

Note that while Eq. (20.9) has three bivalent variables, its linearization increases the number of bivalent variables, and hence the computational effort. In order to overcome this difficulty, an alternate method of solving a nonlinear PB inequality [6] is now given.[2] A similar method for nonlinear equations is easily obtained.

The method is based on linearizing a nonlinear PB function $F(y_1, y_2, \ldots, y_m)$ in any one of its variables, as was done in Eq. (20.1). Thus, if F is made linear in say, y_1, we may write inequality (8) as

$$y_1G_1(y_2, \ldots, y_n) \geq H_1(y_2, \ldots y_m) \tag{20.10}$$

where G_1 and H_1 are nonlinear *PB* functions of reduced order. Inequality given by Eq. (20.10) can be easily solved for y_1, using linear methods, for a given set of values

[2]The method in its preliminary form was presented in 1978 IEEE International Conference on Acoustics, Speech and Signal Processing [6].

of the variables $y_2, \ldots y_m$. The values of these variables are determined by making use of the following set of rules which help to further reduce the order of the problem.

Case-A If G_1 is always[3] >0 and H_1 is always <0, then Eq. (20.10) is a redundant inequality i.e. it is satisfied for all arbitrary y_i.

Case-B If G_1 is always >0 and H_1 is always >0, then there are two possibilities: (i) G_1 always $<H_1$, (i.e. $G_1 - H_1 < 0$) gives no solutions; (ii) otherwise, we need to solve $F_1 = G_1 - H_1 \geq 0$. For all possible solutions of this inequality, $y_1 = 1$.

Case-C If G_1 is always >0 and nothing can be said about H_1, we consider, in H_1, a variable y_k whose corresponding coefficient c_k has the largest magnitude. Let $y_k = 1$ and consider the following two possibilities: (i) If $c_k < 0$ and its magnitude is such that H_1 is always <0, then one solution to Eq. (20.10) is $y_k = 1$, y_i (i = 1, … $k-1, k+1, \ldots, m$) = arbitrary. The other solutions, if any, are obtained by solving Eq. (20.10) after substituting $y_k = 0$ in it. (ii) If $c_k > 0$ and its magnitude is such that H_1 is always $<G_1$ then the solutions, if any, are obtained by solving (20.10) after substituting $y_k = 0$ in it.

Case-D If G_1 is always <0, we replace y_1 in (10) by $(1 - \bar{y}_1)$ so that a new inequality $\bar{y}_1(-G_1) \geq H_1 - G_1$ replaces Eq. (20.10). To this, the procedure is applied again starting with Case *A*.

Case-E If H_1 is always <0 and nothing is known about G_1, then one family of solutions to Eq. (20.10) is $y_1 = 0$ and y_i ($i = 2, \ldots m$) = arbitrary. The remaining solutions are $y_1 = 1$ and solutions, if any, of the inequality $F_1 = G_1 - H_1 \geq 0$.

Case-F If Eq. (20.10) does not fall into any of these cases, we consider the inequality $-F_1 = H_1 \leq 0$ or $F_1 \geq 0$. For each of its solutions, we can find G_1 and H_1 numerically. Then, $y_1 = 0$ if $G_1 < H_1$, while if $G_1 \geq H_1$ y_1 = arbitrary.

Further, if $f_1(y_2, \ldots, y_m)$ is the characteristic function of the solutions of $F_1 \geq 0$, then $f_1(y_2, \ldots, y_m) = 0$ gives all the solutions of the inequality $F_1 < 0$, the two being complementary to each other. For all sets of variables $y_2, \ldots, y_m$ for which $f_1 = 0$ is satisfied, we have $y_1 = 1$ if $G_1 \geq H_1$.

In most of the cases discussed above, a new inequality $F_1(y_2, \ldots, y_m) \geq 0$ is required to be solved. This is similar to Eq. (20.8) but with less number of variables. It can now be made linear with respect to another variable, thus bringing it to the form of Eq. (20.10), and the procedure can be applied again. This process of reducing the order of inequality continues till we obtain

$$y_n G_n \geq H_n, n \leq m \tag{20.11}$$

such that G_n and H_n are merely numerical quantities. Thus Eq. (20.11) is easily solved for y_n using linear PB methods and this value of y_n can be substituted in

[3] Whether a given nonlinear PB function F is always >0 (<0) can easily be determined by considering the constant term in F. If this term is positive (negative) and its magnitude is greater than the sum of all negative (positive) coefficient sin F, then F is always >0 (<0).

$y_{n-1}\ G_{n-1}\ (y_n) \geq H_{n-1}$ to make it a linear inequality. Thus, going backwards, all the solutions to Eq. (20.10) are obtained.

20.5 Minimization of PB Functions

Minimization of objective functions is an important aspect of all programming methods. In the case of bivalent (0, 1) programming the objective function

$$F(y_1, \ldots, y_m) \tag{20.12}$$

will be a PB function. We first consider the *unconstrained* minimization of Eq. (20.12), starting with F expressed as

$$\begin{aligned} F &= F_1(y_1, \ldots y_m) \\ &= y_1 G_1(y_2 \ldots, y_m) + H_1(y_2, \ldots, y_m) \end{aligned} \tag{20.13}$$

i.e. we linearize it with respect to a variable $y_1 \cdot$ The method depends upon the simple fact that a linear PB function $\sum_i a_i y_i$ is minimum when y_i's are chosen equal to 1 if the corresponding $a_i < 0$ and equal to 0 otherwise. Since the right hand side of Eq. (20.13) is linear in y_1, we can choose y_1 as

$$y_1 = \begin{cases} 1 & \text{if } G_1 < 0 \\ 0 & \text{if } G_1 > 0 \\ p_1 = \text{arbitrary} & \text{if } G_1 = 0 \end{cases} \tag{20.14}$$

The last condition in Eq. (20.14) is added because we do not yet know the value of the PB function $G_1(y_2, \ldots, y_m)$. Let, now, $f_1(p_1, y_2, \ldots y_m)$ denote the characteristic function describing the relationship Eq. (20.14); then we may write

$$y_1 = f_1(p_1, y_2 \ldots, y_m) = f_{11}(y_2, \ldots, y_m) + p_1 f_{12}(y_2, \ldots, y_m) \tag{20.15}$$

where, f_{11} and f_{12} are characteristic functions of the solutions of $G_1 < 0$ and $G_1 = 0$ respectively. Substituting Eq. (20.15) into Eq. (20.13) gives

$$F_1 = (f_{11} + p_1 f_{12}) G_1 + H_1 \tag{20.16}$$

Note that $f_{12} G_1$ is always zero since when $G_1 = 0, f_{12} = 1$ while if $G_1 \pm 0, f_{12} = 0$. Thus Eq. (20.16) may be written as

$$F_1 = f_{11} G_1 + H_1 \underline{\Delta} F_2(y_2, \ldots, y_m) \tag{20.17}$$

which is a function only of $y_2, \ldots, y_m$, The above procedure is now applied to F_2 to finally get $F_3\ (y_3, \ldots, y_m) = f_{21} G_2 + H_2$. This is continued till we get,

$$F_m(y_m) = y_m G_m + H_m \tag{20.18}$$

where, G_m and H_m are real numbers. The minimum of F_m is easily obtained and, it can be proved [3] that, this will be the minimum of the original function F also.

We have obtained the minimum value of F but have not got its minimizing points. This is done by finding the value of $y_m = f_m\ (p_m)$ from Eq. (20.18), and substituting it in $y_{m-1} = f_{m-1}\ (p_{m-1}, y_m)$. This gives the value of $y_{m-1} = f_{m-1}\ (p_{m-1}, f_m)$. Thus, going backwards, we finally get $y_1 = f_1(p_1, f_2 \ldots, f_m)$ which can also be written as $y_1 = f_1(p_1, \ldots, p_m)$. This gives all the minimizing points of F.

It should be remarked that $x_i's$ do not always depend upon all the $p_i's$. In general, the number of $p_i's$ coming in the expressions for $y_i's$ is very small. Also, if at any stage of the minimization method, $F_k, k \leq m$, becomes a constant, then the remaining variables $y_{k+1}, \ldots, y_m$ can assume any arbitrary value.

We now consider minimization of Eq. (20.12) subject to the equality constraints

$$F_j(y_1, \ldots, y_m) = 0, \quad j = 1, 2, \ldots, n, \tag{20.19}$$

with the restriction that $F_j \geq 0$ for all sets of values of $(y_1, \ldots, y_m)$. To solve this problem we define S^+ and S^- as the sums of positive and negative coefficients of F. Then it can be proved [4] that a vector $(y_1^*, \ldots, y_m^*)$ minimizing Eq. (20.12), when subjected to Eq. (20.19), also minimizes the PB function:

$$G(y_1, \ldots, y_m) = F + (S^+ - S^- + 1) \sum_{j=1}^{n} F_j \tag{20.20}$$

Further, if (*) minimizes G and $G(y_1^*, \ldots, y_m^*) \leq S^+$, then (*) minimizes F under the constraints (19), while if $G(y_1^* \ldots, y_m^*) > S^+$ then the conditions Eq. (20.19) are inconsistent. Clearly, then the method requires unconstrained minimization of Eq. (20.20) and checking whether, at the minimizing point, $G \leq S^+$.

The general problem of constrained minimization, where (20.12) is to be minimized subject to the constraints.

$$F_j = (y_1, \ldots, y_m) \square\, 0, \quad j = 1, 2, \ldots, n, \tag{20.21}$$

$\square$ being either an equality or an inequality sign, can be solved using any of the following methods.

Method A: The simplest of the methods is to obtain the solutions of the constraints Eq. (20.21) simultaneously and, for each solution, find the value of the objective func-

tion. The solutions will be the solutions to the problem of constrained minimization of *F*.

Method B: In this method also, the first step is to get the solutions satisfying the constraints of Eq. (20.21). If $f(y_1, \ldots, y_m)$ is the characteristic function of these solutions, then the problem reduces to minimizing F subject to the constraint

$$f(y_1, \ldots, y_m) = 1 \tag{20.22}$$

The problem is the same as that described earlier of minimization subject to equality constraints since the Boolean equation (20.22) can be written in its equivalent PB form.

Method C: In this method an upper bound M_o of F is first found. If M_o is not known then it can be taken as equal to $S+$, the sum of all positive coefficients of *F*. We then consider the system of constraints in Eq. (20.21) along with the additional inequality $F \leq M_o$. Let $(y_1^1, y_2^1 \ldots, y_m^1)$ to get the first solution satisfying this new system. Then $F(y_1^1 \ldots y_m^1,) \underline{\underline{\Delta}} M_1$ modifies the value of the upper bound for *F*. We now consider the constraints in Eq. (20.21) along with the inequality $F \leq M_1$. This process is continued till we can no longer modify the bound for *F*. Thus we obtain the minimum of F subject to Eq. (20.21).

PB methods can also be used to minimize a *fractional* PB function

$$F(y_1 \ldots, y_m) = \frac{a_0 + a_1 y_1 + \cdots + a_m y_m}{b_0 + b_1 y_1 \cdots + \cdots + b_m y_m}, \tag{20.23}$$

where, $b_0 > 0$. One algorithm for doing this is as follows. Let I and J be the sets of all indices $i > 0$ and $j > 0$ satisfying $(a_i/b_i) \leq (a_0/b_0)$ and $(a_0/b_0) < (a_j b_j)$. The following set of rules are applied to find the minimum of Eq. (20.23):

(i) Whenever $a_k = 0$ and $b_k > 0$, put $x_k = 1$
(ii) Whenever $a_k > 0$ and $b_k < 0$, put $x_k = 0$
(iii) Determine the sets I and J
(iv) For each $j \in J$, put $x_j = 0$
(v) Let i_1 be the first index for which

$$(a_{i_1}/b_{i_1}) = \min_{i \in I}(a_i/b_i)$$

Then, if $(a_{i_1}/b_{i_1}) < (a_0/b_0)$, put $x_i = 1$ and replace a_0 by $a_0 + a_{i_1}$, b_0 by $b_0 + b_{i_1}$ and go to step 3 while if $(a_{i_1}/b_{i_1}) = (a_0/b_0)$, put $x_i = p_i$ arbitrary for each $i \in I$.

It can also be proved that Eq. (20.23) attains its minimum value if the following are satisfied.

$$Q_i^*(x_1^*, \ldots, x_n^*) \leq 0, i = 1, 2, \ldots, n \tag{20.24}$$

where

$$Q_i^* = (x_i^* - \overline{x_i^*}) \left(b_i \sum_{j=1}^{n} a_j x_j - a_i \sum_{j=1}^{n} b_j x_j \right) \tag{20.25}$$

This forms the basis for another method of minimizing Eq. (20.23) in which we start with an initial set of values $(x_1' \ldots, x_n')$ of the variables for which we compute the quantities Q_i'. We then replace x_i' by $\bar{x}'_i$ for each i for which $Q_i' > 0$. For the new set of the values of the variables, Q_i' are computed again. The process is continued till $Q_i' \leq 0$ for all i; the corresponding set of values of the variables minimizes Eq. (20.23).

After this brief review of the various PB methods, we next outline how they have been effectively employed to find a solution of the coefficient sensitivity problem in digital filters [7].

20.6 Applications to Digital Filters

In most of the digital filter design methods [10, 11], the filter coefficients are assumed to be specified to an infinite accuracy. To implement them on a digital machine, these parameters have to be quantized to a prespecified number of binary digits in order that they may fit the available word length. The resulting deterioration in filter characteristics will depend upon the word length of the filter, and the manner in which the quantized coefficients are chosen in the discrete parameter space. While the word length is decided by the cost of implementation, an optimal choice of the coefficients in discrete parameter space helps in reducing the deterioration in filter characteristics.

The easiest approach to this optimization is an exhaustive search in a predefined region of the discrete-parameter space [12]. Another approach is to use Univariate search methods [13–15]. The problem has also been solved using some continuous [16] and mixed optimization methods [17, 18]. In yet another method, a random-search technique in discrete-parameter space has been employed [19]. Since these approaches do not guarantee a global optimum, efforts have been directed to reducing the computation time needed to obtain a local optimum. The use of PB methods for this problem appreciably reduces the computation time [8, 9]. Besides, the method has various other advantages.

Let $T\left(c_i', \omega\right)$ be a function describing some characteristics of a digital filter. This function is, in general, rational in c_i', $i = 0, 1, \ldots, n$, and in trignometric functions of ω. It is assumed that c_i' are known from any of the continuous parameter design methods. Let c_i be the coefficients obtained by either rounding or truncating c_i' to a desired accuracy of $b + 1$ bits.

We may write

$$c_i = c_{i_0} \sum_{j=1}^{b} c_{ij} 2^{-j}, \quad c_{ij} \in \{0, 1\} \tag{20.26}$$

where c_{io} denotes the sign of c_i. The quantized coefficients result in a deteriorated magnitude response $T(c_i, \omega)$. The purpose of our design method is to modify the last bit of c_i, as

$$c_i \rightarrow c_i + \alpha_{ci} \tag{20.27}$$

such that the maximum of the error $|T(c_i, \omega) - D(\omega)|$, where $D(\omega)$ is the desired magnitude response, is minimized. In Eq. (20.27), $\alpha_{ci} = c_{io}\, S_{ci}\, q y_{ci}$ where $q = 2^{-b}$, is the minimum quantization step size, $y_{ci} \in \{0, 1\}$ are the bivalent variables whose values are to be determined; $S_{ci} \in \{-1, 0, 1\}$ define the direction of modification of c_i. To start with, we set $S_{ci} = \text{sign of}(c_i' - c_i)$.

Let now ω_r be the frequency at which the extrema of the error $|T(c_i, \omega) - D(\omega)| \underline{\underline{\Delta}}\, \delta_r$ is maximum. We may write the following inequality

$$|T(c_i + \alpha_{ci}, \omega_r) - D(\omega_r)| \leq \delta_r \tag{20.28}$$

This is a PB inequality in the variables y_{ci}, and can be solved by the methods described earlier in this chapter. From its solutions, we choose the one giving largest improvement in the error function. This is done by evaluating $T(c_i + \alpha_{ci}, \omega)$ at a grid of frequency points. Corresponding to the best solution, we prepare for the next iteration by modifying c_i as in (20.27), replacing δ_r in Eq. (20.28) by its new lower value and modifying S_{ci} as

$$S_{ci} \rightarrow S_{ci}(2y_{ci} - 1) \tag{20.29}$$

With the new c_i, S_{ci}, ω, and δ_r, inequality Eq. (20.28) can be formulated again, and the procedure is repeated till no more improvement in δ_r can be obtained.

The nature of the *PB* inequalities in Eq. (20.28) is different for the two cases of FIR and IIR filters. For the case of linear phase FIR filters, $T_{(ci,\omega)} = \sum_{i=0}^{n} c_i G(\omega)$ where, n depends upon the order N of the filter and $G(\omega)$ is a trignometric function. For example, for an even-ordered low-pass filter $n = (N/2) - 1$ and $G(\omega) = \sin\left(i + \frac{1}{2}\right)\omega$. For this case, then, Eq. (20.28) will be a linear PB inequality; only linear PB methods need to be applied for their solutions [8].

The magnitude response of an IIR filter is a rational function of the form $P(a_i, \omega)/Q(b_i, \omega)$, where a_i and b_i are the numerator and denominator coefficients, respectively, and P and Q are nonlinear functions of a_i, and b_i, the maximum order of the nonlinearity being N, the order of the filter. Then Eq. (20.28) can be rewritten as

$$P(a_i + \alpha_{ai}, \omega_r) - \delta Q(b_i + \alpha_{bi}, \omega_r) \leq 0 \tag{20.30}$$

where $\delta = D(\omega_r) \pm \delta_r$. The usual method of solving Eq. (20.30), by linearizing each nonlinear variable is computationally very expensive, because of an increase in the number of variables. To overcome this, the alternate method for solving nonlinear inequalities discussed earlier can be applied [8].

As an example of the design, we consider a fourth-order low-pass IIR filter with Chebyshev pass-band characteristics. The filter is to be realized in cascade [10]. With the continuous filter coefficients, the pass band varies between 1.000 and 0.891; these limits are shown in Fig. 20.1. In the same figure, curves *a* and *b* show the characteristics with rounded continuous and with PB modified coefficients, respectively.

The design method described above may be programmed to get many almost-equally-best solutions to the problem. The best solution can then be chosen depending upon an appropriate criterion, like minimum number of 1's in the solution, or their behaviour in the transition region. Also, while solving for the PB inequalities (20.28), no computations are needed on the coefficients of the filter, thus, avoiding any error due to limited word length of the computer. Finally, it is possible to implement the design method with different word lengths for different coefficients by simply modifying the quantizing condition Eq. (20.26).

The design method discussed in this section yields an optimal set of filter coefficients such that the maximum of the error is minimized. In case, the criteria is to minimize the mean-squared error, we can use a similar iterative method in which each iteration modifies the last bit of the coefficient c_i of T(c_i, ω) as in Eq. (20.27) such that

$$F = \int_{-\pi}^{\pi} (|T(c_i + \alpha_{ci}, \omega)| - |D(\omega)|)^2 d\omega \tag{20.31}$$

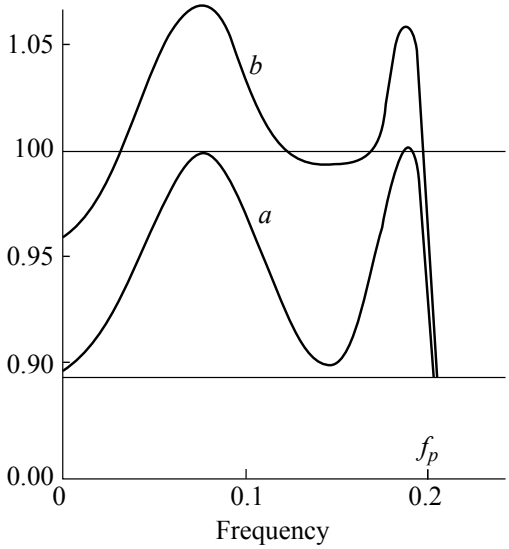


Fig. 20.1 Passband magnitude characteristics of the low-pass filter example with rounded continuous (curve a) and with PB modified (curve b) co-efficients

is minimized. This is a PB function in the bivalent variables y_{ci} whose values can be obtained by using one of the minimization algorithms described in the previous section.

Note that the algorithms used will be different for the design of FIR and IIR filters. In the case of linear phase FIR filters, Eq. (20.31) will be of the form of Eq. (20.12) and we can use the unconstrained minimization method described earlier. On the other hand if $T(c_i, \omega)$ corresponds to an IIR filter, Eq. (20.31) can be written in the following form:

$$F = F_1 + F_2 + \cdots + F_m, \tag{20.32}$$

where, F_j, $j = 1 \ldots, m$, are fractional PB functions in y_{ci}. Starting with an initial guess for the values of y_{ci}, we can evaluate the functions Q_{ij}, as in Eq. (20.25), for each F_j. It can then be proved that F in Eq. (20.32) attains its minimum value if the following are satisfied

$$\sum_{j=1}^{m} Q_{ij} \leq 0, \quad i = 1, 2, \ldots, n. \tag{20.33}$$

The values of the variables y_{ci} that minimize Eq. (20.30) also modify c_i and S_{ci} as in Eqs. (20.27) and (20.29), respectively, for the next iteration. We continue till the optimal set of c_i is obtained.

20.7 Discussion

The PB methods described in this chapter find applications in many fields, some of which we have briefly described. One example is that of problems in coding theory where the algorithms to choose the best codes for transmission over a binary channel, subject to certain error criteria, is to be found. PB methods also find applications in some problems in automata theory, as for example, in minimization of Boolean functions; note that the characteristic functions of the solutions of PB systems are also Boolean functions. These problems can be formulated in terms of minimization of a PB function subject to certain constraints. Another example where PB methods can be applied is that of graph theory, in which the connectedness and path concepts involve only two values 0 and 1.

There are many transportation and scheduling applications, where the problem is naturally formulated in PB terms e.g. in the case of selecting the location for a power generation plant, so as to reduce transmission costs subject to certain feasibility constraints.

We have mentioned earlier that the set of problems having bivalent variables can be solved using integer programming methods. It is also possible to convert the Integer Programing problems in PB terms by replacing each integer variable $\beta_i \leq$

M_i, by several bivalent variables since we can write $\beta_1 = \Sigma \beta_{ij}\, 2^{-j}$, $\beta_{ij} \in \{0, 1\}$. It has also been shown (Hammer and Rudeanu *loc. cit.*) that the linear and quadratic programming problems can also be transformed to solving a system of PB equations simultaneously.

References

1. G.B. Dantzig, *Linear Programming and Extensions* (Princeton University Press, Princeton, New Jersey, 1963)
2. P.L. Ivanescu, *Pseudo-Boolean Programming and Applications* (Springer, Berlin, 1965)
3. P.L. Ivanescu, S. Rudeanu, *Pseudo-Boolean Methods for Bivalent Programing* (Springer, Berlin, 1966)
4. P.L. Hammer, S. Rudeanu, *Boolean Methods in Operations Research and Related Areas* (Springer, New York, 1968)
5. T.L. Saaty, *Optimization in Integers and Related External Problems* (McGraw-Hill, New York, 1970)
6. R.K. Patney, S.C. Dutta Roy, Design of IIR filters using pseudo-Boolean methods, in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing* (1978)
7. R.K. Patney, *Roundoff Noise and Co-efficient Sensitivity in Digital Filters*. PhD thesis, IIT Delhi, 1978
8. R.K. Patney, S.C. Dutta Roy, Design of linear phase FIR filters using pseudo-Boolean methods. IEEE Trans. CAS-**26**, 4, 55–260 (1979)
9. R.K. Patney, S.C. Dutta Roy, A modified pseudo-Boolean method for quantized coefficient design of IIR filters, in *Proceedings of the ICASSP* (1979)
10. A.V. Oppenheim, R.W. Schafer, *Digital Signal Processing* (Prentice-Hall, New Jersey, 1975)
11. L.R. Rabiner, B. Gold, *Theory and Applications of Digital Signal Processing* (Prentice-Hall, New Jersey, 1975)
12. P. Chambon, A. Desblache, Integer co-efficient optimization of digital filters, in *Proceedings of the IEEE International Symposium on Circuits and Systems* (1976), pp. 461–464
13. E. Avenhaus, On the design of digital filters with co-efficients of finite word lengths. IEEE Trans. AU-**20**, 356–373 (1972)
14. E. Avenhaus, W. Schussler, On the approximation problem in the design of digital filters with limited world length. Arch. Elek. Ubertragung. **24**, 571–572 (1970)
15. K. Steiglitz, Designing short word recursive digital filters, in *Proceedings 9th Annual Allerton Conference on Circuit and System* (1971), pp. 778–788
16. R.E. Crochiere, A new statistical approach to the coefficient word length problem for digital filters. IEEE Trans. CAS-**22**, 190–196 (1975)
17. F. Berglez, Digital filter design with co-efficients of reduced wod length, in *Proceedings of the IEEE Symposium on Circuits and Systems* (1977), pp. 52–55
18. C. Charalambous, M.J. Best, Optimization of recursive digital filters with finite word length. IEEE Trans. ASSP- **22**, 424–431 (1974)
19. M. Suk, S.K. Mitra, Computer-aided design of digital filters with finite word length. IEEE Trans. AU-**20**, 356–373 (1972)

Chapter 21
Maximally Linear FIR Digital Differentiators: A Review

Abstract Digital differentiators (DD) are extensively used in various computational systems. An ideal DD has the frequency response given by $\tilde{H}_d(\omega) = j\omega, -\pi \leq \omega \leq \pi$. The most popular FIR design for the differentiators is the minimax relative error (MRE) approximation, proposed by Rabiner and Schafer [1]. The MRE-DDs are, in general, suitable for wideband frequency ranges with moderate accuracy, (say, 100 $\pm$ 1%) but are not flexible enough to be efficiently adopted for limited (low, mid or high) frequency bands. Moreover, the required weighting coefficients for these DDs are computed by using an optimization algorithm [2]. In this chapter, we discuss a new class of digital differentiators, investigated by us, which have maximally linear (ML) frequency response at $\omega = 0$ or $\omega = \pi/2$ or $\omega = \pi$. Correspondingly, this design yields efficient approximations for low, mid or high frequency ranges, giving extremely low relative errors (RE). A universal digital differentiator has also been obtained, which covers the frequency range $0 \leq \omega \leq 0.90\pi$. Mathematical formulas for computing the exact values of the weighting coefficients have also been given.

Keywords Digital filters · Digital differentiators

21.1 Introduction

Digital differentiators (DD) form integral parts of many computational systems and find extensive use in a large number of practical applications, including radars [3, 4], sonars [5], speech processing systems and modulation schemes [6]. For obvious advantages, the DDs are usually realized as Finite Impulse Response (FIR) configurations [7–9], rather than the Infinite Impulse Response (IIR) ones. The literature available on DDs is rather limited. Rabiner and Steiglitz [10] proposed a design of FIR, wideband DDs in 1970, by using an algorithm to minimize the absolute error of the frequency response. However, this design was heavily biased against low frequencies. Rabiner and Schafer [1], in 1974, brought out various properties of FIR-DDs. They designed the differentiators by using the minimax relative error (MRE) as the

Source: Balbir Kumar & S.C. Dutta Roy, "Maximally Linear FIR Digital Differentiators: A Review," *JIETE*, vol 34, pp 347–357, September–October 1988.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_21

criterion of optimization, by making use of McClellan et al. algorithm [2]. Hamming [11], in 1977 suggested a mathematical formula for the weighting coefficients by modifying the Fourier expansion for $\tilde{H}(\omega) = \omega, -\pi \leq \omega \leq \pi$ through Lanczos factor. However, the order, N of the approximation becomes unusually high, if one follows Hamming's method.

The most popular FIR design remains the MRE approximation [1]. However, the MRE-DDs suffer from inflexibility, in the sense that these cannot be efficiently adopted for differentiators required to perform over a limited frequency range at low, mid-band and high frequencies. Moreover, the minimax approach necessarily requires an algorithm [2] for computing the weighting coefficients required in the design. The optimization methods, in general, are very sensitive to finite word length effects, so degradation of frequency response occurs for fixed point implementations. As such, no mathematical formulas are known through which the coefficients could be exactly computed. In this chapter, we present a consolidated review of our work on a different class of design viz. maximally linear DDs (ML-DDs), which are particularly efficient over a limited range at low, mid-band or high frequencies. Mathematical formulas for computing the exact values of the weighting coefficients for the designs have also been given. A novel, efficient structure of DDs for variable frequency range of operation has also been discussed, whereby a differentiator structure designed for odd order N can operate as a 'universal configuration' for all possible odd orders less than or equal to N.

The frequency response of an ideal digital differentiator is given by

$$\tilde{H}_d(\omega) = j\omega \triangleq j\tilde{H}(\omega), -\pi \leq \omega \leq \pi \tag{21.1}$$

where $\tilde{H}(\omega) = \omega$ and is purely real. The curve for $\tilde{H}(\omega)$ is shown in Fig. 21.1. We approximate $\tilde{H}(\omega)$ by using the maximal linearity of frequency response at $\omega = 0$ or $\pi/2$ or π as the optimality criteria of the various frequency bands. The following notations will be used in the sections to follows:

$H_1(\omega)$: DDs for low frequencies, maximally linear at $\omega = 0$.
$H_2(\omega)$: $H_3(\omega)$: DDs for mid-band frequencies, maximally linear at $\omega = \pi/2$.
$H_4(\omega)$: DDs for high frequencies, maximally linear at $\omega = \pi$.
$H_H(\omega)$: digital Hilbert transformer, maximally flat at $\omega = \pi/2$.

The relative error (RE) in the approximations of digital differentiators, say for $H_1(\omega)$, is defined as

$$RE \triangleq |[H_1(\omega) - \tilde{H}(\omega)]/\omega| \tag{21.2}$$

The factor j in (21.1) shows that the differentiators have a constant phase response, equalling $\pi/2$, for all the frequencies. In most of the practical applications, however, the highest accuracy of the magnitude response is more important than the phase response.

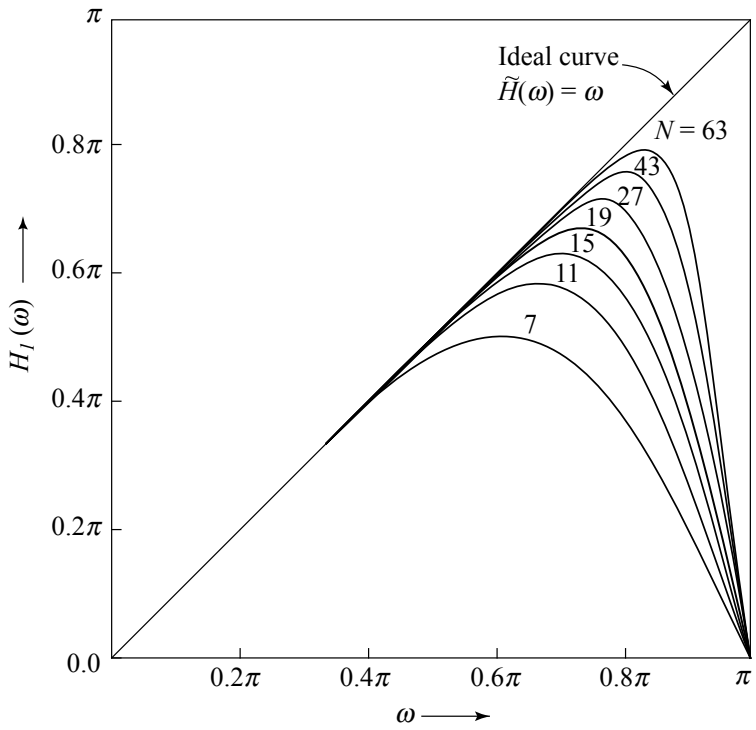


Fig. 21.1 Frequency response, $H_1(\omega)$, of the maximally linear (at $\omega = 0$) digital differentiators for order $N = 7$, 11, 15, 19, 27, 43 and 63

21.2 DDs for Low Frequencies: $H_1(\omega)$

In a host of applications such as autopilotage, autonavigation, autotracking, weapon control systems and in high precision processes [4, 12], differentiation is required to be performed over a limited range in the low-frequency region. Achieving the highest possible accuracy of operation becomes the primary requirement in such cases and as shown in this section, this may be met through the ML-DDs.

21.2.1 Design for Approximation to $H_1(\omega)$

To realize the factor j in Eq. (21.1), we assume the approximation for low frequencies to be of the equation form [1, 13].

$$H_1(\omega) = \sum_{i=1}^{n} d_i \sin i\omega, n = (N-1)/2 \tag{21.3}$$

where a non-recursive configuration has been chosen, and N is the order of the DD, assumed to be odd, for well-known reasons [7]. Maximal linearity of $H_1(\omega)$ at $\omega = 0$ demands:

$$H_1(\omega)|_{\omega=0} = 0 \tag{21.4a}$$

$$dH_1(\omega)/d\omega|_{\omega=0} = 1 \tag{21.4b}$$

$$d^v H_1(\omega)/d\omega^v|_{\omega=0} = 0, v = 2, 3, 4, \ldots, 2n-1 \tag{21.4c}$$

Clearly, $H_1(\omega)$ given by Eq. (21.3) satisfies Eq. (21.4a). Differentiating $H_1(\omega)$, $2n-1$ times successively, and imposing conditions on Eqs. (21.4b) and (21.4c), we get n non-trivial equations. Using the notations $D_i \triangleq id_i$, these equations can be put in the following matrix form:

$$\begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ 1 & 2^2 & 3^2 & \ldots & n^2 \\ 1 & 2^4 & 3^4 & \ldots & n^4 \\ \vdots & \vdots & 3^{2n-2} & & \vdots \\ 1 & 2^{2n-2} & 3^{2n-2} & \ldots & n^{2n-2} \end{bmatrix} \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ \vdots \\ D_n \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{21.5}$$

Equation (21.5) is of the form $[a_{ij}][D_i] = [b_i]$, and can be solved by Crout's method [14]; the solution can be expressed in the following recursive form:

$$d_i = \left\{ (-1)^{i+1} \Big/ \left[i \binom{2i-1}{i-1} \right] \right\} - \sum_{r=1}^{n-i} \binom{2i+r-1}{2i-1} d_{i+r},$$

$$i = n, n-1, n-2, \ldots, 1 \text{(descending order)} \tag{21.6}$$

where

$$\binom{n}{k} \triangleq \frac{n!}{k!(n-k)!} \tag{21.7a}$$

and

$$\binom{0}{k} \triangleq \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases} \tag{21.7b}$$

Following an alternative procedure, it is also possible to derive an explicit formula for $d_i's$ [15].

Table 21.1 The weighting coefficients d_i for maximally linear digital differentiators, $H_1(\omega)$, for low frequencies

n	NF	d_1	$2d_2$	$3d_3$	$4d_4$	$5d_5$	$6d_6$	$7d_7$	$8d_8$	$9d_9$	$10d_{10}$
1	1	1									
2	3	4	− 1			Notes: (i) Normalizing Factor NF $= \binom{2n-1}{n-1}$ (ii) $n = (N - 1)/2$. where N is the order of the filters					
3	10	15	− 6	1							
4	35	56	− 28	8	− 1						
5	126	210	− 120	45	− 10	1					
6	462	792	− 495	220	− 66	12	− 1				
7	1716	3003	− 2002	1001	− 364	91	− 14	1			
8	6435	11,440	− 8008	4368	− 1820	560	120	16	1		
9	24,310	43,758	− 31,824	18,564	− 8568	3060	− 816	153	− 18	1	
10	92,378	167,960	− 125,970	77,520	− 38,760	15,504	− 4845	1140	− 190	20	1

21.2.2 Performance of $H_1(\omega)$

The coefficients d_i calculated by using Eq. (21.6) have been given in Table 21.1. Figure 21.1 shows the curves of $H_1(\omega)$ for selected values of N in the range 7–63. Figure 21.2 gives the relative errors (in decibels) of the approximation $H_1(\omega)$ for $n = 3$–19. As expected, $H_1(\omega)$ achieves extremely high accuracies near $\omega = 0$. In the low-frequency ranges, these DDs are more efficient as compared to the MRE-DDs [1]; for $\omega = 0$ to 0.20π, for example, the minimax design requires an order $N = 127$ to achieve a RE $\leq$ –200 dB, whereas the ML-DDs can accomplish the same with $N = 21$ only. In general, the lower the frequency range of differentiation, the better is the performance of the ML-DDs.

21.3 DDs for Mid-band Frequencies with Zero Phase: $H_2(\omega)$

In many practical cases, it is not necessary to have differentiator approximations to cover the entire band of $\omega = 0$ to π. Almost all the communication systems, for example, demand the highest accuracy over the mid-band frequencies of the spectrum only, since the most vital information is contained therein. Rabiner and Schafer [I] proposed the use of half a sample delay/advance together with an even order N, for the MRE-DDs, to achieve high accuracies. In large signal processing systems, the use of non-integral delays is undesirable. Moreover, choosing an even N deprives the system of possible advantages e.g. accomplishing differentiation through FFT on a

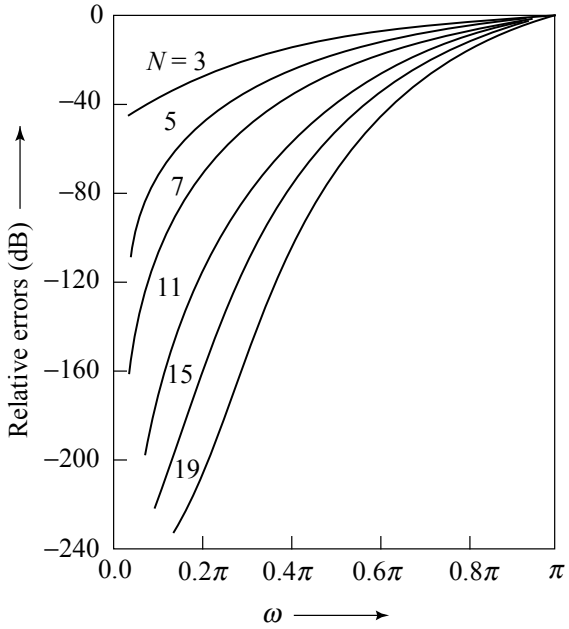


Fig. 21.2 Relative errors (in dB) of $H_1(\omega)$ for order $N =$ 3, 5, 7, 11, 15 and 19

general purpose computer, when N is large. As shown in this section, ML-DDs for mid-band frequencies overcome the aforementioned problems.

21.3.1 Design for Approximation to $\mathbf{H_2(\omega)}$

We aim at a DD approximating the magnitude $|\tilde{H}_d(\omega)| = |\omega|$ in an optimal manner in the mid-band frequency range. Let such an approximation be

$$H_2(\omega) = d_0 + \sum_{i=1}^{n} d_i \cos i\omega, n = (N-1)/2, N \text{ odd} \tag{21.8}$$

Maximal linearity of $H_2(\omega)$ at $\omega = \pi/2$ requires:

$$H_2(\omega)|_{\omega=\pi/2} = \pi/2 \tag{21.9a}$$

$$dH_2(\omega)/d\omega|_{\omega=\pi/2} = 1 \tag{21.9b}$$

$$d^u H_2(\omega)/d\omega^u|_{\omega=\pi/2} = 0, u = 2, 3, 4, \ldots, n \tag{21.9c}$$

lt may be noted that in $(\omega - |\tilde{H}_d(\omega)|)$ plane, $|\tilde{H}_d(\omega)|$ is: (i) periodic in ω with period 2π, (ii) symmetric about the origin (0, 0) and (iii) antisymmetric about the point $(\pi/2, \pi/2)$. We choose an $H_2(\omega)$ which also satisfies these conditions. $H_2(\omega)$ given by Eq. (21.8) is of course symmetric about (0, 0). If we shift the origin of $(\omega, H_2(\omega))$ plane to $(\pi/2, \pi/2)$, and call the new coordinates $(\omega', H_{22}(\omega'))$, then $H_2(\omega) = H_{22}(\omega') + \pi/2$ and $\omega = \omega' + \pi/2$, and we get, from Eq. (21.8)

$$H_{22}(\omega') = (d_0 - \pi/2) + \sum_{i=1}^{n} d_i \cos i(\omega' + \pi/2) \tag{21.10}$$

For $H_{22}(\omega')$ to be antisymmetric function of ω', we must have

$$d_0 = \pi/2; d_i = 0, i = 2, 4, 6, \ldots \tag{21.11}$$

Under the conditions of Eq. (21.11), Eq. (21.8) reduces to

$$H_2(\omega) = \pi/2 + \sum_{\substack{i=1 \\ i \text{ odd}}}^{n} d_i \cos i\omega, n \text{ odd} \tag{21.12}$$

Applying criteria of Eqs. (21.9b) and (21.9c) on Eq. (21.12), and using the notation $B_i = -(2i + l)d_{2i+1}$, $i = 0, 1, 2, \ldots, m$ where $m = (n - 1)/2$, we can write the

resulting equations in the matrix form:

$$\begin{bmatrix} 1 & -1 & 1 & \dots & (-1)^m & (2m+1)^0 \\ 1 & -3^2 & 5^2 & \dots & (-1)^m & (2m+1)^2 \\ 1 & -3^4 & 5^1 & \dots & (-1)^m & (2m+1)^4 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1-3^{2m} & 5^{2m} & \dots & & (-1)^m & (2m+1)^{2m} \end{bmatrix} \begin{bmatrix} B_0 \\ B_1 \\ B_2 \\ \vdots \\ B_m \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{21.13}$$

Solving these equations by Crout's method [14], we finally arrive at the recursive formula [16]:

$$-id_l = \binom{i-2}{(i-1)/2}\left[2^{-(2i-3+\binom{0}{i-1})}\right] + \sum_{k=(i+1)/2}^{(n-1)/2} (-1)^{(2k+i-1)2} \times (2k+1)\binom{(2k+i-1)/2}{(2k-i+1)/2} d_{2k+1}, \tag{21.14}$$

$$i = n, n-2, n-4, \dots, 3, 1 \text{(descending order)}$$

$$n = (N-1)/2;\ n \text{ and } N \text{ odd}$$

Equations (21.11) and (21.14) collectively give all the coefficients d_i, $i = 0, 1, 2, \dots, n$ needed for the approximation $H_2(\omega)$. The weighting coefficients d_i computed by using Eq. (21.14) are shown in Table 21.2 for $n = 1$–15 ($N = 3$–31).

21.3.2 Performance of $\mathbf{H_2(\omega)}$

The frequency response curves for $H_2(\omega)$ are shown in Fig. 21.3 for $n = 1$–17 ($N = 3$–35), and the curves for REs are shown in Fig. 21.4. As expected, the frequency response $H_2(\omega)$ has zero RE at $\omega = \pi/2$ for all N, and extremely low REs are obtainable for narrow mid-band frequencies for reasonably low order N. The ML-DDs, for example, are capable of achieving an RE ≤ -100 dB over the frequency band from 0.45π to 0.55π with just three multiplications (against eight in MRE-DDs) per input sample. These differentiators are more efficient as compared to MRE-DDs for the frequency band $0.15\pi \leq \omega \leq 0.90\pi$ (i.e. over 75% of the complete spectrum). The order N of $H_2(\omega)$ being odd, no non-integral delay is needed.

21.4 DDs for Mid-band Frequencies with $\pi/2$ Phase: $H_3(\omega)$

ln a few typical applications, it is necessary to use a differentiator which realizes the (constant) phase response of $\pi/2$, besides the desired accuracy of the magnitude response $|\tilde{H}_d(\omega)|$. One of the options would be to cascade $H_2(\omega)$ (which gives a zero

Table 21.2 The weighting coefficients d_i for maximally linear digital differentiators, $H_2(\omega)$, for mid-band frequencies (with zero phase response)

$n = (N-1)/2$ (n, N odd)	Normalizing Factor (NF)	d_1	$3d_3$	$5d_5$	$7d_7$	$9d_9$	$11d_{11}$	$13d_{13}$	$15d_{15}$
1	1	– 1					Notes: (i) $d_0 = \pi/2$ for all n (ii) $d_j = 0, j = 2, 4, 6, \ldots$ for all n (iii) $NF \triangleq 2$ Exp $\{(2n-3)\} +$		
3	2^3	– 9	– 1						
5	2^7	– 150	– 25	– 3					
7	2^{11}	– 2450	– 490	– 98	– 10				
9	2^{15}	– 39,690	– 8820	– 2268	– 405	– 35			
11	2^{19}	– 640,332	– 152,460	– 45,738	– 10,890	– 1694	– 126		
13	2^{23}	– 10,306,296	– 2,576,574	– 858,858	– 245,388	– 52,052	– 7098	– 462	
15	2^{27}	– 165,636,900	– 42,942,900	– 15,459,444	– 5,019,300	– 1,301,300	– 245,700	– 29,700	– 1716

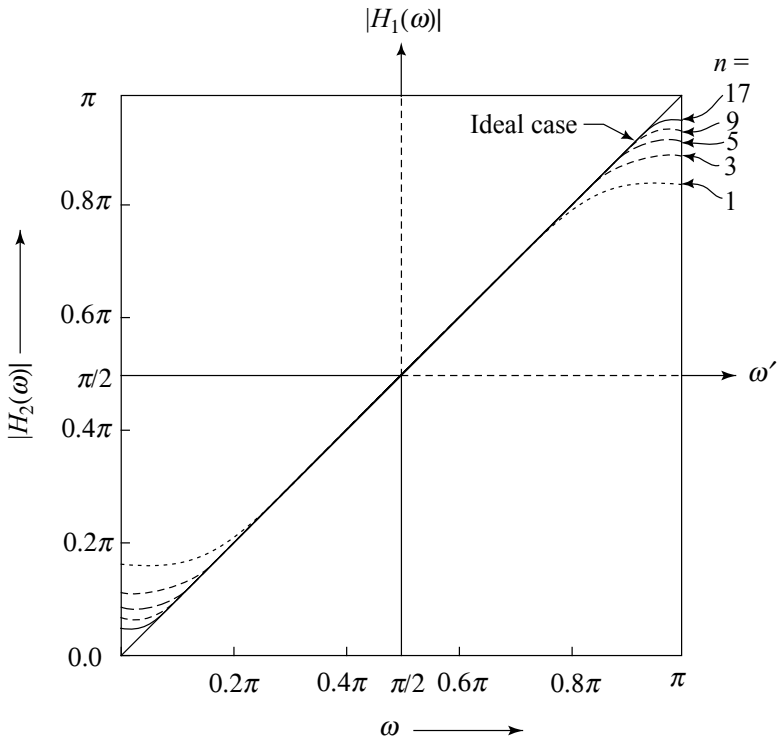


Fig. 21.3 Frequency response, $H_2(\omega)$, of the maximally linear (at $\omega = \pi/2$) digital differentiators for $n = 1, 3, 5, 9$ and 17 (order $N = 2n + 1$). n also corresponds to the tap numbers for the case of variable frequency digital differentiators

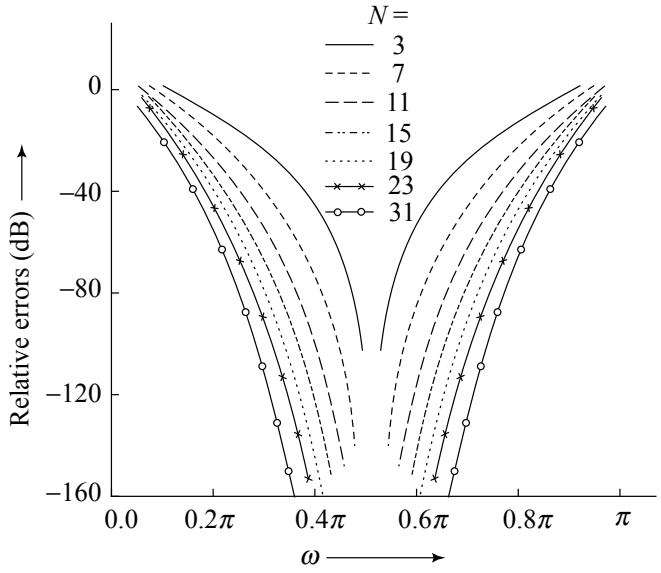


Fig. 21.4 Relative errors (in dB) of $H_2(\omega)$ for order $N =$ 3, 7, 11, 15, 19, 23 and 31

phase response) with a digital Hilbert transformer (DHT). However, it is also feasible to design ML-DDs for mid-band frequency range, meeting the dual requirements of frequency and phase responses [17].

21.4.1 Design for Approximation to $\mathbf{H_3(\omega)}$

To realize the factor j, for $\pi/2$ phase response, in (21.1), $H_3(\omega)$ should have only sine terms [1] i.e.

$$H_3(\omega) = \sum_{i=1}^{n} b_i \sin i\omega,\ n = (N-1)/2,\ N \text{ odd} \tag{21.15}$$

Without loss of generality, we can write Eq. (21.15) as

$$H_3(\omega) = (\pi/2) \sum_{\substack{i=1 \\ i \text{ odd}}}^{n-1} d_i \sin i\omega - (1/2) \sum_{\substack{i=2 \\ i \text{ even}}}^{n} d_i \sin i\omega,\ n \ even \tag{21.16}$$

We have chosen an even value of n, since actual calculations show that it yields a better approximation than that obtainable with an odd n. The criteria of maximal linearity of $H_3(\omega)$ at $\omega = \pi/2$ are similar to those given by Eqs. (21.9a–21.9c). Differentiating $H_3(\omega)$ given by Eq. (21.16), successively, n times and imposing the optimality conditions, we get two sets of linear equations. Solving these, we arrive at the solutions [17]:

$$d_i = \left[\binom{i-2}{(i-1)/2} \Big/ 2^{2i-3+\binom{0}{i-1}} \right] + \sum_{k=(i+1)/2}^{(n-2)/2} (-1)^{(2k+i+1)/2} \times \binom{(2k+i-1)/2}{(2k-i+1)/2} d_{2k+1} \tag{21.17a}$$

$i = n-1, n-3, n-5, \ldots, 5, 3, 1$ (descending order) i odd; n even
and

$$d_i = \left\{ 2/\left[i \binom{i-1}{i/2} \right] \right\} + \sum_{r=1}^{(n-i)/2} (-1)^{r+1} \binom{i+r-1}{r} d_{i+2^r} \tag{21.17b}$$

$i = n, n-2, n-4, \ldots, 6, 4, 2$ (descending order)
i even; n even.

The values of d_i computed by using the formulas of Eqs. (21.17a) and (21.17b) have been given in Table 21.3 for $n = 2$–12 ($N = 5$–25). It may be seen that all the coefficients d_i turn out to be positive by taking $H_3(\omega)$ in the format of Eq. (21.16).

Table 21.3 The weighting coefficients d_i and h_i (i) d_i: for maximally linear digital differentiators, $H_3(\omega)$, for mid-band frequencies (with $\pi/2$ phase response) (ii) h_1: for maximally flat digital Hilbert transformers, $H_H(\omega)$, for mid-band frequencies

n	$NF1$	$-h_1, d_1$	$-h_3, d_3$	$-h_5, d_5$	$-h_7, d_7$	$-h_9, d_9$	$-h_{11}, d_{11}$	$NF2$	d_2	d_4	d_6	d_8	d_{10}	d_{12}
2	1	1		Note 1 NF 1 is normalizing factor for d_i(and h_i), i odd				1		2/2		Note 2 NF 2 is normalizing factor for d_i, i even $\text{NF } 2 = \begin{pmatrix} n-1 \\ n/2-1 \end{pmatrix}$		
4	2^3	9	1					3	8/2	2/4				
6	2^7	150	25	3	$NF1 \triangleq 2\left\lfloor 2n-5+\begin{pmatrix} 0 \\ n-2 \end{pmatrix}\right\rfloor$			10	30/2	12/4	2/6			
8	2^{11}	2450	490	98	10			35	112/2	56/4	16/6	2/8		
10	2^{15}	39,690	8820	2268	405	35		126	420/2	240/4	90/6	20/8	2/10	
12	2^{19}	640,332	152,460	45,738	10,890	1694	126	462	1584/2	990/4	440/6	138/8	24/10	2/12

21.4.2 Performance of $\mathbf{H_3(\omega)}$

The frequency response $H_3(\omega)$ has been shown in Fig. 21.5, and the relative error for these differentiators are shown in Fig. 21.5, for $N = 5$–49. The ML-DDs are again seen to be superior to the MRE-DDs when extremely high accuracies are desired for narrow bands of frequencies around $\omega = \pi/2$. The differentiators $H_3(\omega)$ are particularly efficient over 25% of the mid-band frequency spectrum and may be gainfully exploited in applications such as interferometer radar and phase comparison monopulse radar [3, 4]. If we let

$$H_H(\omega) \triangleq \sum_{\substack{i=1 \\ i \text{ odd}}}^{n-1} h_i \sin i\omega, \quad n \text{ even} \tag{21.18}$$

where

$$h_i = -d_i, \quad i \text{ odd} \tag{21.19}$$

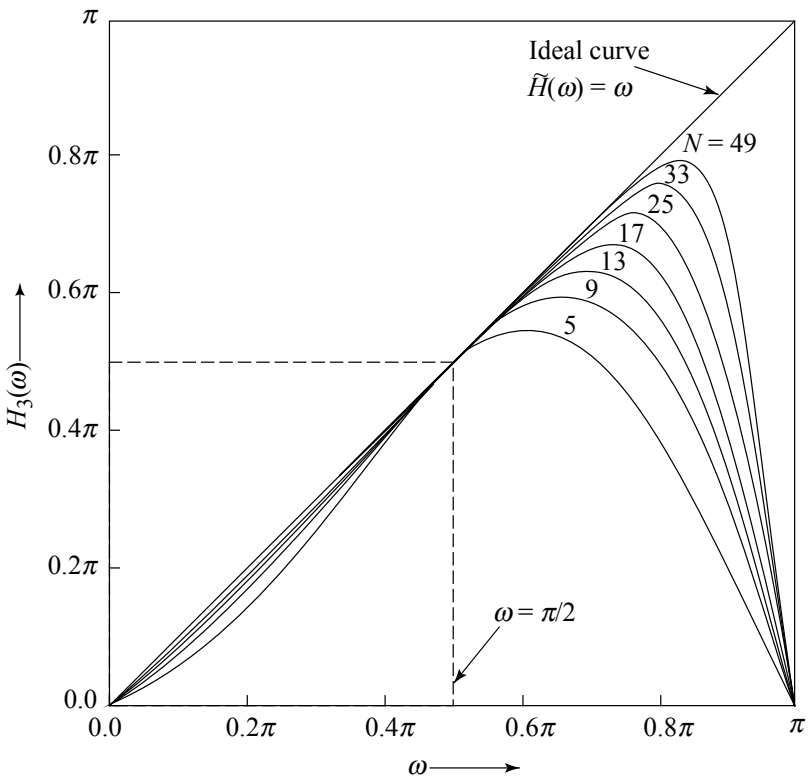


Fig. 21.5 Frequency response, $H_3(\omega)$, of the maximally linear (at $\omega = \pi/2$) digital differentiators, with phase response equalling $\pi/2$, for order $N = 5$, 9, 13, 17, 25, 33 and 49

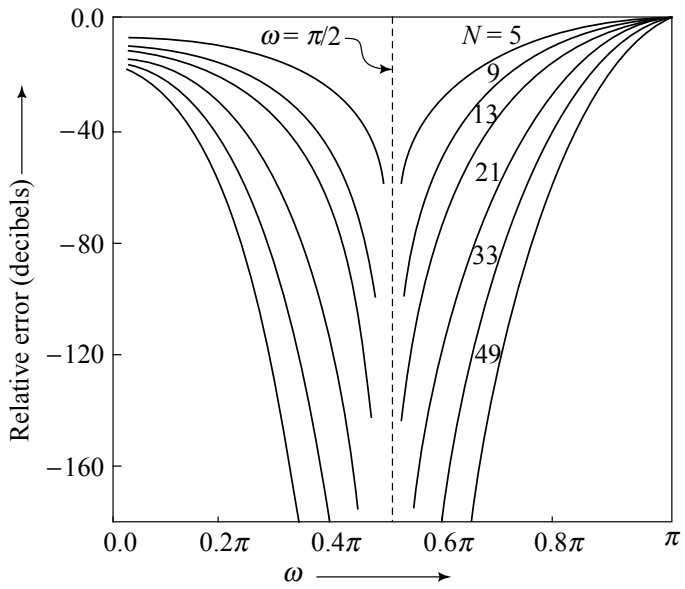


Fig. 21.6 Relative errors (in dB) of $H_3(\omega)$ for order $N = 5, 9, 13, 21, 33$ and 49

and d_i are given by Eq. (21.17a), then it is observed that $H_H(\omega)$ represents DHT's which are maximally flat at $\omega = \pi/2$. The frequency response $H_H(\omega)$ is shown in Fig. 21.7 for values of N from 5 to 49. The cascaded structure '$H_H(\omega) \cdot H_H(\omega)$' gives a differentiator akin to $H_3(\omega)$.

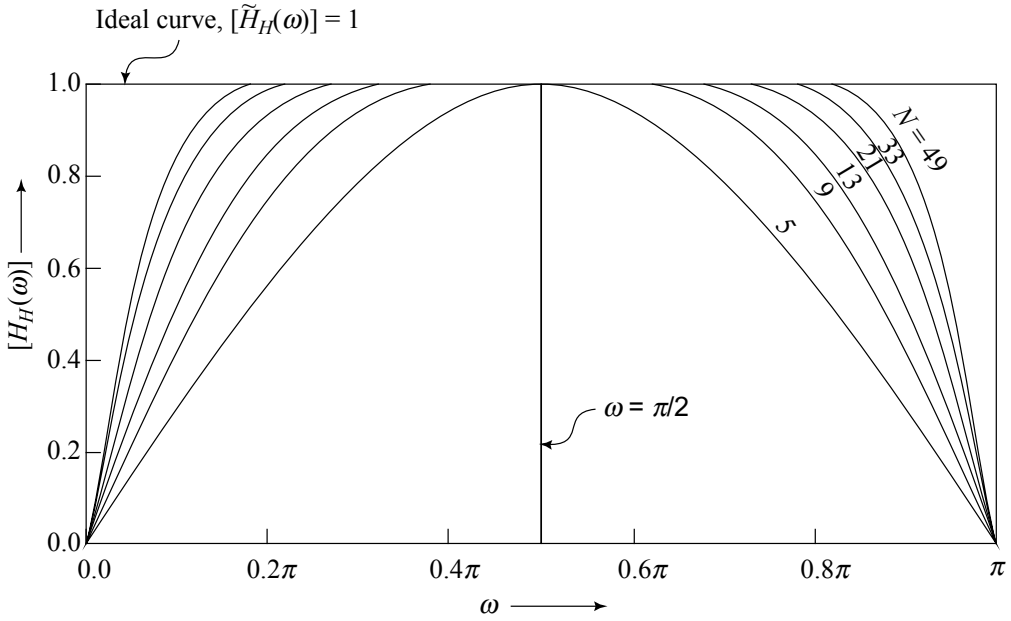


Fig. 21.7 Frequency response, $H_H(\omega)$, of maximally flat (at $\omega = -\pi/2$) digital Hilbert transformers for order $N = 5, 9, 13, 21, 33$ and 49

21.5 DDs for High Frequencies: $H_4(\omega)$

In a number of situations, we need a digital differentiator for high frequencies only. To ensure realization of the factor j in Eq. (21.1), any approximation of $\tilde{H}(\omega)$ must take the form $\sum_{i=1}^{(N-1)/2} \sin i\omega$. Such a summation, of course, is identically zero at $\omega = \pi$, whereas the ideal frequency response $\tilde{H}(\omega)$ demands the highest value (=π) at $\omega = \pi$. To overcome this problem, introduction of half a sample delay (i.e. a factor $z^{-1/2}$) in the structure becomes necessary [1]. For achieving very high accuracy of differentiation for the frequency band $0.5\pi \leq \omega \leq \pi$, the MRE-DDs assume very high order N, whereas the ML-DDs behave more efficiently, as shown in this section.

21.5.1 Design for Approximation to $\mathbf{H_4(\omega)}$

Let the approximation for high frequencies be given by

$$H_4(\omega) = \pi \sum_{i=1}^{n} c_i \sin(i - 1/2)\omega - \sum_{i=1}^{n} d_i \sin i\omega, n = N/2 \tag{21.20}$$

where c_i and d_i are the weighting coefficients and N is the order of the approximation. The terms in the first summation will have additional delays of half sample (i.e. $z^{-1/2}$), when realized. Maximal linearity of $H_4(\omega)$ at $\omega = \pi$ demands:

$$H_4(\omega)|_{\omega=\pi} = \pi \tag{21.21a}$$

$$dH_4(\omega)/d\omega|_{\omega=\pi} = 1 \tag{21.21b}$$

$$d^u H_4(\omega)/d\omega^u|_{\omega=\pi} = 0, u = 2, 3, 4, \ldots, 2n - 1 \tag{21.21c}$$

Forcing conditions given by Eqs. (21.21a–21.21c) on Eq. (21.20) give us two sets of linear equations. Solving these, we finally arrive at the following results [18]:

$$c_i \left[\binom{2i-3}{i-1} \Big/ 2^{\left(4i-5+\binom{0}{i-1}\right)} \right] + \sum_{k=i}^{n-1} (-1)^{k+i} \binom{k+i-1}{k-i+1} c_{k+1}, \tag{21.22a}$$

$i = n, n-1, n-2, \ldots, 2, 1$ (descending order)
and

$$d_i = \left\{ 1 \Big/ \left[i \binom{2i-1}{i-1} \right] \right\} + \sum_{r=1}^{n-i} (-1)^{r-1} \binom{2i+r-1}{2i-1} d_{i+r}, \tag{21.22b}$$

$i = n, n-1, n-2, \ldots, 2, 1$ (descending order)

The values of c_i and d_i, calculated by using Eqs. (21.22a) and (21.22b) have been given in Table 21.4 for $n = 1$–6 ($N = 2$–12).

The transfer function, $G(z)$, corresponding to causal realization of $H_4(\omega)$ is given by

$$\begin{aligned} G(z) &= jz^{-n} H_4(\omega)|_{e^{j\omega}=z} \\ &= (\pi/2)z^{-1/2} \sum_{i=1}^{n} c_i z^{-n+i}(1 - z^{-2i+1}) \\ &\quad - (1/2) \sum_{i=1}^{n} d_i z^{-n+i}(1 - z^{-2i}) \end{aligned} \tag{21.23}$$

A possible structure of $G(z)$, for $n = 3$ ($N = 6$), is shown in Fig. 21.8. We require only one half sample delay ($z^{-1/2}$) to realize $G(z)$, as in the case of MRE-DDs [1] (Fig. 21.6).

21.5.2 Performance of $\mathbf{H_4(\omega)}$

The frequency response of $H_4(\omega)$ is shown in Fig. 21.9 for $n = 1$–10 ($N = 2$–20) and the relative errors are shown in Fig. 21.10. Extremely low REs are indeed achievable near $\omega = \pi$ through the differentiators $H_4(\omega)$, for attractively low orders; for example, with $N = 16$, the ML-DDs can achieve an RE $\leq$ –60 dB for frequency range $\omega = 0.5\pi$ to π, whereas the minimax DDs would require $N \geq 128$ to accomplish a comparable performance. The ML-DDs, $H_4(\omega)$, are therefore, particularly suitable in homomorphic processes [8] and image enhancement applications [19].

21.6 Variable Frequency DDs: Universal Differentiators

It is often of interest to implement, in hardware, a digital differentiator for which the operating range of frequencies can be varied. A number of techniques [20–23] are available to transform a prototype filter to variable cutoff ones. However, all these transformations are applicable only to such filters, in which the ideal frequency response of the prototype filter is piecewise constant.

These cannot, therefore, be meaningfully adopted for achieving variable frequency range digital differentiators. We have obtained a universal canonic structure which can function as a ML-DD at $\omega = 0$ or at $\omega = \pi/2$, and which provides a variable operating frequency range around these values, simply by changing the point at which the output is derived.

Table 21.4 The weighting coefficients (c_i and d_i). *For* maximally linear digital differentiators, $H_4(\omega)$, *for* high frequencies

N	$NF1$ (for c_i)	c_1	c_2	c_3	c_4	c_5	c_6	$NF2$ (for d_i)	d_1	$2d_2$	$3d_3$	$4d_4$	$5d_5$	$6d_6$
1	1	1		$\text{NF}\,1 \triangleq 2\left\lfloor 4n-5+\binom{0}{n-1}\right\rfloor$				1	1			$\text{NF}\,2 \triangleq \binom{2n-1}{n-1}$		
2	2^3	9	1					3	4	1				
3	2^7	150	25	3				10	15	6	1			
4	2^{11}	2450	490	98	10			35	56	28	8	1		
5	2^{15}	39,690	8820	2268	405	35		126	210	120	45	10	1	
6	2^{19}	640,332	152,460	45,738	10,890	1694	126	462	792	495	220	66	12	1

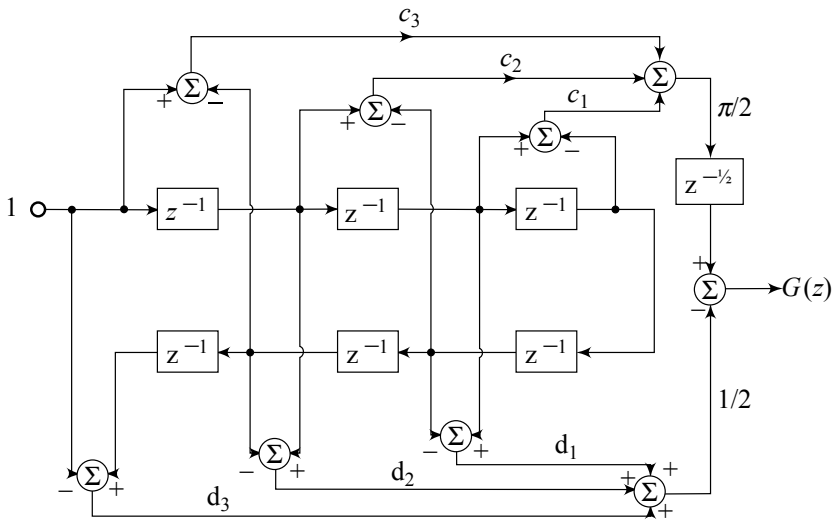


Fig. 21.8 Realization of transfer function $G(z)$ for maximally linear (at $\omega = \pi$) digital differentiators with order $N = 6$

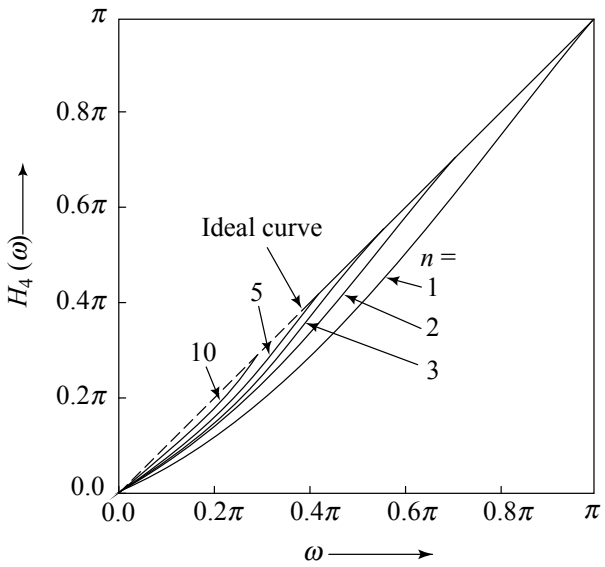


Fig. 21.9 Frequency response, $H_4(\omega)$, of maximally linear (at $\omega = \pi$) digital differentiators for $n = 1, 2, 3, 5$ and 10 (order $N = 2n$)

21.6.1 Design of Universal Digital Differentiators

The DDs with frequency response $H_2(\omega)$, are maximally linear at $\omega = \pi/2$ and give efficient performance in the mid-band frequencies. Now

$$H_2(\omega) = \sum_{i=0}^{n} d_i \cos i\omega, \quad n = (N-1)/2 \tag{21.24}$$

can be expressed as

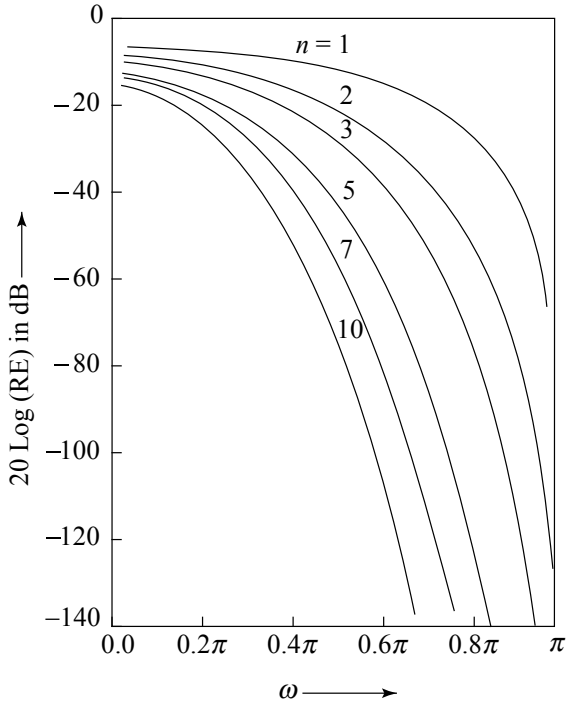


Fig. 21.10 Relative errors (in dB) of $H_4(\omega)$ for $n = 1, 2, 3, 5, 7$ and 10 (order $N = 2n$)

$$H_2(\omega) = \sum_{i=0}^{n} p_i (\cos\omega)^i, \quad n = (N-1)/2 \tag{21.25}$$

where N is the order of the differentiator and d_i's are given by Eqs. (21.11) and (21.14). The coefficients p_i are related to d_i through the coefficients of the Chebyshev polynomials. We can derive an explicit relation between p_i and d_i, provided an explicit formula for d_i (rather than the recursive formula Eq. (21.14)) is available. Following involved calculations and lengthy mathematical manipulations [24], the values of p_i are finally obtained as:

$$\left.\begin{aligned} p_0 &= \pi/2;\ p_1 = -1 \\ p_j &= 0,\ j = 2, 4, 6, \ldots \\ p_i &= -[1, 3, 5, \ldots, (i-2)]^2/i! \\ i &= 3, 5, 7, \ldots \end{aligned}\right] \tag{21.26}$$

It is important to note that the coefficients p_i are independent of n. We exploit this property for realization of variable bandwidth differentiators. The values of the first few p_i's, calculated by using Eq. (21.26) are:

$$\begin{aligned} p_3 = -1/6 \qquad p_5 = -3/40, \qquad p_7 = -5/112 \\ p_9 = -35/1152,\ p_{11} = -63/2816,\ p_{13} = -231/13312, \end{aligned}$$

Now if we change the coordinate axes from $[\omega, H_2(\omega)]$ to $[\omega', H_1(\omega')]$ by using the transformation (see Fig. 21.3 for new axes):

$$\omega = \omega' + \pi/2;\ H_2(\omega) = H_1(\omega') + \pi/2 \tag{21.27}$$

the criterion of maximal linearity would hold at $\omega' = 0$. Using Eqs. (21.26) and (21.27) in Eq. (21.25) and dropping the primes, we get

$$H_1(\omega) = \sum_{\substack{i=1 \\ i \text{ odd}}}^{n} -p_i (\sin\omega)^i, \quad n = (N-1)/2 \tag{21.28}$$

Using Eqs. (21.25) and (21.26), the transfer function of causal, FlR digital differentiators corresponding to $H_2(\omega)$ (i.e. ML-DDs for mid-band frequencies) is given by

$$\begin{aligned} \hat{H}_2(z) &= z^{-n} H_2(\omega)|_{e^{i\omega}=z} \\ &= (\pi/2)2^{-n} + \sum_{\substack{i=1 \\ i \text{ odd}}}^{n} p_i z^{-n+i} \left(\frac{1+z^{-2}}{2}\right)^i \end{aligned} \tag{21.29}$$

Proceeding in a similar manner, the transfer function of causal, FIR digital differentiators for low frequencies is found to be

$$\hat{H}_1(z) = \sum_{\substack{i=1 \\ i \text{ odd}}}^{n} (-1)^{(i+1)/2} p_i z^{-n+i} \left(\frac{1-z^{-2}}{2}\right)^i \tag{21.30}$$

$\hat{H}_1(z)$ and $\hat{H}_2(z)$ can be realized through a composite structure as shown in Fig. 21.11. The outputs from the tappings numbered 1, 3, 5, …, etc. correspond to those of differentiators with $n = 1, 3, 5, \ldots$, respectively.

21.6.2 Performance of Universal Differentiators

The expression for $H_2(\omega)$ in Eqs. (21.24) and (21.25) are mathematically equivalent. Hence the transfer function $\hat{H}_2(z)$ gives ML (at $\omega = \pi/2$)-DDs for mid-band frequencies and the corresponding frequency response is as shown in Fig. 21.3. $\hat{H}_1(z)$ gives maximally linear (at $\omega = 0$)-DDs for low frequencies, as shown in Fig. 21.12. Due to the criteria of maximal linearity, the allied advantages, as brought out in the earlier sections, are also available in the universal DDs. Table 21.5 gives the performance of universal differentiators for REs of |±0.1%|, |±0.5%| and |±1.0%|. This Table may be used to select the appropriate tapping point (i.e. the value of n) to achieve the desired accuracy of differentiation in the specified frequency band from $\omega = 0$ to 0.90π. No change of the weighting coefficients p_i *is* required in the structure. A DD structure of order N can be used for all odd orders from 3 to N. The proposed

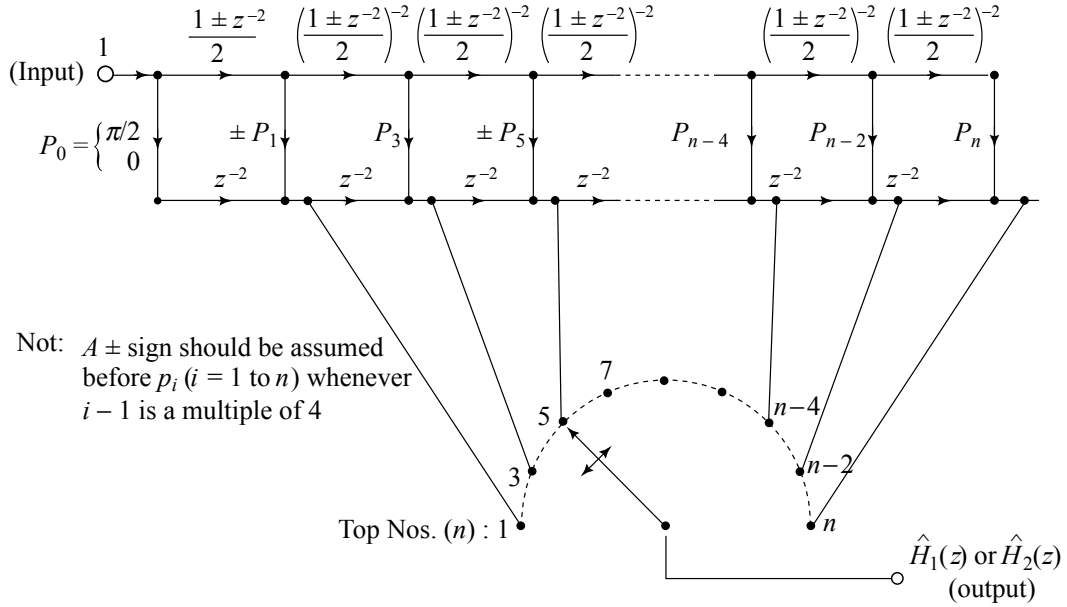


Fig. 21.11 Realization of transfer functions for variable frequency, FIR digital differentiators $\hat{H}_1(z)$: for low frequencies (take $p_0 = 0$ and –ve for $\pm$ signs) $\hat{H}_2(z)$: for midband frequencies (take $p_0 = \pi/2$ and +ve for $\pm$ signs)

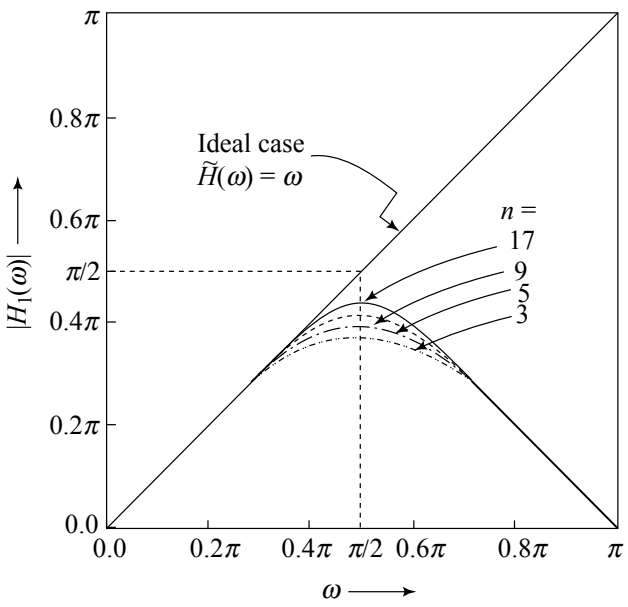


Fig. 21.12 Frequency response, $\hat{H}_2(e^{j\omega})$, of variable frequency digital differentiators for low frequencies at output tap Nos. 3, 5, 9 and 17 (order $N = 2n + 1$)

universal DD would find ready use in various applications in communication and instrumentation such as electronic scanning, beam steering [4] and adaptive image enhancement processes [25], to mention a few.

Table 21.5 Frequency ranges for different relative errors with varying values of *n,* of operation of the universal digital differentiators

$n \downarrow$	Relative error $\rightarrow$					
	(Frequency band)/π for $H_1(\omega)$			(Frequency band)/π for $H_2(\omega)$		
	0.1%	0.5%	1.0%	0.1%	0.5%	1.0%
5	0–0.15	0–0.20	0–0.25	0.35–0.70	0.30–0.75	0.25–0.80
7	0–0.20	0–0.25	0–0.30	0.30–0.75	0.25–0.80	0.25–0.85
9	0–0.25	0–0.30	0–0.32	0.25–0.75	0.22–0.82	0.20–0.85
13	0–0.28	0–0.33	0–0.36	0.23–0.81	0.19–0.86	0.17–0.89
17	0–0.32	0–0.37	0–0.38	0.20–0.83	0.17–0.88	0.14–0.92

21.7 Conclusions

A new class of digital differentiators covering low ($0 \leq \omega \leq 0.5\pi$), mid ($0.15\pi \leq \omega \leq 0.90\pi$) and high ($0.5\pi \leq \omega \leq \pi$) frequency bands of operation has been discussed. A universal differentiator performing over $0 \leq \omega \leq 0.90\pi$ has also been presented. Exact mathematical formulas for computing the weighting coefficients, required in the structures, have been given. The designed differentiators are shown to be highly efficient in their respective frequency bands, and are capable of achieving extremely low relative errors with attractively low orders. The flexibility available in the choice of designs for various situations makes these differentiators particularly suitable for numerous applications.

References

1. L.R. Rabiner, R.W. Schafer, On the behaviour of minimax relative error digital differentiators. Bell Syst. Tech. J. **53**, 333–362 (1974)
2. J.H. McClellan, T.W. Parks, L.R. Rabiner, A computer program for designing optimum FIR linear phase digital filters. IEEE Trans. Audio Electroacoust. AU-**21**, 506–526 (1973)
3. M.I. Skolnik, *Introduction to Radar Systems*, 2nd edn. (McGraw-Hill International Book Co., 1980)
4. M.I. Skolnik, *Radar Handbook* (McGraw-Hill Book Co., New York, 1970)
5. R.J. Urick, *Principles of Underwater Sound* (Mc-Graw-Hill Inc., 1975)
6. R.E. Crochiere, L.R. Rabiner, *Multirate Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, New Jersey, 1983)
7. L.R. Rabiner, B. Gold, *Theory and Applications of Digital Signal Processing* (Prentice-Hall, Englewood Cliffs, New Jersey, 1975)
8. A.V. Oppenheim, R.W. Schafer, *Digital Signal Processing* (Prentice Hall, Englewood Cliffs, New Jersey, 1975)
9. A. Antoniou, *Digital Filter Analysis and Design* (Tata McGraw-Hill Publishing Co, New Delhi, 1983)
10. L.R. Rabiner, K. Steiglitz, The design of wideband recursive and non-recursive digital differentiators. IEEE Trans. Audio Electroacoust. AU-**18**, 204–209 (1970)
11. R.W. Hamming, *Digital Filters* (Prentice-Hall, Englewood Cliffs, New Jersey, 1977)

12. B.C. Kuo, *Digital Control Systems* (Holt, Rinehart and Winston Inc., New York, 1980)
13. B. Kumar, S.C. Dutta Roy, Design of digital differentiators for low frequencies. Proc. IEEE **76**, 287–289 (1988)
14. C. Froberg, *Introduction to Numerical Analysis* (Addison Wesley, Reading, Massachusetts, 1969)
15. B. Kumar, S.C. Dutta Roy, Coefficients of maximally linear, FIR digital differentiators for low frequencies. Electron. Lett. **24**, 563–565 (1988)
16. B. Kumar, S.C. Dutta Roy, Maximally linear FIR digital differentiators for midband frequencies. Int. J. Circuit Theory Appl. **17**, 21–27 (1989)
17. B. Kumar, S.C. Dutta Roy, Design of maximally linear FIR digital differentiators and maximally flat FIR digital Hilbert transformers for midband frequency ranges. Int. J. Circuit Theory Appl. **17**, 483–486 (1989)
18. B. Kumar, S.C Dutta. Roy, Optimal, FIR digital differentiators for high frequencies. IEEE Trans. Circuits Syst. CAS-**36**, 314–318 (1989)
19. L.R. Rabiner, C.M. Rader, *Digital Signal Processing*, Selected Reprint Series under the sponsorship of the IEEE Audio and Electroacoustic Group (IEEE Press, 1972)
20. A.V. Oppenheim, W.F.G. Mecklenbrauker, R.M. Mersereau, Variable cut-off linear phase digital filters. IEEE Trans. Circuits Syst. CAS-**23**, 199–203 (1976)
21. R.E. Crochiere, L.R. Rabiner, On the properties of frequency transformation for variable cut-off linear phase digital filters. IEEE Trans. Circuits Syst. CAS-**23**, 684–686 (1976)
22. S.C. Dutta Roy, S.S. Ahuja, Frequency transformation for linear-phase variable-cut off digital filters. IEEE Trans. Circuits Syst. CAS-**26**, 73–75 (1979)
23. S.N. Hazra, Linear phase bandpass digital filters with variable cut-off frequencies. IEEE Trans. Circuits Syst. CAS-**31**, 661–663 (1984)
24. B. Kumar, S.C. Dutta Roy, Design of universal variable frequency range FlR digital differentiators. Circuit Syst. Signal Process. **11**, 431–439 (1992)
25. A.V. Oppenheim, et al., *Selected Papers in Digital Processing, II*, ed. by Digital Signal Processing Committee, IEEE Acoustics, Speech, and Signal Processing Society (IEEE Press, 1976)

Chapter 22
Design of Maximally Flat and Monotonic FIR Filters Using the Bernstein Polynomial

Abstract This chapter presents a recently introduced technique for the design of Maximally flat (MAXFLAT) FIR filters by using the Bernstein polynomial, and reviews its applications in (i) establishing the equivalence between the earlier known methods, and (ii) formulating a matrix approach for determining the coefficients of MAXFLAT FIR filters efficiently. We also review a new, optimal design procedure for MAXFLAT filters and extend the method to generate monotonic FIR filters with arbitrary magnitude specifications, for which, presently, no method exists. Further, these concepts have been used here to design Quadrature Mirror Filters (QMF), with extremely low reconstruction error.

Keywords Digital filters · Maximally-flat FIR filters · Monotonic FIR filters · Quadrature mirror filters

22.1 Introduction

Linear-phase FIR filters find wide applications where frequency dispersion due to nonlinear phase is harmful, as in speech processing and data transmission. Presently, a variety of design procedures are available for the design of this class of filters, and most of them result in ripples in the passband, as well as the stopband. However, there are some applications, as in the design of filter banks, where we desire monotonic frequency response, to overcome the problem of inter-channel interference. Maximally Flat (MAXFLAT) FIR filters, characterized by flat response in both passband and stopband, are useful in such applications, and has been a topic of interest for quite some time [1–12]. Besides direct applications, MAXFLAT FIR filters have been used to design a class of monotonic FIR filters meeting the prescribed specification of the passband cutoff frequency exactly [13, 14]. Vaidyanathan [15] used this class of filters as building blocks to design filters with monotonic frequency response and narrow transition width. He has further extended the use of MAXFLAT FIR filters to improve the performance of optimal filters [16].

Source: L R Rajagopal & S.C. Dutta Roy, "Design of Maximally-Flat and Monotonic FIR Filters using the Bernstein Polynomial" *JIETE,* vol 34, pp 358–373, September-October 1988.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_22

Herrmann [1], who was the first to propose a design method for this class of filters, arrived at the desired transfer function through the use of general Hermite interpolation. Miller [2] identified the design problem to that of a classical analog maximally flat filter, while Fahmy [3] used a special integral to generate a suitable MAXFLAT polynomial. None of these contributions give a method to determine the order of the filter, an essential parameter required in the design. Kaiser and Reed [6] gave a design method, based on empirical relations, for a specific case of 95 to 5% transition width, which is widely used to design this class of filters. Indeed, even at this time, this is the only method available, and, as yet, no method exists for arbitrary magnitude specifications. This deficiency, in effect, has greatly restricted the use of this class of filters. Besides, although a number of direct and indirect methods [5–13] are available for computing the coefficients of MAXFLAT FIR filters, all of them are involved, and require a large dynamic range of the computing machine.

In this chapter, we start with a review of the work done by us on the design of MAXFLAT filters, and then extend these results to design a new class of monotonic and Quadrature Mirror Filters (QMF). Thus, the aim of this chapter is to enable one to start from the given arbitrary magnitude specifications, and arrive at the impulse response weights required to implement the filter; a concise work of this nature was not so far available in the literature. The chapter is organized as follows. In the next section, we briefly review the existing methods for the design of MAXFLAT and monotonic FIR filters. We introduce the new technique for the design of MAXFLAT filters, and show the equivalence between this and the existing techniques. A matrix approach is also presented for determining the coefficients of MAXFLAT filters. An optimal design procedure, suited for the design of this class of filters with arbitrary magnitude specifications, is also given. Then we present a new technique for the design of FIR filters with monotonic frequency response, which meets the specification of magnitude at the passband and stopband edges exactly. We have shown that, for a set of given specifications, this class of filters yields considerable saving in order of filter, as compared to MAXFLAT filters. Finally, we consider the design of QMFs with monotonic response and low reconstruction errors.

22.2 Review

The transfer function of an FIR filter of order *2N*, expressed in terms of its impulse response coefficients, is

$$H_D(z) = \sum_{i=0}^{2N} h_i \, z^{-1}, \tag{22.1}$$

The phase response of the transfer function $H_D(z)$ could be made linear by constraining the impulse response to be symmetrical or antisymmetrical. However, in the discussions to follow, we shall consider only the symmetrical case, that is,

$$h_i = h_{2N-1},\ i = 0 \text{ to } N-1. \tag{22.2}$$

Although a similar relationship can be obtained for odd order filters, we shall restrict our discussions to only even order filters. The transfer function of a symmetric (linear-phase) FIR filter of order $2N$ is obtained by combining Eqs. (22.1) and (22.2), as

$$H_D(z) = z^{-N} \sum_{i=0}^{N} b_i \frac{(z^i + z^{-i})}{2} = z^{-N} H(z), \tag{22.3}$$

where

$$b_0 = h_0 \text{ and } b_i = 2h_i,\ i = 1 \text{ to } N.$$

The factor z^{-N} in Eq. (22.3) accounts for the linear-phase characteristic, while the zero-phase transfer function $H(z)$ determines the magnitude characteristic. As noted earlier, the symmetric impulse response guarantees linear-phase response and hence we shall be primarily concerned with designing a suitable zero-phase transfer function $H(z)$ meeting the given magnitude specifications. By putting $z = e^{j\omega}$, the frequency response of $H(z)$ can be written in terms of the angular frequency ω, as

$$H(e^{j\omega}) = \sum_{i=0}^{N} b_i \cos(i\omega). \tag{22.4}$$

Specifically for some applications, it is convenient to express the transfer function in the form

$$H(e^{j\omega}) = \sum_{i=0}^{N} a_i (\cos\omega)^i. \tag{22.5}$$

Thus, the design of a linear-phase FIR digital filter starts with the given set of specifications and aims at finding a suitable transfer function $H(z)$, defined by a value of N and a set of coefficients, b_i's or a_i's which meet the specifications exactly or 'to the best possible extent'. In this chapter, we are concerned with the design of suitable MAXFLAT and monotonic FIR filters meeting the given specifications.

A MAXFLAT lowpass (or highpass) filter is characterized by the maximum possible total order of tangencies at either of the extreme points, $\omega = 0$ and $\omega = \pi$. The filter is said to be lowpass and maximally flat if the following conditions are satisfied:

$$H(e^{j\omega})|_{\omega=0} = 1, \tag{22.6a}$$

$$H(e^{j\omega})|_{\omega=\pi} = 0, \tag{22.6b}$$

$$\left.\frac{d^n H(e^{j\omega})}{d\omega^n}\right|_{\omega=0} = 0, \text{ for } n = 1, 2, \ldots, 2L-1, \tag{22.6c}$$

and

$$\left.\frac{d^n H(e^{j\omega})}{d\omega^n}\right|_{\omega=\pi} = 0, \text{ for } n = 1, 2, \ldots, 2K-1, \tag{22.6d}$$

L and K represent the orders of flatness at $\omega = 0$ and $\omega = \pi$ respectively, and are related to N by

$$N = L + K - 1 \tag{22.6e}$$

Herrmann [1] used the transformation

$$\cos\omega = 1 - 2x \tag{22.7}$$

on $H(e^{j\omega})$ of Eq. (22.4) to yield an N-th order polynomial. The coefficients of this polynomial and hence the impulse response weights, are obtained by using the general Hermite interpolation [17]. Besides, Herrmann gave the following empirical relation to determine the order of tangency, L, at $\omega = 0$ required to select one of the N possible magnitude characteristics:

$$L = int(NX_c) + 1. \tag{22.8}$$

where *int* (p) denotes the integer part of p. On the other hand, Miller [2] used the conformal mapping

$$z^{-1} = \frac{1-q}{1+q}, q = \sum + j\Omega, \tag{22.9}$$

on $H(z)$ of Eq. (22.3) to get a rational function. By imposing the necessary derivative conditions and substituting $q = j\Omega$, the resulting analog MAXFLAT transfer function is obtained as

$$H_a(j\Omega) = \frac{\sum_{i=0}^{L} \binom{N}{i} \Omega^{2i}}{(1+\Omega^2)^N}, \tag{22.10}$$

where L is the order of flatness in the passband. Miller then reverted back to the z-plane to obtain the desired MAXFLAT FIR filter transfer function.

Although the methods of Herrmann [1] and Miller [2] give a design approach, they do not give a method to determine the order of the filter required to meet a set of given specifications. Kaiser and Reed [6] gave the empirical relation

$$N = \left[int \geq (\pi/\Delta\omega)^2 \right] \tag{22.11}$$

where $\Delta\omega$ is the transition band as measured by the region, where the filter response magnitude varies from 95% (passband edge) down to 5% (stopband edge). Once the order is estimated, the orders of tangency required at $\omega = 0$ can be determined by using the results of Herrmann [1]. Generally, it has been observed that the filters derived through these steps do not meet the given specifications satisfactorily. Kaiser and Reed [6] proposed a design procedure for a special case of specified 95 to 5% transition width. In this, N is computed by using Eq. (22.11), but in determining L by using Eq. (22.8), the centre frequency of the transition band is used as the parameter, rather than the cutoff frequency, ω_c. Besides, with these initial values of N and L, the filter order is permitted to change by a factor of two over that estimated by using Eq. (22.11), in order to adjust the transition band centre to as close to the prescribed value as possible. Though this method gives satisfactory results over the first, it cannot obviously be used for arbitrary magnitude specifications.

Having estimated the values of N and L, we need to compute the coefficients to realize the transfer function. In the usual situation, we are interested in the coefficients of MAXFLAT filters, expressed in terms of the cosines of multiple angles, i.e. b_i's of Eq. (22.4). Kaiser and Reed [6] proposed a method in which the filter coefficients are determined by first evaluating the frequency response characteristic at N equally spaced set of frequency points, and then performing an inverse discrete Fourier transform on this point set. The results related to MAXFLAT FIR filter in [6] are consolidated in [7]. Gumacos [8] derived a closed form solution for the coefficients of MAXFLAT lowpass filter, having equal passband and stopband widths. Bromba and Zeigler [9] developed a general formula for the coefficients by the use of modified Krawtchouk polynomials, and claimed to have overcome the limitations of [8]. However, it is observed that the results of [9] are advantageous for either high or low passband and stopband width ratios. Jinaga and Dutta Roy proposed two methods—an explicit expression [10] and a recurrence relation [11]—for computing the coefficients directly, in contrast to Kaiser's indirect method. Both of these methods require a large dynamic range of the computing machine.

Fahmy [3] has expressed the coefficients a_i's of Eq. (22.5) in terms of N and L; however the formula is computationally expensive. Explicit expressions for the coefficients a_i's have also been reported by Jinaga and Dutta Roy [12]. The method of derivation was based on Kaiser-Hamming amplitude change function, and is neither straight forward nor simple. Thus, it can be concluded that the existing methods are neither simplistic in their approach, nor computationally efficient.

Generally, in the filter design problem, we are interested in meeting the passband edge specifications exactly. However, in the design procedure outlined earlier, more often, this condition is not met. One of the solutions to this problem is to design suboptimal filters, which meet the requirement of passband edge exactly. Thajchayapong

et al. [13] obtained a transitional filter between two adjacent values of L, say L_1 and L_2 with $L_2 = L_1 + 1$, by adding an extra term in the numerator of (22.10) to get

$$H_0(j\Omega) = \frac{\sum_{i=0}^{L} \binom{N}{i} \Omega^{2i} + v\Omega^{2L_2}}{(1+\Omega^2)^N}, \tag{22.12}$$

If in Eq. (22.12), the value of v is allowed to vary in the range $0 \leq v \leq \binom{N}{L_2}$, infinite number of transitional filters whose magnitude responses lie between those with flatness parameters L_1 and L_2 are obtained. Thus the coefficient v in Eq. (22.12) can be chosen to meet the magnitude specification at any one desired point and the authors of [13] have specifically considered the case of 3 dB cutoff.

A different method to generate this class of filters was reported in [14], where the authors worked directly on the digital transfer function. They considered two transfer functions $H_1(z)$ and $H_2(z)$, orders of flatness at $\omega = 0$ being L_1 and $L_1 + 1$, respectively. They generate a class of filters $H_0(z)$ by considering a weighted linear combination of $H_1(z)$ and $H_2(z)$ as,

$$H_0(z) = \frac{H_1(z) + \beta H_2(z)}{1+\beta}, \tag{22.13}$$

where the weight β is adjusted to make the magnitude response pass through a prescribed cutoff frequency. Later, in this chapter, we shall prove that, in this class of filters, the magnitude characteristic is forced to pass through an arbitrary cutoff frequency at the cost of one degree of flatness. Thus, these transitional filters of [13, 14], although claimed as MAXFLAT, are really suboptimal; hence the problem of designing MAXFLAT FIR filters meeting arbitrary specifications is yet an unsolved one.

The theory of MAXFLAT FIR filters has been extended to design other class of digital filters like the MAXFLAT QMF [18]. QMFs are largely used in the field of speech transmission/reception, where the signal is split into two or more sub-bands at the transmitter and recombined at the receiver with minimum aliasing error. Two sub-band coding is achieved by a pair of orthogonal lowpass and highpass filters. Bergeron [18] proposed a numerical method for the design of MAXFLAT QMF, based on the work of Herrmann [1]. In brief, the method aims at searching all the orders of filters over the entire range of L and pick out those characteristics, which approximate a pair of orthogonal filters, with fairly low aliasing errors. He has left the optimal search procedure as an open problem.

22.3 The Bernstein Polynomial and Design of Maxflat Filters

Let $f(x)$ be a function defined in the interval (0, 1). The N-th ($N \geq 1$) order Bernstein polynomial for $f(x)$ is given by [17] and [19]:

$$B_N(x) = \sum_{i=0}^{N} f\left(\frac{i}{N}\right)\binom{N}{i}x^i(1-x)^{N-1}. \tag{22.14}$$

Here, the interval (0, 1) is divided into N equal intervals and the values $f(x)$ at these $(N + 1)$ points only are used to evaluate the polynomial $B_N(x)$.

The Bernstein polynomial has some very interesting properties and has been extensively studied by Lorentz [19]; some of the properties of this polynomial in real domain are dealt with by Davis in [17]. A few of the important properties that are of interest in the present context of design of MAXFLAT FIR filters are:

(i) At the extreme points, the approximating polynomial is exactly equal to the value of the function, that is,

$$B_N(0) = f(0) \text{ and } B_N(1) = f(1) \tag{22.15}$$

(ii) This polynomial yields smooth approximants, in contrast to the well-known Chebyshev approximation, which is characterized by ripple behaviour.

(iii) This polynomial provides simultaneous approximations of the function and its derivatives, that is, if the p-th derivative of $f(x)$ exists, then,

$$\lim_{N\to\infty} B_N^{(p)}(x) = f^{(p)}(x), \tag{22.16}$$

where $f^{(p)}$ denotes the p-th derivative of the function f.

Thus the Bernstein polynomial 'mimics the behaviour of the function to a remarkable degree' [17]. In fact, it is the statement of Davis, which motivated us to study the possibility of using this polynomial to design MAXFLAT FIR filters. Although this polynomial has extremely good approximation properties, the disadvantage is that it has very slow convergence, and this fact has precluded any numerical application of this polynomial. However, as we shall demonstrate, the use of this polynomial has added a new dimension to the design of MAXFLAT and monotonic filters, by generating numerous interesting and useful results.

Consider the approximation problem of a lowpass function, shown in Fig. 22.1, by using the Bernstein polynomial. If L is the number of successive discrete points at which the function is unit valued, then

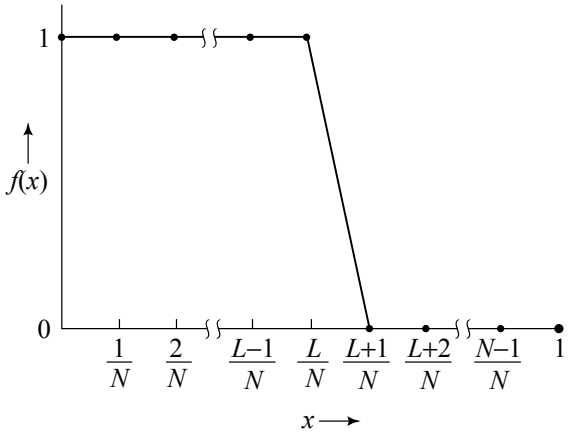


Fig. 22.1 Lowpass function

$$f\left(\frac{i}{N}\right) = \begin{cases} 1, & 0 \le i \le L \\ 0, & L+1 \le i \le N \end{cases}. \tag{22.17}$$

Substituting the functional values of Eq. (22.17) in Eq. (22.14), and by using an additional subscript L for $B_N(x)$, we get

$$B_{N,L}(x) = \sum_{i=0}^{L} \binom{N}{i} x^i (1-x)^{N-i} \tag{22.18}$$

In [20], it is shown that, by using suitable transformations, the expression given in Eq. (22.18) reduces to those obtained by Herrmann [1] and Miller [2] for a MAXFLAT filter with $(L+1)$ as the degree of tangency at $\omega = 0$, thus establishing an equivalence between the three techniques. Further, as noted in [20], this new approach has led to an analytical assertion of the empirical relation of Eq. (22.8).

22.4 Transformation Matrix

In the previous Section, it was shown that the MAXFLAT lowpass FIR filters can be easily designed by using a functional vector, whose entries are either 1 or 0, where the number of 1s and 0s determine the degrees of flatness at $\omega = 0$ and $\omega = \pi$ respectively. In this Section, we shall relate this simple functional vector and the coefficients of MAXFLAT filter, either a_i's of Eq. (22.5) or b_i's of Eq. (22.4), by a transformation matrix.

22.4.1 *Coefficients* $\mathbf{a_i}$*'s*

Let the coefficient vector $[a] = [a_0, a_1, \ldots, a_N]^t$ be related to the functional vector $[f] = [f(0), f(1/N), f(2/N), \ldots, f(N/N)]^t$ via a constant matrix $[P_N]$ of order $(N + 1)$, such that

$$[a] = \alpha[P_N][f], \tag{22.19}$$

where α is a normalizing constant, so chosen that all the elements of $[P_N]$ are integers. We are interested in finding the entries of the matrix $[P_N]$.

Starting with the explicit formula for the coefficient given in [12], we arrive at a simple matrix relation [21]:

$$[P_N] = [Q_N][D_N] \text{and} \, \alpha = 2^{-N} \tag{22.20}$$

where $[Q_N]$ is the matrix used in the bilinear transformation [22–24], and $[D_N]$ is a diagonal matrix with nonzero elements, equal to $\binom{N}{m}$ for $m = 0, 1, \ldots, N$. It is interesting to note that the elements of $[D_N]$ are the same as the elements of the first column of $[Q_N]$, and hence need not be computed explicitly. Besides, in [22, 23], it has been shown that the knowledge of the elements in the first column is sufficient to generate all the other entries of $[Q_N]$. The simple relation, repeated here for the sake of completeness, is

$$\begin{aligned} Q_{0,j} &= 1 \; j = 0 \text{ to } N \\ Q_{i,j} &= Q_{i,j-1} - Q_{i-1,j-1} - Q_{i-1,j} \quad i = 1 \text{ to } N \\ &\qquad\qquad\qquad\qquad\quad j = 1 \text{ to } N. \end{aligned} \tag{22.21}$$

By observing the symmetry properties of $[Q_N]$, only about half the elements need to be computed. The simplicity of this approach would soon be illustrated by an example.

22.4.2 *Coefficients* $\mathbf{b_i}$*'s*

Here, the problem can be formulated in terms of a transformation matrix $[R_N]$ as

$$[b] = \beta[R_N][f] \tag{22.22}$$

where β is a normalizing constant. It is known that the b_i's of Eq. (22.4) are related to a_i's of Eq. (22.5) through Chebyshev polynomials [25]. Taking advantage of this observation and the earlier results of this Section, we have

$$[R_N] = [C_N][P_N] \tag{22.23a}$$

and

$$\beta = 2^{2N-1}, Z \tag{22.23b}$$

where $[C_N]$ is the Chebyshev matrix, which is an upper triangular one, with merely one-fourth nonzero elements, and hence its pre-multiplication with $[P_N]$ can be done efficiently. Further, $[C_N]$ can be decomposed into product of two simpler matrices:

$$[C_N] = [\hat{C}_N][\hat{D}_N], \tag{22.24}$$

where the elements of $[\hat{C}_N]$ are derived from the binomial coefficients, and $[\hat{D}_N]$ is a diagonal matrix whose nonzero elements are expressible as powers of 2. A simple method to generate the matrices in Eqs. (22.20) and (22.24) is given in [26].

Example 1 Consider the case when $N = 4$, then

$$[Q_4] = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 4 & 2 & 0 & -2 & -4 \\ 6 & 0 & -2 & 0 & 6 \\ 4 & -2 & 0 & 2 & -4 \\ 1 & -1 & 1 & -1 & 1 \end{bmatrix}, \quad [D_4] = \begin{bmatrix} 1 & & & & \\ & 4 & & & \\ & & 6 & & \\ & & & 4 & \\ & & & & 1 \end{bmatrix}$$

$$[\hat{C}_4] = \begin{bmatrix} 1 & 0 & 1 & 0 & 3 \\ 0 & 1 & 0 & 3 & 0 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad [\hat{D}_4] = \begin{bmatrix} 8 & & & & \\ & 8 & & & \\ & & 4 & & \\ & & & 2 & \\ & & & & 1 \end{bmatrix} \tag{22.25}$$

By multiplying $[Q_4]$ and $[\hat{D}_4]$, we get $[P_4]$ as

$$[P_4] = \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 8 & 0 & -8 & -4 \\ 6 & 0 & -12 & 0 & 6 \\ 4 & -8 & 0 & 8 & -4 \\ 1 & -4 & 6 & -4 & 1 \end{bmatrix}, \tag{22.26}$$

Pre-multiplying $[P_4]$ in Eq. (22.26), with $[C_4]$ and $[D_4]$ of (22.25), we get $[R_4]$ as

$$[R_4] = \begin{bmatrix} 35 & 20 & 18 & 20 & 35 \\ 56 & 16 & 0 & -16 & -56 \\ 28 & -16 & -24 & -16 & 28 \\ 8 & -16 & 0 & 16 & -8 \\ 1 & -4 & 6 & -4 & 1 \end{bmatrix}, \tag{22.27}$$

For $L = 1$, by using Eq. (22.26), we get $[a] = [5, 12, 6, -4, -3]^t$ with $\alpha = 16$ and by using Eq. (22.25), we get $[b] = [55, 72, 12, -8, -3]^t$ with $\beta = 128$.

A recursive method to generate the transformation matrix $[R_N]$ is given in [26]; the advantage of this method is that it does not require all the matrices in Eq. (22.25) to be computed independently. Thus, with the results of this Section and those in [26], given a functional vector corresponding to the desired N and L of MAXFLAT FIR filter, we can generate the required coefficient vector. However, one needs to address a more general problem, that is, for a given set of specifications, 'what are the optimal values of N and L?'. In the next Section, we discuss this problem and propose a solution.

22.5 Maxflat FIR Filters

Consider the design of a lowpass linear-phase MAXFLAT FIR filter transfer function $H(z)$ to satisfy the following arbitrary magnitude specifications (Fig. 22.2):

$$H(e^{j\omega}) \begin{cases} > \delta_1, & 0 \le \omega < \omega_c, \\ = \delta_1, & \omega = \omega_c, \\ \le \delta_2, & \omega_s \le \omega \le \pi \end{cases} \tag{22.28a, b, c}$$

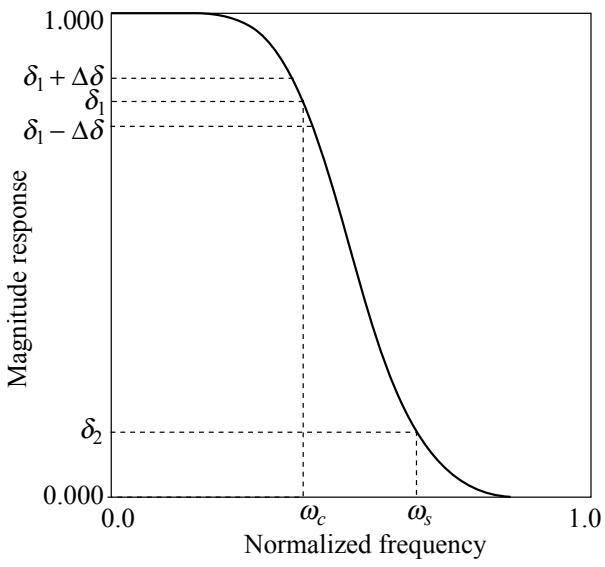


Fig. 22.2 A typical MAXFLAT FIR lowpass filter specifications

Besides these, the filter is, of course, required to meet the constraints of Eq. (22.6a, 22.6b, 22.6c, 22.6d, 22.6e). As mentioned earlier, for a given N-th order MAXFLAT filter, there are only N possible characteristics, and we need to choose one amongst these, parameterized by L, which meets the specifications given in Eq. (22.28a, b, c). Generally, it has been observed that the specification of Eq. (22.28b) cannot be met exactly with nominal order of filter. In view of this, we allow a magnitude tolerance of $\Delta\delta$ at the passband edge ω_c; consequently, Eq. (22.28b) is modified as

$$\left|H\left(e^{j\omega}\right) - \delta_1\right| \leq \Delta\delta, \omega = \omega_c. \tag{22.28d}$$

Now the problem of design of MAXFLAT filter is the same as choosing a MAXFLAT characteristic, parameterized by N and L, which satisfies the conditions of Eq. (22.28a, b, c).

The proposed method is based on the use of Bernstein polynomial for the design of MAXFLAT FIR filters. Now, the filter design problem is reduced to identifying a suitable Bernstein polynomial, which satisfies the equivalent conditions of Eq. (22.28a, b, c). The optimal method, proposed for this purpose, uses the following set of recurrence relations [27–29]:

$$B_{N+1,\,L}(x) = B_{N,\,L}(x) - \binom{N}{L} \times b_{N,\,L}(x) \tag{22.29a}$$

and

$$B_{N+1,L+1}(x) = B_{N,L}(x) - \binom{N}{L+1} \times b_{N,L}(x), \tag{22.29b}$$

where the auxiliary polynomial $b_{N,L}(x)$ satisfies

$$b_{N,L}(x) = (1-x)b_{N-1,L}(x) \tag{22.29c}$$

We start the search procedure with second-order Bernstein polynomial, and search for the optimum order of filter and the orders of tangency at $\omega = 0$ and $\omega = \pi$, simultaneously, in a 'stair case' like manner. The procedure is depicted in the flow-diagram of Fig. 22.3. This being an optimal method, it results in the minimum order of the filter required to meet the given specifications, and a detailed comparison with the earlier methods establishes the superiority of this approach. The details of this method and exhaustive comparative study based on examples is discussed in [28, 29].

Generally, in the filter design problem, we are interested in meeting the passband edge specifications exactly. However, in the design procedure outlined in this Section, we need to accept a tolerance of $\Delta\delta$ at the passband edge and more often, if this value is small, then the required N is very high. One of the solutions to this problem is to design suboptimal filters, which meet the requirement of passband and stopband edges exactly.

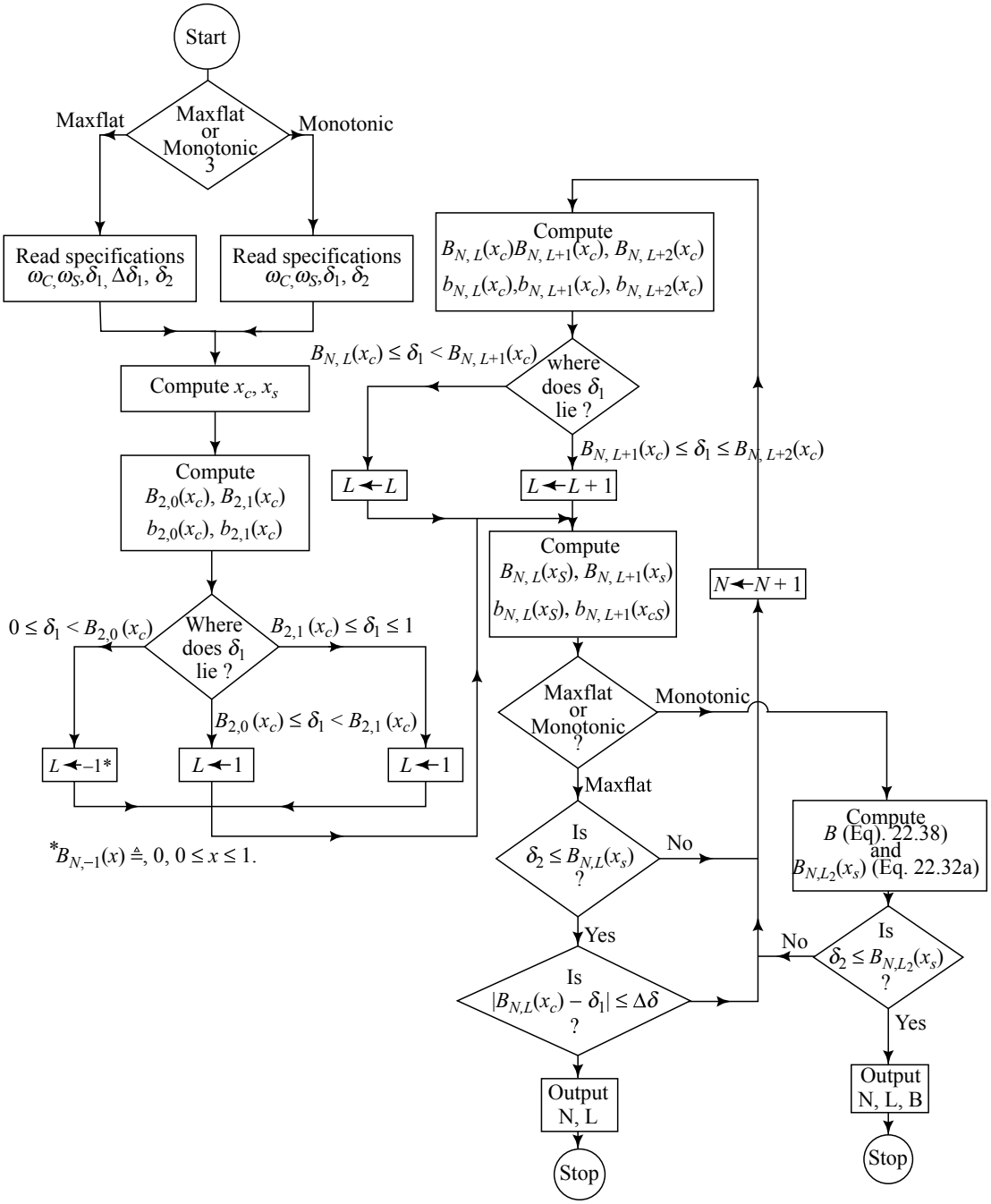


Fig. 22.3 Flow diagram depicting the procedure for designing MAXFLAT and monotonic FIR Iowpass filters

22.6 Monotonic FIR Filters

As is clear from earlier discussions, the design of MAXFLAT 1owpass filter evolves from the Bernstein polynomial approximation of an ideal 1owpass function, with transition width 1/*N* (Fig. 22.1) and obviously, we can not exercise control over the behaviour of the characteristic in the transition region. However, it is possible to

overcome this difficulty by increasing the ideal transition width of the approximating function to either $2/N$ or $3/N$, thus, resulting in a new class of monotonic filters.

22.6.1 Monotonic Filters with Transition Width 2/N

Consider the approximation problem of a lowpass function, defined in the interval (0, 1), as shown in Fig. 22.4. Here,

$$f\left(\frac{i}{N}\right) = \begin{cases} 1, & 0 \le i \le L \\ \alpha, & i = L+1, 0 < \alpha < 1 \\ 0, & L+2 \le i \le N. \end{cases} \tag{22.30}$$

The Bernstein polynomial for this lowpass function, obtained by using the functional values of Eq. (22.30) in Eq. (22.14), and is given by

$$B_{N,L_2}(x) = \sum_{i=0}^{L} \binom{N}{i} x^i (1-x)^{N-i} + \alpha \binom{N}{L+1} x^{L+1} (1-x)^{N-L-1} \tag{22.31}$$

where the suffix L_2 has been added to indicate an ideal transition width of $2/N$. Here, we have a parameter α, which can be suitably selected to control the behaviour of the response in the transition region as well. Simplifying Eq. (22.31), we get

$$B_{N,L_2}(x) = \frac{B_{N,L}(x) + \beta B_{N,L+1}(x)}{1+\beta}, \tag{22.32a}$$

where

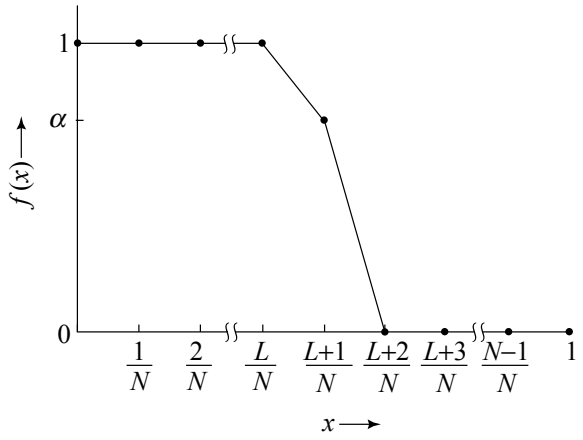


Fig. 22.4 Lowpass function with transition width $2/N$

$$\beta = \frac{\alpha}{1-\alpha} \tag{22.32b}$$

The order of flatness at the extreme points, $x = 0$ and $x = 1$, found by the simple method given in [20], are L and N–L–2, respectively; a maximum degree of tangency $(N - 2)$ only, in contrast to $(N - 1)$ for MAXFLAT function. Hence, although this approximating polynomial is monotonic, it is not maximally flat, because the total degree of flatness is one less than the maximum possible.

Now we shall show that two of the previous works, [13] and [14], are special cases of this class of filters, and thus establish a link amongst the three methods. Using the conformal mapping given by Eqs. (22.9), (22.31) simplifies to

$$H_0(\Omega) = \frac{\sum_{i=0}^{L} \binom{N}{i} \Omega^{2i} + v\Omega^{2(L+1)}}{(1+\Omega^2)^N} \tag{22.33a}$$

where

$$v = \alpha \binom{N}{L+1}. \tag{22.33b}$$

It is seen that Eq. (22.33a) is identical to the result in [13]. Now, it is obvious that the 'transitional maximally flat filters' of [13] are not strictly maximally flat.

Transforming equation (22.32a) to the z-domain by using the inverse-transformation of Eq. (22.7), we get,

$$H_0(z) = \frac{H_{N,L}(z) + \beta H_{N,L+1}(z)}{1+\beta}. \tag{22.34}$$

This result is identical to the one in [14]. Both the methods discussed so far have been used to generate monotonic filters with specified cutoff frequency, by a suitable choice of v and β. Though v or β can be computed for arbitrary magnitude at the cutoff frequency, the choice of N and L need to be known before-hand. Since no method exists for estimating N and L for arbitrary magnitude specifications, the authors of [13] and [14] have restricted the use of this class of filters to meet only a specified 3 dB cutoff frequency. Before we discuss the design procedure for this class of filters, let us extend this concept to generate another class of filters with monotonic response, but with both the passband and stopband edge specifications being met exactly. The special interest in this class of filters would become evident from the requirements of the next Section on QMF.

22.6.2 Monotonic Filters with Transition Width 3/N

Consider the approximation of a lowpass function with an ideal transition width 3/*N*, as shown in Fig. 22.5. The Bernstein polynomial for this lowpass function is given by

$$\begin{aligned} B_{N,L_3}(x) = &\sum_{i=0}^{L} \binom{N}{i} x^i (1-x)^{N-i} \\ &+ \alpha_1 \binom{N}{L+1} x^{L+1} (1-x)^{N-L-1} \\ &+ \alpha_2 \binom{N}{L+2} x^{L+2} (1-x)^{N-L-2}, \end{aligned} \tag{22.35}$$

where L_3 has been used to indicate the ideal transition width of 3/*N*. After simplification, we get,

$$B_{N,L_3}(x) = \frac{B_{N,L}(x) + \beta_1 B_{N,L+1}(x) + \beta_2 B_{N,L+2}(x)}{1 + \beta_1 + \beta_2}, \tag{22.36}$$

where α_i's and β_i's are related by

$$\beta_1 = \frac{\alpha_1 - \alpha_2}{1 - \alpha_1} \quad \beta_2 = \frac{\alpha_2}{1 - \alpha_1} \tag{22.37}$$

Here, for $B_{N,L3}(x)$ to be monotonic, the conditions $0 \le \alpha_2 \le \alpha_1 \le 1$ need to be satisfied, and this requires both β_1 and β_2 to be positive.

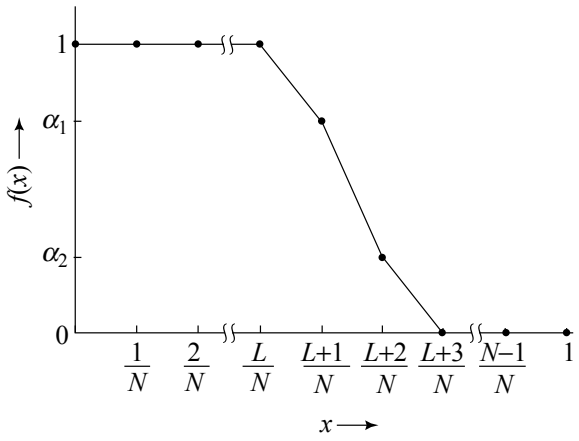


Fig. 22.5 Lowpass function with transition width 3/*N*

22.6.3 The New Method

First, we discuss the design procedure for generating monotonic filters meeting only the passband cutoff frequency exactly. The procedure is conceptually similar to that used for designing optimal MAXFLAT filters [27–29]. We start the search procedure with all the possible second-order Bernstein polynomials, and evaluate the magnitudes at $x = x_c$ for all of them. The two values of L are chosen for which the desired magnitude δ_1 lies in one of the following ranges $\{0, B_{2,0}(x_c)\}$, $\{B_{2,0}(x_c), B_{2,1}(x_c)\}$, $\{B_{2,1}(x_c), B_{2,2}(x_c)\}$. The value of β is computed by using the relation

$$\beta = \frac{\delta_1 - B_{2,L}(x_c)}{B_{2,L+1}(x_c) - \delta_1}. \tag{22.38}$$

With this value of β, α is computed from Eqs. (22.32a, 22.32b) and is used to obtain the required polynomial $B_{2,L2}(x)$ satisfying the magnitude specifications at the passband cutoff frequency exactly. It should be pointed out that $B_{2,L2}(x)$, which is derived through linear combinations of $B_{2,L}(x)$ and $B_{2,L+1}(x)$, will be monotonic if and only if β is positive (which is satisfied if $0 \leq \alpha \leq 1$). It is obvious that in this design, this is always guaranteed. Though the magnitude at the passband cutoff frequency is met exactly, the stopband specifications may not be met. Hence, the magnitude at $x = x_s$ is computed for those two values of L which are of interest. Using the values of β, $B_{2,L}(x_s)$ and $B_{2,L+1}(x_s)$, we compute the magnitude at x_s by using Eq. (22.32a). If this magnitude is less than or equal to δ_2, then the desired conditions have been met. If not, the order of the polynomial is incremented by one and a set of three values of L are chosen so that δ_1 lies in either of the following two ranges $\{B_{3,L}(x_c), B_{3,L+1}(x_c)\}$, $\{B_{3,L+1}(x_c), B_{3,L+2}(x_c)\}$. The three values of magnitude at $x = x_c$ are computed by using the recurrence relations (22.29a, 22.29b, 22.29c), and two values of L are chosen, corresponding to the range which includes δ_1. Now we compute β, and hence the magnitude at $x = x_s$, and check if the stopband conditions are satisfied. If not, the search procedure continues with the higher order of polynomial till the desired conditions are met. The design procedure is depicted in the flow-diagram shown in Fig. 22.3. We now consider an example to illustrate the design procedure.

Example 2 Consider the design of a monotonic FIR filter with the following specifications: $\omega_c = 1.047198$, $\omega_s = 2.094395$, $\delta_1 = 0.707$ and $\delta_2 = 0.04$. The transformed passband and stopband edges, in the interval (0, 1), are 0.25 and 0.75 respectively. The estimated order of optimal MAXFLAT filter is 11, that too, with an error of $\Delta\delta = 0.00631$ at the passband cutoff frequency. However, by the method proposed here, the same specifications are met with a sixth order filter. The intermediate results obtained from various steps are shown in Table 22.1. The three frequency responses, $H_{6,2}(e^{j\omega})$, $H_{6,3}(e^{j\omega})$ and $H_{6,2_2}(e^{j\omega})$ are shown in Fig. 22.6. Here, $H_{6,2_2}(e^{j\omega})$ denotes the linear combination of two transfer functions of same order ($N = 6$), but with different degrees of tangency at $\omega = 0$, that is, $L = 2$ and $L = 3$. It should be noted that the value of L associated with digital transfer function $H_{N,L}(^{ej\omega})$ is one higher than the corresponding value of $B_{N,L}(x)$; this is due to the fact that the transformation

Table 22.1 Example 2. Details of intermediate steps for the design meeting the specification of pass band exactly

N	Pass band cutoff point (x_c)					Stop band edge (x_s)		
	L	$B_{N,L}(x_c)$	$B_{N,L+1}(x_c)$	$B_{N,L+2}(x_c)$	β	L	$B_{N,L}(x_s)$	$B_{N,L+1}(x_s)$
3	0	0.42188	0.84375	0.98438	2.08501	0	0.01563	0.15625
4	0	0.31641	0.73828	0.94922	12.48637	0	0.00391	0.05078
5	1	0.23730	0.73281	0.89648	0.39152	1	0.01563	0.10352
6	1	0.53394	0.83057	0.96240	1.40057	1	0.00464	0.03760

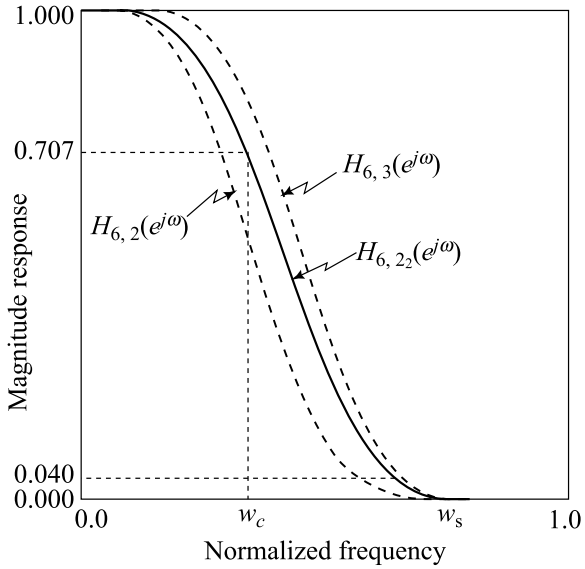


Fig. 22.6 (Example 2). The magnitude response of MAXFLAT filters and monotonic filter meeting the prescribed passband cut-off specification exactly

of Eq. (22.7) increases the degree of flatness at the extreme points by one [5]. The coefficients of the transfer function given by Eq. (22.4) are:

$$\begin{aligned} b_0 &= 0.4084574 \quad b_1 = 0.5615292 \\ b_2 &= 0.1251097 \quad b_3 = -0.0610438 \\ b_4 &= -0.0353989 \; b_5 = -0.0004853 \\ b_6 &= 0.0018318 \end{aligned}$$

Now let us consider the design of a monotonic filter to meet the specifications of both passband and stopband edge frequencies exactly. The conditions to be satisfied are:

$$\frac{B_{N,L}(x) + \beta_1 B_{N,L+1}(x) + \beta_2 B_{N,L+2}(x)}{1 + \beta_1 + \beta_2} = \delta_1 \tag{22.39a}$$

and

$$\frac{B_{N,L}(x) + \beta_1 B_{N,L+1}(x) + \beta_2 B_{N,L+2}(x)}{1 + \beta_1 + \beta_2} = \delta_2 \tag{22.39b}$$

The design procedure has already been outlined; here, in each iteration, the magnitude at x_s is computed for all the three values of L, unlike in the previous case where only two values were used. The search stops when both β_1 and β_2 turn out to be positive. Using these values of β_i's, α_i's are computed and the necessary Bernstein polynomial satisfying the given conditions is derived by using Eq. (22.35). For the example considered earlier, this method results in $N = 6, L = 2, \beta_1 = 0.80501$ and $\beta_2 = 0.28814$. The various steps involved are shown in Table 22.2. The four frequency responses are shown in Fig. 22.7. As defined earlier, $H_6{,}2_3\ (e^{j\omega})$ denotes the linear combination of three transfer functions of same order ($N = 6$), but with different degrees of tangency at $\omega = 0$, that is, $L = 2$, $L = 3$ and $L = 4$. The coefficients of the transfer function given by Eq. (22.4) are:

$$\begin{aligned} b_0 &= 0.4156528 \quad b_1 = 0.5579589 \\ b_2 &= 0.1125626 \quad b_3 = -0.0556665 \\ b_4 &= -0.0282291 \quad b_5 = -0.0022778 \\ b_6 &= 0.0000394. \end{aligned}$$

22.6.4 Comparative Study

We illustrate the advantages of this new design procedure by three sets of examples with diverse specifications. The specifications and the results of all the filters for transition widths $1/N$ (MAXFLAT design), $2/N$ and $3/N$ are shown in Table 22.3. The value of $\Delta\delta$ corresponds to the maximum permissible error at the passband cutoff frequency for the MAXFLAT filters.

It is observed that savings in the order of the filter are significant for the new method, and is, of course, at the cost of one degree of flatness. For Examples 3a and 3b in Table 22.3, the filters with 3 dB cutoff, the saving in N is significant. However, for Examples 4a and 4b, the design criteria being 95 to 5% transition width, the gain in N is not significant. This is obvious, because the magnitude characteristics are closely spaced at $\delta_1 = 0.95$. The saving in N becomes significant if lower values of $\Delta\delta$ are specified. This fact is clearly brought out by Examples (5a and 5b).

Apparently there seems to be no advantage if we use three polynomials, but by this method, with the same order of filter, we are able to exercise control over the magnitude characteristics in the stopband as well. Although this method is not of importance for simple lowpass/highpass filter design, undoubtedly, it has a significant use in orthogonalizing the bands of Quadrature Mirror filters; this is dealt in details in the next Section.

Table 22.2 Example 2. Details of intermediate steps for the design meeting the specification of passband and stopband exactly

N	Pass band cutoff point (x_c)						Stop band edge (x_s)			
	L	$B_{N,L}(x_c)$	$B_{N,L+1}(x_c)$	$B_{N,L+2}(x_c)$	β_1	β_2	L	$B_{N,L}(x_s)$	$B_{N,L+1}(x_s)$	$B_{N,L+2}(x_s)$
3	0	0.42188	0.84375	0.98438	3.54760	−0.72108	0	0.01563	0.15625	0.57813
4	0	0.31641	0.73828	0.94922	18.00510	−0.71272	0	0.00391	0.05078	0.26172
5	1	0.23730	0.63281	0.89648	−240.37800	−91.63390	1	0.00098	0.01563	0.10352
6	1	0.53394	0.83057	0.96240	0.80500	0.288141	1	0.00464	0.03760	0.03760

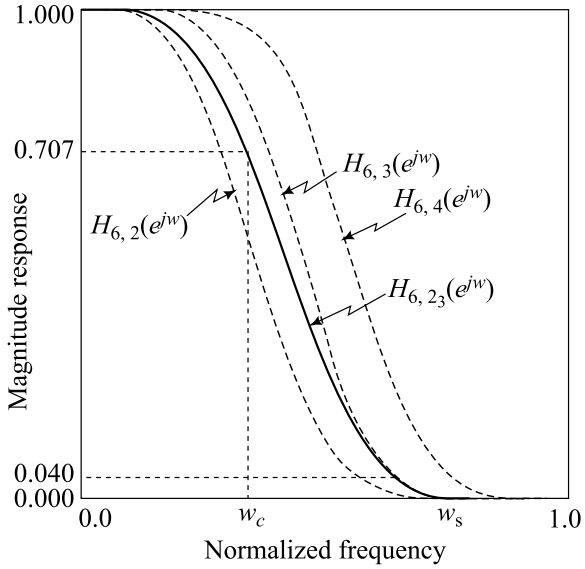


Fig. 22.7 (Example 2). The magnitude response of MAXFLAT filters and monotonic filter meeting the specifications at both passband and stopband exactly

22.7 Quadrature Mirror Filter

The theory of MAXFLAT FIR filters developed earlier can be used to design other classes of filters like the QMF, multi-band filters and two-dimensional MAXFLAT FIR filters. In this paper, we restrict our study to the first case in details and propose an efficient method for designing MAXFLAT and monotonic QMFs.

As mentioned earlier, Burgeron [18] proposed a computationally inefficient numerical method for the design of MAXFLAT QMF and arrived at two specific values of $N = 13$ and 30. As we shall see soon, the method described earlier, for the design of MAXFLAT filters, can easily be extended to design MAXFLAT QMF efficiently. However, the disadvantage of MAXFLAT QMF is that there are very few magnitude characteristics, which qualify for the purpose, and hence they are of little practical significance. Here, we propose a method by which a monotonic QMF structure with fairly low reconstruction error is generated for all values of N, by considering a linear combination of two MAXFLAT characteristics. The reconstruction error can be further reduced by the linear combination of three MAXFLAT characteristics.

A basic two-band QMF structure is shown in Fig. 22.8. In the two-band system, the aliasing error due to decimation in the analysis structure needs to be exactly cancelled by the imaging due to interpolation in the synthesis filter. To meet this end, the overall design of analysis/synthesis can be easily translated to a prototype lowpass filter $H(^{ej\omega})$, satisfying the following conditions [30]:

$$H_{lp}(e^{j\omega}) = \quad H(e^{j\omega}) \tag{22.40a}$$

$$H_{hp}(e^{j\omega}) = \quad H(e^{j\omega+\pi}) \tag{22.40b}$$

$$G_{lp}(e^{j\omega}) = \quad 2H(e^{j\omega}) \tag{22.40c}$$

Table 22.3 Comparative study based on Examples (3a)–(5b)

Example	ω_c	ω_s	δ_1	$\Delta\delta$	δ_2	Transition width					
						1/*N* (MAXFLAT)		2/*N*		3/*N*	
						N	*L*	*N*	*L*	*N*	*L*
22.3(a)	1.159279	1.570796	0.707	0.005	0.05	40	15	30	10	30	10
22.3(b)	1.884955	2.347821	0.707	0.005	0.05	35	25	22	16	22	15
22.4(a)	1.159280	1.772154	0.950	0.005	0.05	28	13	28	13	20	13
22.4(b)	1.884955	2.347821	0.950	0.005	0.05	52	40	51	39	51	38
22.5(a)	1.152791	2.150000	0.850	0.001	0.001	30	12	17	7	17	7
22.5(b)	0.927295	1.813162	0.900	0.001	0.001	45	13	25	8	25	7

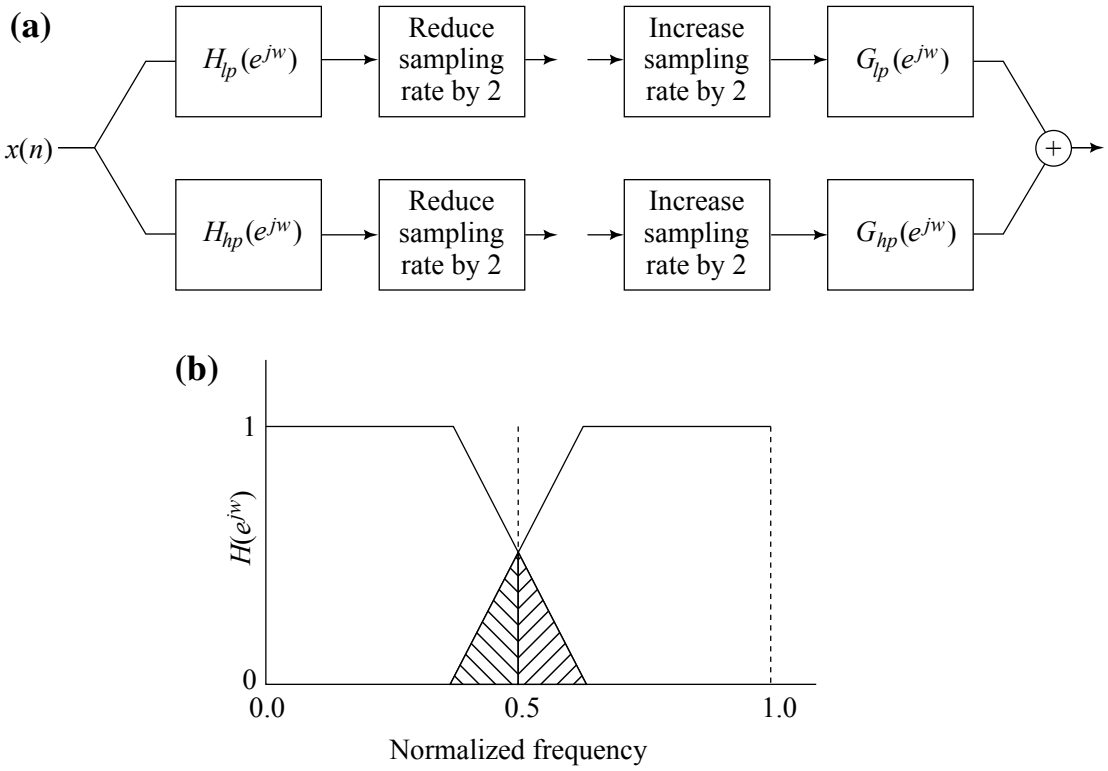


Fig. 22.8 Two-band analysis/synthesis structure for QMF design

$$G_{hp}(e^{j\omega}) = \quad -2H(e^{j\omega+\pi}) \tag{22.40d}$$

The conditions of Eqs. (22.40a) and (22.40b) imply that the filters $Hlp(e^{j\omega})$ and $H_{hp}(^{ej\omega})$ have mirror-image symmetry about the frequency $\omega = \pi/2$, as shown in Fig. 22.8b. The shaded parts indicate the regions where aliasing and imaging between the two bands occur.

The prototype lowpass filter $H(^{ej\omega})$ could be designed as either an FIR or an IIR structure. Specifically, a symmetric FIR filter with even-length impulse response can be designed to satisfy the condition

$$\left|H(e^{j\omega})\right|^2 + \left|H(e^{j\omega+\pi})\right|^2 = 1, 0 \leq \omega \leq \pi, \tag{22.41}$$

which is required to yield a QMF structure with zero reconstruction error and an overall gain of unity. Besides, the prototype lowpass $H(^{ej\omega})$ should approximate the ideal lowpass condition i.e.

$$H(e^{j\omega}) = \begin{cases} 1, & 0 \leq \omega \leq \pi/2 \\ 0, & \pi/2 \leq \omega \leq \pi. \end{cases} \tag{22.42}$$

Conditions given by Eqs. (22.41) and (22.42) impose constraints not only on the behaviour of $H(^{ej\omega})$ in the passband and stopband but also on its behaviour in the transition band.

It is well known that both the conditions of an ideal QMF and lowpass filter cannot be met *exactly;* although, a two tap FIR filter meets Eq. (22.41) exactly, its lowpass characteristics are poor. Therefore, we are required to design a symmetric even-length prototype lowpass filter with frequency response.

$$H(e^{j\omega}) = \sum_{i=1}^{N} b_i \cos[(2i-1)\omega/2] \tag{22.43}$$

which closely *approximates* the ideal requirements of Eqs. (22.41) and (22.42). It should be remembered that in all the designs discussed earlier, the resulting transfer function given by Eq. (22.4) of the MAXFLAT and monotonic FIR filters has always an odd-length impulse response. Since we intend to use the earlier methods for designing QMFs, naturally, we would arrive at unrealizable QMF transfer functions. Therefore, we need to convert the impulse response of Eq. (22.4) into those of Eq. (22.43). This conversion can easily be accomplished by interpolating the odd-length impulse response of (22.4), by a factor of two and subsampling to yield an equivalent even-length [18]. An alternate and simpler approach, would be to use a method similar to the one used by Kaiser and Reed [6]. In this method, we first compute the magnitude at $2N + 1$ points, then pad the sequence with $2N + 1$ zeros, and, finally, perform IDFT on this augmented point set. By suitably subsampling this sequence, we can easily generate the desired even-length impulse response. Although both the methods yield identical results [31], the latter is computationally simpler; hence, we shall use this method to compute the coefficients of the realizable QMF structure. Thus, in the designs to follow, first, we shall design an odd-length filter (on the assumption that it is possible to generate a QMF structure), and subsequently convert it to an equivalent even-length, by using the method just outlined. Of course, this process of conversion introduces an error in the derived magnitude characteristics.

22.7.1 Design of MAXFLAT QMF

In the design of MAXFLAT QMF, it has been well established that we are unable to exercise control over the response in the transition band, while facility for control is a positive requirement to generate QMF structure with low reconstruction error. Owing to this limitation, we need to select only those characteristics, which approximate conditions Eqs. (22.41) and (22.42) satisfactorily. The numerical method of Bergeron [18], as already mentioned, involves searching the entire two-dimensional plane of N and L, and select the characteristics of interest.

The new design method being proposed here is based on the use of Bernstein polynomial for designing MAXFLAT FIR filters [35]. Corresponding to the term QMF in z-domain, we shall use the term quadrature mirror function (qmf) in x. Since the function (22.18) is monotonic in the interval (0, 1), we have

$$E_{N,L}(x) = [B_{N,L}(x)]^2 + [B_{N,L}(1-x)]^2 - 1 \tag{22.44}$$

where $E_{N,L}(x)$ represents the error in approximating (22.41). As a first step in the design, we shall locate the points of maximum error, and then attempt to minimize the errors at these points. The problem of locating the points of maxima/minima of $E_{N,L}(x)$ is reduced to finding the roots of the polynomial $dE_{N,L}(x)/dx$ (denoted by $E'_{N,L}(x)$). Differentiating Eq. (22.44), and equating to zero, we get

$$E'_{N,L}(x) \triangleq e_{lp}(x) + e_{hp}(x) = 0 \tag{22.45a}$$

where

$$e_{lp}(x) \triangleq B'_{N,L}(x)B_{N,L}(x) \tag{22.45b}$$

and

$$e_{hp}(x) \triangleq B'_{N,L}(1-x)B_{N,L}(1-x). \tag{22.45c}$$

It can be shown that for MAXFLAT qmf (22.24), the error function can have atmost three maxima/minima, including the one at $x = 0.5$ (Appendix A). Generally, in the designs that have been verified, for $N = 5$ to 55, the absolute maximum error occurs at $x = 0.5$, with the exceptions of $N = 13$, 30, 32 and 34. As a result of this observation, we shall first design MAXFLAT qmf structures with minimum error at $x = 0.5$, and extend this design to include the exceptional cases as well.

The design of N-th order MAXFLAT QMF, with minimum reconstruction error, involves selecting the correct value of L for which the magnitude characteristic is close to –3 dB at $\omega = \pi/2$. The same search procedure, as used for designing MAXFLAT filters, can be used for generating MAXFLAT QMFs. Besides the two specific values of $N = 13$ and 30, we have $N = 51$, 53 and 55, with extremely low errors (in the range searched $N = 5$ to 55. The plot of error at $\omega = \pi/2$ versus N is shown in Fig. 22.9. As will be seen, if one permits a slightly higher error tolerance specification, say 0.025, then there are many values of N which qualify for the purpose. Some of the values of N and L which generate satisfactory responses are given in Table 22.4. In the new design procedure, the search for the values of L, which generate minimum error, is done recursively, and is confined to a maximum of three values for every N. Hence, this method is computationally more efficient than the exhaustive search procedure of Bergeron [18].

For a given N, the method just outlined yields the optimal L for which the error at $x = 0.5$ is minimum. As mentioned earlier, there are cases in which the maximum error does not occur at $x = 0.5$, and hence this value of L may not be optimum. Here, we shall propose a method to choose a suitable L for these special cases. Starting with the value of L, which gives minimum error at $x = 0.5$, we compute the maximum error for $E_{N,L+1}(x)$ and $E_{N,L-1}(x)$. If the maximum error in either of these two cases is less than that of $E_{N,L}(x)$, then the value of L is further incremented or decremented

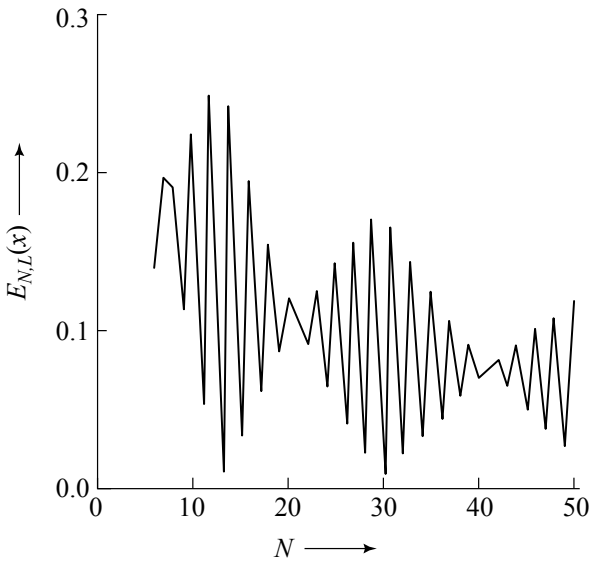


Fig. 22.9 Plot of error at $\omega = \pi/2$ versus N for MAXFLAT QMFs

Table 22.4 Some of the MAXFLAT QMFs with low reconstruction errors

N	L	Maximum reconstruction error
13	8	0.009864
28	16	0.020180
30	17	0.012409
32	18	0.020149
49	27	0.025067
51	28	0.014104
53	29	0.011994
55	30	0.015814

depending on whether the best choice was $E_{N,L+1}(x)$ or $E_{N,L-1}(x)$ respectively. It should be pointed out that in all the designs that have been verified, $N = 5$ to 55, the optimal L obtained for minimum error condition at $x = 0.5$, was also the optimum for these special cases. Hence there may not be a requirement to initiate the search procedure outlined here; the procedure has been mentioned mainly for the sake of completeness.

22.7.2 Monotonic QMF with Zero Reconstruction Error at the End and Mid Points

In the previous subsection, it was shown that generally, MAXFLAT QMFs have maximum error $\omega = \pi/2$, and, therefore, we opted to choose the characteristic which is close to -3 dB at this point. As observed, there are very few values of N which qualify for the purpose of QMF structure and hence this class of QMF has limited

applications. It would be a useful proposition to generate QMFs with low reconstruction errors for all N, and, in particular, QMFs with monotonic response. In designing such monotonic QMFs, we use the result of the earlier section, where we evolved a method to control the behaviour of the magnitude characteristics in the transition band as well—a feature not available in MAXFLAT filters.

Earlier, we have discussed a method for designing monotonic filters satisfying the specified cutoff frequency exactly by using the expression of Eq. (22.32a). In designing monotonic QMF with zero reconstruction error at $\omega = \pi/2$, we need to satisfy only one condition and hence the procedure outlined in the previous section can be used directly. Since we are interested in generating QMFs for all N, the test for stopband specification may be inhibited.

The magnitude and error characteristics of the QMF for $N = 7$, $L = 3$ and $\beta = 3.12232$ ($\beta = 0.757419$) are shown in Fig. 22.10. The error curve exhibits some interesting features and has a definite number of minima and maxima. At this point, it is worth mentioning that in the computer-aided design procedure as in [32, 33], the error surface has many local minima and the optimization routine can converge to any of the local minima rather than the global minimum. Hence, special care should be taken in the choice of the initial starting points and manual intervention may be desired while the optimization is in progress. However, in this design, as we shall see soon, we do not encounter such problems.

Let $E_{N,L2}(x)$ represent the error function corresponding to the lowpass function given by Eq. (22.32a). It can be shown that $E_{N,L2}(x)$ can have at most five max-

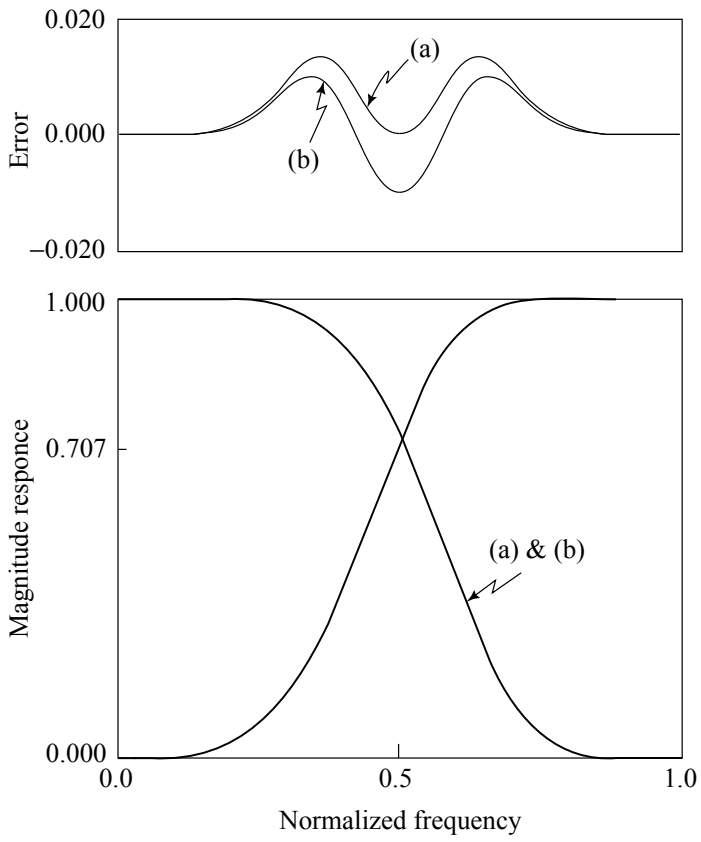


Fig. 22.10 The magnitude and error response of QMF with $N = 7$ and $L = 3$. **a** $\beta = 3.12232$ **b** $\beta = 3.35825$

ima/minima in the open interval (0, 1), (Appendix B). Since $E_{N,L2}(x)$ is constrained to have zero error at $x = 0.5$, and also observing that $E'_{N,L2}(x)$ is an anti-symmetric function about the same point, we can conclude that there can be two maxima and three minima in the interval (0, 1). The absolute value of the minima, on either extremes, is greatly reduced by the multiplying factors x^L and $(1 - x)^L$ in $E_{N,L2}(x)$. Hence the absolute maximum error is given by the value at the two maxima x_m and $(1 - x_m)$, located symmetrically about $x = 0.5$. This assertion has been verified in all the orders $N = 5$ to 55. A plot of $E'_{N,L2}(x)$ for $N = 7$ and $L = 3$, with different values of α., is shown in Fig. 22.11. It is seen that, for all values of α except for $\alpha = 0.8$, the maximum error occurs at $x = 0.5$, while for $\alpha = 0.8$, there are three such points. Although the upper limit on the number of maxima/minima is five, generally it is confined to a figure of only three or one.

The maximum error of some of the designs obtained by the method discussed in this Section is listed in Table 22.5, along with those of MAXFLAT QMFs. It is observed that the maximum error of monotonic QMF is lower than that of MAXFLAT QMF for all N, except for $N = 13$. This is expected, because the MAXFLAT error function for this specific N has positive value at $x = 0.34$, and negative value at $x = 0.5$; thus, while minimizing the absolute error at $x = 0.5$, the error at $x = 0.34$ increases. Hence, for some values of N, the monotonic QMF designed by this method, does not yield improvement over that of MAXFLAT QMF. In view of this, it would be desirable to distribute the peak error at x_m equally on either side of zero and thus minimize the maximum error. This can easily be achieved by iteratively adjusting the value of β, obtained for zero error at $x = 0.5$. Adopting this method for the case $N = 13$, we get $\beta = 0.005025$ and an error of 0.009150; the improvement is not, therefore, significant. However for $N = 7$ with $\beta = 3.35825$, the peak error is 0.010133, which amounts to a reduction of about 30%. The magnitude and error characteristics for this case are also shown in Fig. 22.10. It will be observed that

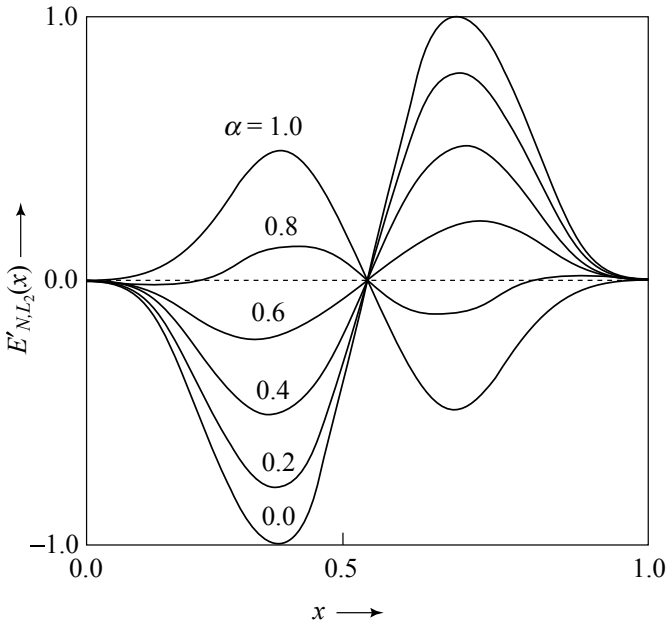


Fig. 22.11 Plot of $E'_{N,L2}(x)$ for $N = 7$ and $L = 3$

Table 22.5 A comparative study of MAXFLAT and monotonic QMF structures

N	L	Maximum reconstruction error		
		MAXFLAT QMF	Monotonic QMF	Monotonic QMF
5	4	0.320313	0.011055	0.000876[a]
6	4	0.138672	0.005591	0.001449[a]
7	5	0.196411	0.013383	0.000020[a]
8	5	0.189178	0.007945	0.000805[a]
9	6	0.113312	0.013544	0.000894[a]
10	6	0.223625	0.009857	0.000342[a]
11	7	0.052950	0.013052	0.004593
12	7	0.248970	0.011179	0.000077[a]
13	8	0.009864	0.012333	0.008230
14	9	0.241966	0.012091	0.002052
15	9	0.030108	0.011792	0.010243
16	10	0.194288	0.012615	0.004339
17	10	0.060260	0.011698	0.009016
18	11	0.154163	0.013018	0.006231
19	11	0.085515	0.011835	0.008535
20	12	0.119839	0.013241	0.007935

[a]Designs in which $E_{N,L3}(x_m)$ can be reduced to zero

the magnitude response has not changed appreciably. To sum up the procedure for designing monotonic QMF, the first step is to obtain the value of β for which the reconstruction error is zero at $\omega = \pi/2$, and adjust it to distribute the maximum error equally on either side of zero. Although the maximum error of monotonic QMF designed by this method is smaller compared to that of MAXFLAT QMF, the improvement is not a drastic one. In the following subsection, we shall attempt to minimize the maximum error at x_m.

22.7.3 Minimization of the Reconstruction Error at $\mathbf{x_m}$

In the previous subsection, we found that the points of maximum error are placed symmetrically about $x = 0.5$ and this results in QMF structures with maximum error at two points, x_m and $1 - x_m$. Here, we shall minimize the maximum error at these two points, with an added constraint of zero error at $x = 0.5$. Since the error function is symmetric, it is sufficient to reduce the error at x_m only.

Consider the monotonic lowpass function given by Eq. (22.36) discussed in previous section. Here, we have two parameters β_1 and β_2 and hence two conditions can be satisfied. For a given N and L, a linear condition relating β_1 and β_2 which

satisfies zero error condition at $x = 0.5$ can easily be obtained from Eq. (22.36). Also, as observed earlier, for $B_{N,L3}(x)$ to be monotonic, it requires both β_1 and β_2 to be positive; hence, the range of β_1 and β_2 of interest is restricted to a line segment in the first quadrant. The point of intersection of this line segment with β_1 axis corresponds to the previous case, where only one condition of zero error at $x = 0.5$ was met. Thus, it is possible to find a point along this line for which the error at x_m can be made lower than that of the previous case.

Let $E_{N,L3}(x)$ denote the error function corresponding to the lowpass function given by Eq. (22.36). The various steps involved in minimizing the error are:

(i) Locate the point x_m at which the error is maximum for the previous design.
(ii) Evaluate the magnitudes of the functions $B_{N,L}(x)$, $B_{N,L+1}(x)$ and $B_{N,L+2}(x)$ at x $= 0.5$ and x_m.
(iii) Obtain the linear condition for β_1 and β_2 for which the reconstruction error is zero at $x = 0.5$.
(iv) Minimize $E_{N,L3}(x)$ at x_m, by using Hooke and Jeeves optimization [34] routine and arrive at the optimal values of β_1 and β_2. The search procedure terminates when $E_{N,L3}(x_m)$ reduces to zero, or the condition for monotonicity is violated, that is, β_1 becomes negative.

At this point, a question that naturally arises is as follows: by minimizing the error only at x_m, does the point of maximum error shift and result in a value greater than the one with which we started at x_m? Although no analytical reason could be formulated, we have not found any contradiction to the assumption in any of the designs which have been verified for $N = 5$ to 55. The plot of the magnitude and the error characteristics of the QMF structure for $N = 7$, designed by using the present method, is shown in Fig. 22.12; as will be seen, the error is virtually zero. The maximum error of some of the designs, obtained by the method discussed, is also listed in Table 22.5. As is obvious, the maximum error in this design is much lower than in the previous one; the reduction is particularly significant for lower order filters.

Generally, it has been found that for most cases, the error at x_m could not be reduced to zero, because the limiting value of $\beta_1 = 0$ is attained first. The error can be reduced further, provided we compromise on the monotonic behaviour of the response characteristic; this is not a topic of interest in the present discussion. However, in those cases where we have been able to reduce the error at x_m to zero, it has been observed that the maximum error occurs at four other points but the values are extremely low, as seen from Table 22.5.

In all the results discussed so far, we have made the basic assumption that it is possible to design the QMF structure with an odd-length impulse response. As stated earlier, we need to generate an equivalent even-length for a realizable QMF structure, which satisfies Eq. (22.41). The magnitude and the error characteristics of an equivalent even-length for $N = 7$, is also shown in Fig. 22.12. It is observed that the process of conversion has introduced ripples in the error function, but the transition width has decreased considerably, a compromise that always exists.

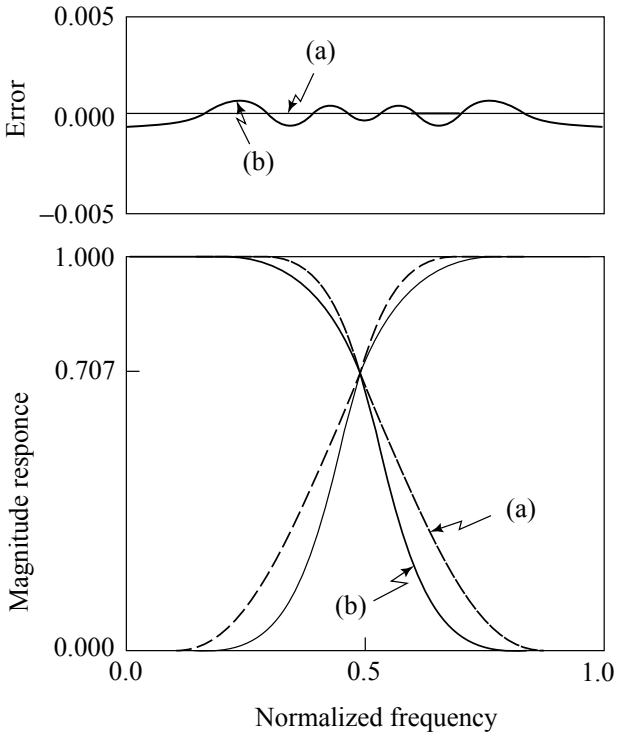


Fig. 22.12 The magnitude and error response of a monotonic QMF with $N = 7$, **a** odd-length $H(z)$ **b** even-length $H(z)$

A true comparison of performance between the class of QMF studied here and those obtained by using the optimization techniques cannot really be carried out because of the difference in the formulation of the problems. In the designs developed here, we are primarily interested in the monotonic behaviour of the magnitude characteristics, while the optimization techniques aim at minimizing the average error. Obviously, the numerical methods can result in ripples in both passband and stopband. However as an illustrative example, different responses of a 16th order filter are shown in Fig. 22.13. The designs named 16A and 16B are those which are obtained by the method in [32]; the coefficients for these designs are readily available in [30]. The QMF structures developed here generally have lower reconstruction error and greater transition width than those resulting from numerical methods.

22.8 Conclusion

In this chapter, we have discussed a new approach for the design of MAXFLAT FIR filters by using the Bernstein polynomial and have established a simple relationship between the other known methods for this purpose. We proposed a novel method to design MAXFLAT filters with arbitrary magnitude specifications. A new, matrix approach for determining the coefficients of MAXFLAT filters has also been discussed. For the case where the coefficients are associated with powers of cosine, it was shown that the transformation matrix is a product of a Q-matrix and a sim-

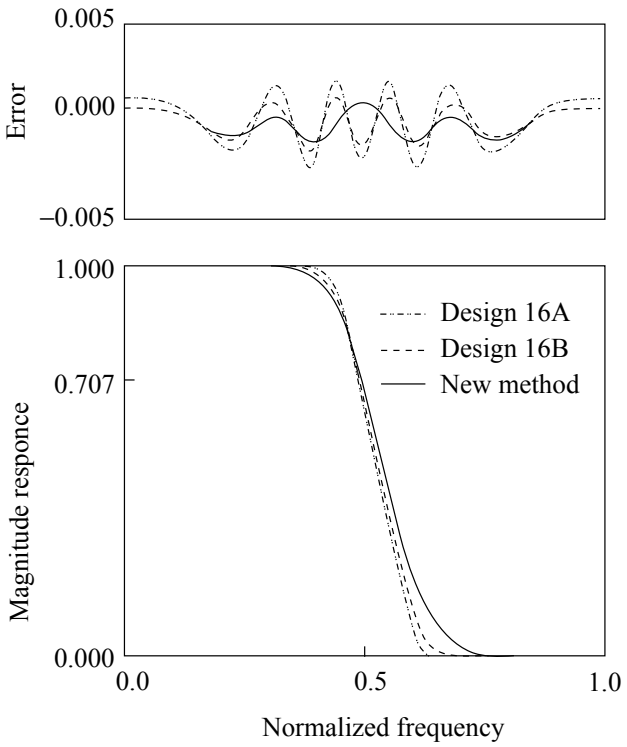


Fig. 22.13 The magnitude and error response of QMFs with $N = 8$

ple diagonal matrix. The transformation matrix required to generate b_i's could be generated either directly or recursively.

We proposed a new technique for the design of monotonic filters meeting the desired magnitude specification at the passband edge exactly. We have also established a link between the earlier works in this area. We have brought out the advantages of the new method by example of diverse specifications. The application of MAXFLAT and monotonic FIR filters to the design of QMFs is discussed in details. First, we discussed an optimal method for designing MAXFLAT QMF structures. Next, we extended the design to generate a new class of monotonic QMFs with zero reconstruction error at $\omega = \pi/2$, by using a linear combination of two MAXFLAT characteristics. Finally, we minimized the maximum error resulting from the previous design, by a simple two-variable optimization technique. Monotonic QMFs, discussed here, can become a good starting point for the optimization techniques, a problem generally faced in these methods.

References

1. O. Herrmann, On the approximation problem in nonrecursive digital filter design. IEEE Trans. CT-**18**, 411–413 (1971)
2. J.A. Miller, Maximally flat nonrecursive digital filters. Electron. Lett. **8**, 157–158 (1972)
3. M.F. Fahmy, Maximally flat nonrecursive digital filters. Int. J. Circuit Theory Appl. **4**, 311–313 (1976)
4. J.F. Kaiser, Comments on maximally flat nonrecursive digital filters. Int. J. Circuit Theory Appl. **5**, 103 (1977)

5. R.W. Hamming, *Digital Filters* (Englewood Cliffs, New Jersey; Prentice-Hall, 1977)
6. J.F. Kaiser, W.A. Reed, Data smoothing using lowpass digital filters. Rev. Sci. lnstrum. **48**, 1447–1457 (1977)
7. C. Gumacos, Weighting coefficients for certain maximally flat nonrecursive digital filters. IEEE Trans. CAS-**25**, 234–235 (1978)
8. J.F. Kaiser, Design subroutine (MAXFLAT) for symmetric FIR low pass digital filters, with maximally-flat pass and stop bands, in *Programs for Digital Signal Processing* (IEEE Press, New York, 1979), pp. 5.3-1–5.3-6
9. M.U.A. Bromba, H. Zeigler, Explicit formula for filter function of maximally flat nonrccursive digital filters. Electron. Lett. **16**, 905–906 (1980)
10. B.C. Jinaga, S.C. Dutta Roy, Explicit formulas for the weighting coefficients of maximally flat nonrecursive digital filters. Proc. IEEE **72**, 1092 (1984)
11. B.C. Jinaga, S.C. Dutta Roy, Coefficients of maximally-flat nonrecursive digital filters. Sig. Process. **7**, 185–189 (1984)
12. B.C. Jinaga, S.C. Dutta Roy, Explicit formula for the coefficients of maximally flat nonrecursive digital filter transfer function expressed in powers of cos ω. Proc. IEEE **73**, 1135–1136 (1985)
13. P. Thajchayapong, M. Puangpool, S. Banjongjit, Maximally flat FIR filters with prescribed cut-off frequency. Electron. Lett. **16**, 514–515 (1980)
14. B.C. Jinaga, S.C. Dutta Roy, Coefficients of maximally flat low and high pass nonrecursive digital filters with specified cut-off frequency. Sig. Process. **9**, 121–124 (1985)
15. P.P. Vaidyanathan, Efficient and multiplierless design of FIR filters with very sharp cutoff via maximally flat building blocks. IEEE Trans. CAS-**32**, 236–244 (1985)
16. P.P. Vaidyanathan, Optimal design of linear phase FIR digital filters with very flat passband and equiripple stopbands. IEEE Trans. CAS-**32**, 904–907 (1985)
17. P.J. Davis, *Interpolation and Approximation* (Dover, New York, 1975)
18. L.E. Bergeron, A maximally-flat filter design algorithm for Quadrature Mirror Filter (QMF), in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Washington, D.C. (1979), pp. 828–831
19. G.G. Lorentz, *Bernstein Polynomial* (Toronto 1953)
20. L.R. Rajagopal, S.C. Dutta Roy, Design of maximally-flat FIR filters using the Bernstein polynomial. IEEE Trans. CAS-**34**, 1587–1590 (1987)
21. L.R. Rajagopal, S.C. Dutta Roy, A matrix approach for the coefficients of maximally-flat FIR filters expressed in powers of cos ω. Electron. Lett. **18**, 391–393 (1987)
22. C.J. Greaves, G.A. Gagne, C.W. Bordner, Evaluation of integrals appearing in minimization problems of discrete-data systems. IEEE Trans. AC-**11**, 145–148 (1966)
23. E.I. Jury, W.C. Chen, Combinatorial Rules for some useful transformations. IEEE Trans. CT-**10**, 476–480 (1973)
24. N.K. Bose, Properties of the Q-matrix in bilinear transformation. Proc. IEEE **71**, 1110–1111 (1983)
25. A.V. Oppenheim, W.F.G. Mecklenbrauker, R.M. Mersereau, Variable cutoff linear phase digital filters. IEEE Trans. CAS-**23**, 199–203 (1976)
26. L.R. Rajagopal, S.C. Dutta Roy, A matrix approach for the coefficients of maximally-flat FIR filter transfer functions. IEEE Trans. ASSP, ASSP-**36**, 1914–1917 (1988)
27. L.R. Rajagopal, S.C. Dutta Roy, Design of optimal maximally flat FIR filters using the Bernstein polynomial, in *International Symposium on Electronic Devices, Circuits and Systems*, Kharagpur, India (1987)
28. S.C. Dutta Roy, L.R. Rajagopal, A new approach to the design of maximally-flat and monotonic FIR filters, in *Indo-US workshop on Systems and Signal Processing*, Bangalore (1988)
29. L.R. Rajagopal, S.C. Dutta Roy, Design of optimal maximally flat FIR filters with arbitrary magnitude specifications. IEEE Trans. ASSP, ASSP-**37**, 4, 512–218 (1989)
30. R.E. Crochiere, L.R. Rabiner, *Multirate Digital Signal Processing* (Englewood Cliffs, New Jersey, Prentice-Hall, 1983)
31. S.M. Kay, S.L. Marple, Spectrum analysis—a modern perspective. Proc. IEEE **69**, 1380–1419 (1981)

32. J.D. Johnston, A filter family designed for use in quadrature mirror filter banks, in *IEEE International Conference on Acoustics, Speech, and Signal Processing* (1980), pp. 291–294
33. V.K. Jain, R.E. Crochiere, A novel approach to the design of analysis/synthesis filter banks, in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Boston (1983), pp. 228–231
34. L.C.W. Dixon, *Nonlinear Optimization* (The English Universities Press, London, 1972)
35. L.R. Rajagopal, S.C. Dutta Roy, A new class of monotone quadrature mirror filters, in Symposium on Signals, Systems and Sonars, vol. 88, Cochin, India (1988)

Chapter 23
On the Design of FIR Notch Filters

Abstract A novel approach for designing linear phase FIR notch filters has been proposed. Mathematical formulas for computing the weights required in the design have been derived. Exact notch frequency point, as specified, can be achieved through the proposed technique. Design examples have also been given.

Keywords Notch filters · FIR digital filters · Signal processing

23.1 Introduction

Digital notch filters find extensive use in various fields such as automatic control, biomedical engineering, communication and radar. A variety of approaches are available for the design of IIR as well as FIR notch filters [1–3]. We present here a novel approach for designing linear phase FIR notch filters, which is based on the transformation of a lowpass filter. The method allows realization of the desired notch frequency (ω_d) exactly besides resulting in a 3-dB rejection bandwidth ($\overline{\text{BW}}$), which is better than the specified one. The notch filter may be designed with a maximally flat (MF) or equiripple characteristics. We shall discuss the MF design in detail, and then briefly mention the equiripple case.

23.2 Design

Let $H_{LP}(\omega)$ be the frequency response of a zero-phase lowpass FIR filter such that

$$H_{LP}(0) = 1, \tag{23.1a}$$

$$H_{LP}(\omega_d) = 1/2, \tag{23.1b}$$

Source: B. Kumar, S. B. Jain & S.C. Dutta Roy, "On the Design of FIR Notch Filters," *IETE Journal of Research*, vol 43, pp 65–68, January–February 1997.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_23

and

$$H_{LP}(\pi) = 0 \tag{23.1c}$$

We can obtain a notch filter by using the transformation [2]

$$H(\omega) = 2H_{LP}(\omega) - 1 \tag{23.2}$$

Using Eq. (23.1a–23.1c) in Eq. (23.2), results in the notch filter with

$$H(0) = 1, \tag{23.3a}$$

$$H(\pi) = -1, \tag{23.3b}$$

and

$$H(\omega_d) = 0, \tag{23.3c}$$

If $H_{LP}(\omega)$ is maximally flat with flatness distributed between $\omega = 0$, and $\omega = \pi$, then so is $H(\omega)$. The design problem of a notch filter is thus reduced to that of designing a maximally flat lowpass filter $H_{LP}(\omega)$ such that $H_{LP}(\omega) = 1/2$ at $\omega = \omega_d$.

A maximally flat lowpass filter $H_{LP}(\omega)$ can be obtained from the technique suggested by Thajchayapong et al. [4] which is a modification of Miller's design [5]. Miller has suggested transforming the zero phase polynomial

$$H_0(z) = \sum_{i=0}^{n} C_i(z^i + z^{-i}) \tag{23.4}$$

into a rational function $\hat{H}(q)(q = \sigma + j\Omega)$ through the bilinear transformation $z^{-1} = (1 - q)/(1 + q)$. Then by imposing the condition of maximal flatness of $\hat{H}(q)$, for $q = j\Omega$ with mth degree of flatness at $\Omega = \infty$ (i.e., $\omega = \pi$; because $\Omega = \tan \omega/2$), the resulting analog filter has the transfer function [4], Eq. (23.2)

$$\hat{H}(j\Omega) = \frac{1}{(1 + \Omega^2)^n} \left\{ \sum_{l=0}^{n-m} \binom{n}{l} \Omega^{2l} \right\} \tag{23.5}$$

Clearly, for a given n, Eq. (23.5) yields n different lowpass filters as m varies from 1 to n. Thajchayapong et al. [4] have suggested a method of obtaining transitional filters between two adjacent values of m, say m_1 and $m_2 = m_1 - 1$, by modifying Eq. (23.5) as follows:

$$\hat{H}_1(j\Omega) = \frac{1}{(1 + \Omega^2)^n} \left\{ \sum_{l=0}^{n-m_1} \binom{n}{l} \Omega^{2l} + C_{n-m_2} \Omega^{2(n-m_2)} \right\} \tag{23.6}$$

The value of constant C_{n-m_2} in Eq. (23.6) is found by forcing $\hat{H}_1(j\Omega)|_{\Omega=\Omega_d} = 1/2$, where $\Omega_d = \tan(\omega_d/2)$. This gives us

$$C_{n-m_2} = \frac{1}{\Omega_d^{2(n-m_2)}}\left[\frac{(1+\Omega_d^2)^n}{2} - \sum_{l=0}^{n-m_1}\binom{n}{l}\Omega_d^{2l}\right] \tag{23.7}$$

Equation (23.6) is now transformed back to the z-plane by using the transformation $j\Omega = (1-z^{-1})/(1+z^{-1})$,

$$\left.\begin{aligned}\bar{H}_1(z) &= \hat{H}_1(j\Omega)|_{j\Omega=(1-z^{-1})/(1+z^{-1})} \\ &= 2^{-2n}(1+z^{-1})^{2n}z^n\left\{\sum_{l=0}^{n-m_1}\binom{n}{l}(-1)^l((1-z^{-1})/(1+z^{-1}))^{2l}\right. \\ &\quad \left. + C_{n-m_2}(-1)^{(n-m_2)}((1-z^{-1})/(1+z^{-1}))^{2(n-m_2)}\right\}\end{aligned}\right. \tag{23.8}$$

Using the Binomial expansion for generic power k, i.e.

$$(1+z^{-1})^k = \sum_{q=0}^{k}\binom{k}{q}z^{-q} \tag{23.9}$$

and rearranging the terms (after somewhat involved algebraic manipulations), we can write Eq. (23.8) as

$$\begin{aligned}\bar{H}_1(z) = 2^{-2n}z^n\Bigg[\sum_{i=0}^{2n}\Bigg\{&\sum_{l=0}^{n-m_1}\sum_{q=0}^{2l}(-1)^{l+q}\binom{n}{l}\binom{2l}{q} \\ &\times\binom{2n-2l}{i-q} + C_{n-m_2}(-1)^{n-m_2}\sum_{q=0}^{2m_2}(-1)^{i-q}\binom{2m_2}{q} \\ &\binom{2n-2m_2}{i-q}\Bigg\}z^{-i}\Bigg]\end{aligned} \tag{23.10}$$

A causal, FIR lowpass filter $\bar{H}_2(z)$ is related to a zero phase FIR lowpass filter $\bar{H}_1(z)$ by

$$\bar{H}_2(z) = z^{-n}\bar{H}_1(z) \tag{23.11}$$

Also, we know,

$$\bar{H}_2(z) = \sum_{i=0}^{2n}h(i)z^{-i},\ n = (N-1)/2 \tag{23.12}$$

where N is the filter length. Comparing Eqs. (23.10), (23.11) and (23.12), we obtain

$$h(i) = 2^{-2n}\left[\sum_{l=0}^{n-m_1}\sum_{q=0}^{2l}(-1)^{l+q}\binom{n}{l}\binom{2l}{q}\binom{2n-2l}{i-q} + C_{n-m_2}(-1)^{n-m_2}\sum_{q=0}^{2m_2}(-1)^{i-q}\binom{2m_2}{q}\binom{2n-2m_2}{i-q}\right] \tag{23.13}$$

Changing i to $(n-i)$ in Eq. (23.13), we may express $h(n-i)$ as

$$h(n-i) = A(i) + C_{n-m_2}B(i) \tag{23.14a}$$

where

$$A(i) = 2^{-2n}\sum_{l=0}^{n-m_1}\sum_{q=0}^{2l}(-1)^{l+q}\binom{n}{l}\binom{2l}{q}\binom{2n-2l}{n-i-q} \tag{23.14b}$$

$$B(i) = 2^{-2n}\sum_{q=0}^{2m_2}(-1)^{q+m_2+i}\binom{2m_2}{q}\binom{2n-2m_2}{n-i-q} \tag{23.14c}$$

The frequency response $H_{LP}(\omega)$ of zero phase lowpass filter ($\bar{H}_1(z)$) is given by

$$H_{LP}(\omega) = \bar{H}_1(z)|_{z=e^{j\omega}} \triangleq \sum_{i=0}^{n} a_i \cos(i\omega) \tag{23.15}$$

Also, a zero phase, notch filter $H(\omega)$ is expressed as

$$H(\omega) = \sum_{i=0}^{n} d_i \cos(i\omega) \tag{23.16}$$

Using Eqs. (23.2), (23.15) and (23.16), we finally obtain

$$d_i = \begin{cases} 2h(n) - 1, & i = 0 \\ 4h(n-i), & i = 1, 2, \ldots, n \end{cases} \tag{23.17}$$

Hence, the coefficients d_i of the desired notch filter (having notch at $\omega = \omega_d$) are computed from Eqs. (23.14a–23.14c) and (23.17).

The values of n and m_1, required for computation of $h(n-i)$ are obtained from the empirical formulas, proposed earlier by the authors in [6], and reproduced here for completeness.

$$n \geq \text{Integer part of } [1/2\{(\pi/\overline{\text{BW}})^2 - (\pi/\overline{\text{BW}}) + 3\}] \tag{23.18}$$

$$m_1 \geq \text{Integer part of } [n(0.55 + 0.5\cos\omega_d)] \tag{23.19}$$

We now study the performance of the aforementioned design of the notch filter.

23.3 Design Examples and Performance

Example Design a linear phase FIR notch filter with notch frequency $\omega_d = 0.2\pi$ radian; and 3-dB rejection bandwidth $\overline{\text{BW}} \leq 0.5$ rad

(i) Compute the values of n and m_1 from Eqs. (23.18) and (23.19), respectively. These are found to be $n = 18$ and $m_1 = 17$.

(ii) By using $\Omega_d = \tan(\omega_d/2)$, we get $\Omega_d = \tan(0.1\pi) = 0.324919$ rad. From Eq. (23.7), we obtain the value of C_{n-m_2}, with $n = 18$ and $m_2 = m_1 - 1 = 16$ as 12.953.

(iii) By using Eqs. (23.14a–23.14c), we get $A(i)$, $B(i)$ and $h(n-i)$, for $i = 0$–18 and then compute d_i from Eq. (23.17). These values are given in Table 23.1.

Table 23.1 Values of $A(i)$, $B(i)$, $h(n-i)$ and weights d_i, $i = 0, 1, 2, \ldots, 18$ for a notch filter example with $\omega_d = 0.2\pi$ rad and $(\overline{\text{BW}}) \leq 0.5$ rad

i	$A(i)$	$B(i)$	$h(n-i)$	d_i
0	0.19997748D + 00	0.34301454D − 03	0.20442070D + 00	−0.59115859D + 00
1	0.18230320D + 00	0.25274756D − 03	0.18557716D + 00	0.74230863D + 00
2	0.13672740D + 00	0.45133493D − 04	0.13731204D + 00	0.54924815D + 00
3	0.81023646D − 01	−0.14442718D − 03	0.79152815D − 01	0.31661126D + 00
4	0.33146037D − 01	−0.21664076D − 03	0.30339791D − 01	0.12135916D + 00
5	0.28822641D − 02	−0.16954495D − 03	0.68607132D − 03	0.27442853D − 02
6	−0.98877671D − 02	−0.71428658D − 04	−0.10813015D − 01	−0.43252060D − 01
7	−0. 11240830D − 01	0.81632752D − 05	−0.11135087D − 01	−0.44540349D − 01
8	−0.79262262D − 02	0.41758292D − 04	−0.73853121D − 02	−0.29541248D − 01
9	−0.42664625D − 02	0.399791 17D − 04	−0.37485948D − 02	−0.14994379D − 01
10	−0.18494528D − 02	0.25065441D − 04	−0.15247687D − 02	−0.60990749D − 02
11	−0.65596355D − 03	0.11793803D − 04	−0.50319305D − 03	−0.20127722D − 02
12	−0.19031041D − 03	0.43234322D − 05	−0. 13430703 D − 03	−0.53722812D − 03
13	−0.44671353D − 04	0.12419187D − 05	−0.28584216D − 04	−0.11433686D − 03
14	−0.83024061 D − 05	0.27603528D − 06	−0.47267956D − 05	−0.18907182D − 04
15	−0.11785305D − 05	0.46042260D − 07	−0.58212419D − 06	−0.23284968D − 05
16	−0.12022792D − 06	0.54424163D − 08	−0.49729831D − 07	−0.19891933D − 06
17	−0.78580342D − 08	0.40745363D − 09	−0.25801022D − 08	−0.10320409D − 07
18	−0.24738256D − 09	0.14551915D − 10	−0.58884986D − 10	−0.23553994D − 09

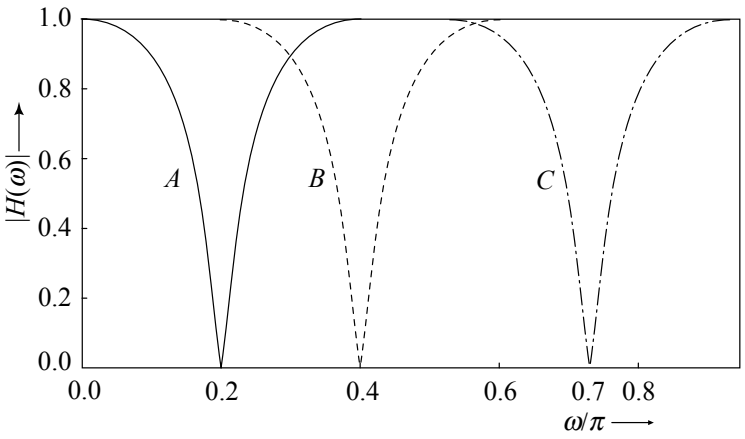


Fig. 23.1 Frequency response $|H(\omega)|$ of the designed notch filters A, B, C, whose parameters are given in Table 23.2

Table 23.2 Various parameters for Fig. 23.1

Response curve	ω_d (rad)	$\overline{\text{BW}}$ (rad)		$\left(n = \frac{N-1}{2}\right)$	m_1	C_{n-m_2}
		Specified	Realized			
A	0.2π	0.50	0.490	18	17	12.953
B	0.4π	0.45	0.446	22	15	15912.430
C	0.7π	0.40	0.399	28	7	310229.079

(iv) The magnitude response of the $\left(\sum_{i=0}^{18} d_i \cos i\omega\right)$ designed notch filter is shown in Fig. 23.1 (curve A). The value of 3-dB rejection bandwidth ($\overline{\text{BW}}$) is found to be 0.49 rad, which is better than the specified one. The magnitude responses for a few other designs using the above methodology are also shown in the same figure (curves B and C) and the related parameters are given in Table 23.2.

23.3.1 Equiripple Design

The proposed design procedure can also be adapted for an equiripple notch filter, $\bar{H}(\omega)$ by using an equiripple LPF $\bar{H}_{LP}(\omega)$ in Eq. (23.2). If we constrain $\bar{H}_{LP}(\omega)$ as

$$\bar{H}_{LP}(\omega) = \begin{cases} 1/2(2 \pm \delta_1), & 0 < |\omega| < B_1 \\ 1/2(1 \pm \delta_2), & |\omega| = \omega_d \\ \pm\delta_1/2, & B_2 < |\omega| < \pi \end{cases} \tag{23.20}$$

where $B_2 - B_1 = (\overline{\text{BW}})$, δ_1 and δ_2 are, respectively, the maximum ripples in the passband and the stopband of the LPF, then the resulting notch filter will have

$$\bar{H}(\omega) = \begin{cases} 1 \pm \delta_1, & 0 < |\omega| < B_1 \\ \pm\delta_2, & |\omega| = \omega_d \\ -1 \pm \delta_1, & B_2 < |\omega| < \pi \end{cases} \tag{23.21}$$

The equiripple linear phase LPF $\bar{H}_{LP}(\omega)$ may be obtained by any of the conventional methods such as the McClellan and Parks algorithm.

23.4 Conclusion

An analytical approach by which explicit formulas for computation of the exact weights of a linear phase FIR notch filter has been proposed. This design approach is particularly attractive when the available memory is limited. Also, we obtain the exact notch frequency by the suggested method. Moreover, the design procedure presented in this chapter has an edge over the design methods, based on optimization/iterative techniques, since the latter requires more time of computation.

Acknowledgements The authors were thankful to the referees for their constructive suggestions.

References

1. K. Hirano, S. Nishimura, S.K. Mitra, Design of digital notch filters. IEEE Trans. Circuits Syst. CAS-**21**, 540–546 (1974)
2. T.H. Yu, S.K. Mitra, H. Babic, Design of linear phase FIR notch filters. Sādhanā **15**, 133–155 (1990)
3. M.H. Er, Designing notch filter with controlled null width. Signal Process. **24**, 319–329 (1991)
4. P. Thajchayapong, M. Puangpool, S. Banjongjit, Maximally flat FIR filter with prescribed cut off frequency. Electron. Lett. **16**, 514–515 (1980)
5. J.A. Miller, Maximally flat nonrecursive digital filters. Electron. Lett. **8**, 157–158 (1972)
6. S.C. Dutta Roy, S.B. Jain, B. Kumar, Design of digital FIR notch filters. IEE Proc. Vis. Image Signal Process. **14**, 334–338 (1994)

Chapter 24
A Simple Solution for the Analytic Inversion of Van der Monde and Confluent Van der Monde Matrices

Abstract In this chapter, a simple and elegant approach is suggested for the explicit, analytic inversion of the Van der Monde and Confluent Van der Monde matrices, based on known results for certain polynomial interpolation problems. These matrices arise in a number of signal processing applications, and the closed-form results should be of interest to the signal processing community.

Keywords Matrix inversion · Van der Monde matrix

24.1 Introduction

Inversion of a Van der Monde matrix is often required in many signal processing situations, such as signal reconstruction, spectral estimation, DOA estimation via sensor array, system identification, etc. [1, 2]. Neagoe [3] has recently derived an exact expression, for the inverse of a Van der Monde matrix, using the classical approach of determinants.

In this chapter, we suggest a much simpler, alternative approach for the explicit inversion of this matrix, which makes use of the well known connection between Van der Monde matrices and the Lagrange interpolation formula [1, 4]. It is shown that this approach can also be easily extended for the explicit inversion of a related matrix, viz, the Confluent Van der Monde matrix [4] associated with a class of interpolation problems, with derivative constraints, which are often formulated in beamforming problems [2].

Source: Monika Agrawal, Surendra Prasad & S.C. Dutta Roy, "A Simple Solution for the Analytic Inversion of Van der Monde and Confluent Van der Monde Matrices," *IETE Journal of Research*, vol 47, pp 217–219, September–October 2001.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_24

24.2 Inversion Formula for Van der Monde Matrix

Let **A** be a Van der Monde matrix of size $N \times N$ over C^N, given by

$$\mathbf{A} = \begin{bmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{N-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{N-1} \\ & & & & \\ & & & & \\ 1 & a_N & a_N^2 & \cdots & a_N^{N-1} \end{bmatrix} \quad a_i \neq a_j \quad \forall i \neq j \tag{24.1}$$

Consider the set of linear equations given by

$$\mathbf{Ay} = \mathbf{b} \tag{24.2a}$$

so that

$$\mathbf{y} = \mathbf{A}^{-1}\mathbf{b}, \tag{24.2b}$$

which is guaranteed to exist for distinct $\{a_i, i = 1, \ldots, N\}$. It is well known that the equation set (24.2a, 24.2b) arises in the problem of finding the coefficients of a polynomial of order $N - 1$ which passes through N given points. Let $f(x)$ be the required polynomial s.t.

$$f(a_i) = b_i \; i = 1, 2, \ldots, N \tag{24.3}$$

The Lagrange interpolation formula yields the unique and well-known solution for such a polynomial viz,

$$f(x) = \sum_{i=1}^{N} b_i l_i(x) \tag{24.4a}$$

where b, is the ith element of the vector **b** and $l_i(x)$ is given by

$$l_i(x) = \prod_{j-1, j \neq i}^{N} \frac{x - a_j}{a_i - a_j} \quad i = 1, 2, \ldots, \; N \tag{24.4b}$$

The polynomial $l_i(x)$ can be expanded as

$$l_i(x) = \alpha_0^{(i)} x^{N-1} - \alpha_1^{(i)} x^{N-2} + \cdots + (-1)^k \alpha_k^{(i)} x^{N-k-1} + \cdots + (-1)^{N-1} \alpha_{N-1}^{(i)}, \quad i = 1, 2, \ldots, N \tag{24.5a}$$

where

$$\alpha_k^{(i)} = \frac{1}{\prod\limits_{j=i, j\neq i} (a_i - a_j)} \sum_{i_1 \neq i}^{N} \sum_{i_2 \neq (i_1, i)}^{N} \cdots \sum_{i_k \neq (i_1 \cdots i_{k-1}, i)}^{N} a_{i_1} a_{i_2} \ldots a_{i_k}$$
$$i = 1, 2, \ldots N; k = 1, 2, \ldots N + 1 \tag{24.5b}$$

Using Eq. (24.5a) in Eq. (24.4a), and combining the terms of same power, we get

$$\begin{aligned} f(x) = {} & x^{N-1}(\alpha_0^{(1)} b_1 + (\alpha_0^{(2)} b_2 + \cdots + a_0^{(N)} b_N) - x^{N-2}(\alpha_1^{(1)} b_1 \\ & + \alpha_1^{(2)} b_2 + \cdots + \alpha_1^{(N)} b_N) + \cdots + (-1)^k x^{N-k-1}(\alpha_k^{(1)} b_1 \\ & + \alpha_k^{(2)} b_2 + \cdots + \alpha_k^{(N)} b_N) + \cdots + (-1)^{N-1}(\alpha_{N-1}^{(1)} b_1 + \alpha_{N-1}^{(2)} b_2 \cdots + \alpha_{N-1}^{(N)} b_N) \end{aligned} \tag{24.6}$$

Comparison of Eq. (24.6) with Eq. (24.2a, 24.2b) yields the ith coefficient of $f(x)$ (i.e. coefficient of the term x^i) which forms the $(i + 1)$th element of the vector y. Hence Eq. (24.3) becomes

$$\mathbf{y} = \begin{bmatrix} (-1)^{N-1}\alpha_{N-1}^{(1)} & (-1)^{N-1} & \alpha_{N-1}^{(2)} & \cdots & (-1)^{N-1}\alpha_{N-1}^{(m)} & \cdots & (-1)^{N-1} & \alpha_{N-1}^{(N)} \\ (-1)^{N-k}\alpha_{N-k}^{(1)} & (-1)^{N-k} & \alpha_{N-k}^{(2)} & \cdots & (-1)^{N-k}\alpha_{N-k}^{(m)} & \cdots & (-1)^{N-k} & \alpha_{N-k}^{(N)} \\ \alpha_0^{(1)} & & \alpha_0^{(2)} & \ldots & \alpha_0^{(m)} & & & \alpha_0^{(N)} \end{bmatrix} \begin{bmatrix} b_1 \\ b_k \\ b_N \end{bmatrix} \tag{24.7}$$

This gives

$$\mathbf{A^{-1}} = \begin{bmatrix} (-1)^{N-1}\alpha_{N-1}^{(1)} & (-1)^{N-1} & \alpha_{N-1}^{(2)} & \cdots & (-1)^{N-1} & \alpha_{N-1}^{(m)} & \cdots & (-1)^{N-1} & \alpha_{N-1}^{(N)} \\ (-1)^{N-k}\alpha_{N-k}^{(1)} & (-1)^{N-k} & \alpha_{N-k}^{(2)} & \cdots & (-1)^{N-k} & \alpha_{N-k}^{(m)} & \cdots & (-1)^{N-k} & \alpha_{N-k}^{(N)} \\ \alpha_0^{(1)} & & & \alpha_0^{(2)} & \ldots & \alpha_0^{(m)} & & & \alpha_0^{(N)} \end{bmatrix} \tag{24.8}$$

which is identical to the result in [3].

24.3 Inversion Formula for the Confluent Van der Monde Matrix

An interesting and useful application of this method arises in the inversion of a matrix **A** defined by

$$\mathbf{A} = \begin{vmatrix} 1 & a_1 & a_1^2 & \cdots & a_1^{2N-1} \\ 1 & a_2 & a_2^2 & \cdots & a_2^{2N-1} \\ . & . & . & . & . \\ 1 & a_N & a_N^2 & \cdots & a_N^{2N-1} \\ 0 & 1 & 2a_1 & \cdots & (2N-1)a_1^{2N-2} \\ 0 & 1 & 2a_2 & \cdots & (2N-1)a_2^{2N-2} \\ . & . & . & . & . \\ 0 & 1 & 2a_N & \cdots & (2N-1)a_N^{2N-2} \end{vmatrix} a_i \neq a_j, \forall i \neq j \tag{24.9}$$

where the top N rows correspond to a Van der Monde matrix and the bottom N rows to the derivatives of top rows w.r.t. $\{a_1, a_2, \ldots, a_N\}$, respectively.

This matrix is the well-known Confluent Van der Monde [4] matrix. It arises in the following kind of polynomial interpolation problems: find a polynomial $f(x)$ of degree $(2N-1)$ such that

$$f(a_i) = b_i; \quad f'(a_i) = b_{N+i} 1 \leq i \leq N \tag{24.10}$$

Such problems are often formulated in the context of beamforming when, in addition to steering nulls of a sensor array in certain desired directions, one may want to flatten these nulls so as to handle wideband and moving jammers [2].

Returning to the inversion of A, it is well known that the solution to the interpolation problem of (24.10) is obtained by choosing $f(x)$ to be the Hermite polynomial [5] given by

$$f(x) = \sum_{i=1}^{N} [1 - 2(x - a_i) l_i^2(a_i)]\, l_i'(x) b_i + \sum_{i=1}^{N} (x - a_i) \quad l_i^2(x) b'_{N+1} \tag{24.11}$$

Using this result and repeating the steps earlier, it is easy to obtain the inverse of A as given in Eq. (24.12a).

$$A^{-1} = \begin{bmatrix} \beta_0^{(1)} t_1 & \cdots & \beta_0^{(m)} t_m & \cdots & \beta_0^{(N)} t_N - \beta_0^{(1)} a_1 & \cdots & -\beta_0^{(m)} a_m & \cdots & -\beta_0^{(N)} a_N \\ \beta_k^{(1)} t_1 + 2l_1' \beta_{k-1}^{(1)} \cdots \beta_k^{(m)} t_m + 2l_m' \beta_{k-1}^{(m)} \cdots \beta_k^{(N)} t_N + 2l_N' \beta_{k-1}^{(N)} - (\beta_k^{(1)} a_1 + \beta_{k-1}^{(1)}) \cdots - (\beta_k^{(m)} a_m + \beta_{k-1}^{(m)}) \ldots - (\beta_k^{(N)} a_N' + \beta_{k-1}^{(N)}) \\ -2l_1' \beta_{2N-1}^{(1)} & \cdots & -2l_m' \beta_{2N-1}^{(m)} & \cdots & -2l_N' \beta_{2N-1}^{(N)} & \beta_{2N-1}^{(1)} & \cdots & \beta_{2N-1}^{(m)} & \cdots & \beta_{2N-1}^{(N)} \end{bmatrix} \tag{24.12a}$$

where

$$\beta_k^{(i)} = (-1)^{k+1} \sum_{j=1}^{N} \alpha_j^{(i)} \alpha_{2N-1\cdot k-j}^{(i)} \quad i = 1, 2, \ldots, N; k = 0, 1, \ldots, 2N-1 \tag{24.12b}$$

and

$$l_i' = l_i'(x)|_{x=a_i}; \quad t_i = 1 + 2a_i l_i' \quad i = 1, 2, \ldots, N \tag{24.12c}$$

References

1. N.K. Bose, *Digital Fillers: Theory and Application* (North Holland, New York, 1985)
2. A.B. Gershman, G.V. Serebryakov, J.F. Bohme, Constrained hung-turner adaptive beam-forming algorithm with additional robustness to wideband and moving jammers. IEEE Trans. Antennas Propag. **44**, 361–366 (1996)
3. V.E. Neagoe, Inversion of the Van der Monde matrix. IEEE Signal Process. Lett. **3**, 19–120 (1996)
4. G.H. Golumb, C.F. Vanloan, *Matrix Computations* (The Johns Hopkins University Press, Baltimore, Maryland, 1989)
5. R.A. Horn, C.R. Johnson, *Topics in Matrix Analysis* (Cambridge University Press, 1995)

Chapter 25
A New Canonic Lattice Realization of Arbitrary FIR Transfer Functions

Abstract A new lattice structure is presented for realizing a pair of arbitrary FIR transfer functions of the same order, which is canonic in multipliers as well as delays. By a simple modification of the procedure, one can also obtain a canonic realization of a single FIR transfer function. The problem cases have been discussed, along with their solutions, and the limitations of the method have been pointed out. Examples have been worked out to illustrate the theory.

Keywords FIR lattice · Canonic realization · Linear phase design · Digital filter

25.1 Introduction

The FIR lattice has received considerable attention in the digital signal processing literature because of modularity and robustness to coefficient quantization errors [1, 2]. The most popular form of the structure is shown in Fig. 25.1, with $k_a^{(i)} = k_b^{(i)} = k_i$ and $H_0(z) = G_0(z) = 1$. It realizes the transfer function

$$H_N(z) = X_N(z)/X_0(z) = 1 + a_1^{(N)} z^{-1} + a_2^{(N)} z^{-2} + \cdots + a_N^{(N)} z^{-N} \tag{25.1}$$

and its mirror image transfer function

$$G_N(z) = X'_N(z)/X_0(z) = z^{-N} H_N(z^{-1}) \tag{25.2}$$

Note that $H_N(z)$ and $G_N(z)$ have the same magnitude but equal and opposite phase on $|z| = 1$. Hence for designing a single digital filter with prescribed magnitude specifications, it is wasteful in multipliers, requiring double the number as compared to canonic realization. By a clever manipulation of the relevant equations, Makhoul [3] showed that each two multiplier section can be converted to a single multiplier one, but scaling is needed for the overall lattice—either at the input or at the output or at selected nodes, and thereby achieve near canonicity.

Source: S.C. Dutta Roy, 'A New Canonic Lattice Realization of Arbitrary FIR Transfer Functions', *IETE Journal of Research*, vol 53, pp 13–18, January–February 2007.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_25

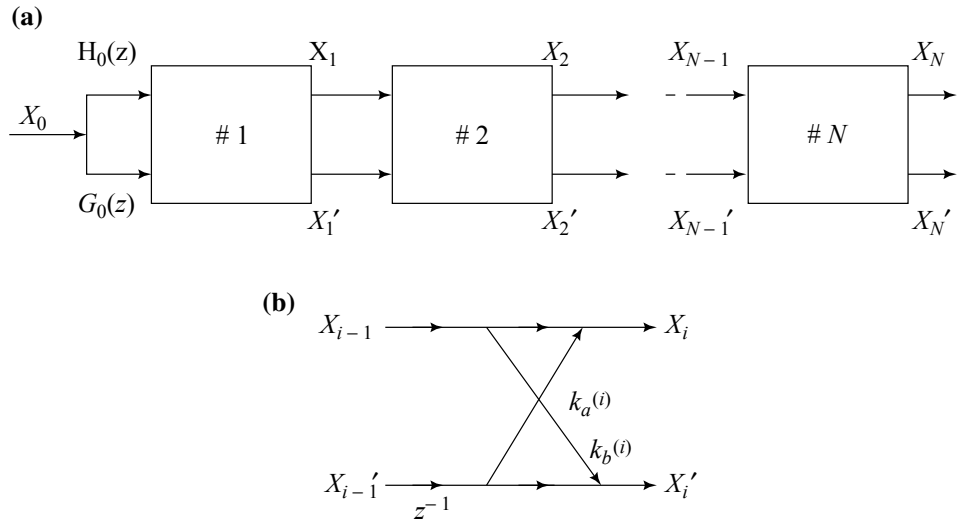


Fig. 25.1 **a** A generalized FIR lattice structure; **b** composition of the i-th section of (**a**). For the conventional structure, $H_0(z) = G_0(z) = 1$ and $k_a^{(i)} = k_b^{(i)} = k_i$

The usual treatment of the FIR lattice is to assume the structure of Fig. 25.1, with $k_a^{(i)} = k_b^{(i)} = k_i$ and $H_0(z) = G_0(z) = 1$, and then to analyse it to derive the following properties and design equations:

$$H_i(z) = X_i(z)/X_0(z) = 1 + a_1^{(i)} z^{-1} + a_2^{(i)} z^{-2} + \cdots + a_i^{(i)} z^{-i} \tag{25.3}$$

$$G_i(z) = X_i'(z)/X_0(z) = z^{-i} H_i(z^{-1}), \tag{25.4}$$

$$k_i = a_i^{(i)}, \tag{25.5}$$

$$a_n^{(i-1)} = [a_n^{(i)} - k_i a_{i-n}^{(i)}]/(1 - k_i^2), \quad 1 \le n \le (i-1), \tag{25.6}$$

$$i = N, N-1, \ldots, 1. \tag{25.7}$$

In [4], this lattice structure was derived from the digital two-pair approach of Mitra [1], where the question was also raised as to whether relationships other than that given by Eq. (25.4) are possible to be realized. In a subsequent paper [5], this question was answered in the affirmative by deriving a new lattice structure for satisfying the relationship

$$G_i(z) = z^{-i} H_i(-z^{-1}) \tag{25.8}$$

Equation (25.8) implies that $H_i(z)$ and $G_i(z)$ are complementary filters, i.e. if $H_i(z)$ is low-pass, then $G_i(z)$ is high-pass and vice versa. Such filters, as is well known, are required in cross-over networks for digital stereo and in multi-rate signal processing. This structure has cascade sections of the form of Fig. 25.1 with $H_0(z) = G_0(z) = 1$

and $k_a^{(i)} = -(1)^i k_b^{(i)}$; obviously, if i is even, then the two multipliers are equal, as is the case for Eq. (25.2), while if i is odd, then they are equal and opposite.

The present chapter is a generalization of [5], in which $H_N(z)$ and $G_N(z)$ are two arbitrary transfer functions, given by

$$H_N(z) = a_0^{(N)} + a_1^{(N)} z^{-1} + a_2^{(N)} z^{-2} + \cdots + a_N^{(N)} z^{-N}, \tag{25.9}$$

$$G_N(z) = b_o^{(N)} + b_1^{(N)} z^{-1} + b_2^{(N)} z^{-2} + \cdots + b_N^{(N)} z^{-N}. \tag{25.10}$$

This is achieved by using the structure of Fig. 25.1 as it is i.e. by using two different parameters in each lattice section and by keeping $H_0(z)$ and $G_0(z)$ unspecified. It is shown that the resulting structure is canonic in delays as well as multipliers. A simple modification of the procedure is shown to be applicable to a situation where a single transfer function is required to be realized canonically. Some examples are included to illustrate the application of the theory.

There exist some cases where the new procedure cannot be applied directly. It is shown that simple modifications obviate the problems without changing canonicity. There exist some other cases where the new procedure yields non-canonic solutions only; these have been pointed out.

25.2 Derivation of the New Structure

Consider the lattice section shown in Fig. 25.1b with $i = N$ and $H_N(z) = X_N(z)/X_0(z)$ and $G_N(z) = X'_N(z)/X_0(z)$ as given by Eqs. (25.9) and (25.10), respectively. It will now be explored whether by choosing $k_a^{(N)}$ and $k_b^{(N)}$ appropriately, one can make the transfer functions $H'(z) = X_{N-1}(z)/X_0(z)$ and $G'(z) = X'_{N-1}(z)/X_0(z)$ to be of order $N - 1$. It is easily shown that

$$H_N(z) = H'(z) + z^{-1} k_a^{(N)} G'(z) \tag{25.11}$$

$$G_N(z) = z^{-1} G'(z) + k_b^{(N)} H'(z) \tag{25.12}$$

Solving Eqs. (25.11) and (25.12), we get

$$H'(z) = K_N [H_N(z) - k_a^{(N)} G_N(z)] \tag{25.13}$$

$$G'(z) = K_N z [G_N(z) - k_b^{(N)} H_N(z)] \tag{25.14}$$

where

$$K_N = 1/[1 - k_a^{(N)} k_b^{(N)}] \tag{25.15}$$

Substituting for $H_N(z)$ and $G_N(z)$ from Eqs. (25.9) and (25.10), we get

$$H'(z) = K_N\{[a_0^{(N)} - k_a^{(N)}b_0^{(N)}] + [a_1^{(N)} - k_a^{(N)}b_1^{(N)}]z^{-1} + \cdots + [a_N^{(N)} - k_a^{(N)}b_N^{(N)}]z^{-N}\} \tag{25.16}$$

$$G'(z) = K_N z\{[b_0^{(N)} - k_b^{(N)}a_0^{(N)}] + [b_1^{(N)} - k_b^{(N)}a_1^{(N)}]z^{-1} + \cdots + [b_N^{(N)} - k_b^{(N)}a_N^{(N)}]z^{-N}\} \tag{25.17}$$

Note that if one chooses

$$k_a^{(N)} = a_N^{(N)}/b_N^{(N)} \text{ and } k_b^{(N)} = b_0^{(N)}/a_0^{(N)} \tag{25.18}$$

then both $H'(z)$ and $G'(z)$ will be of order N—l and can be re-designated as $H_{N-1}(z)$ and $G_{N-1}(z)$, respectively. Also note that the coefficients of the two transfer functions are given by

$$\begin{aligned} a_n^{(N-1)} &= K_N[a_n^{(N)} - k_a^{(N)}b_n^{(N)}], \\ b_n^{(N-1)} &= K_N[b_{n+1}^{(N)} - k_b^{(N)}a_{n+1}^{(N)}] \\ n &= 0, 1, 2, \ldots, N-1 \end{aligned} \tag{25.19}$$

We have thus achieved our objective and the procedure can now be repeated till we get

$$H_1(z) = a_0^{(1)} + a_1^{(1)}z^{-1}, G_1(z) = b_0^{(1)} + b_1^{(1)}z^{-1} \tag{25.20}$$

so that by applying Eqs. (25.18) and (25.19) with $N = 1$, we get

$$\begin{aligned} &k_a^{(1)} = a_1^{(1)}/b_1^{(1)}, k_b^{(1)} = b_0^{(1)}/a_0^{(1)}, \\ &H_0(z) = a_0^{(0)} = a_0^{(1)} \text{ and } G_0(z) = b_0^{(0)} = b_1^{(1)}. \end{aligned} \tag{25.21}$$

Summarizing, then, the design equations for the FIR lattice are as follows:

$$\begin{aligned} k_a^{(i)} &= a_i^{(i)}/b_i^{(i)} \text{ and } k_b^{(i)} = b_0^{(i)}/a_0^{(i)}, \\ i &= N, N-1, \ldots 1 \\ a_n^{(i-1)} &= [a_n^{(i)} - k_a^{(i)}b_n^{(i)}]/[1 - k_a^{(i)}k_b^{(i)}], \\ b_n^{(i-1)} &= b_{n+1}^{(i)} - k_b^{(i)}a_{n+1}^{(i)}]/[1 - k_a^{(i)}k_b^{(i)}], \\ n &= 0, 1, 2, \ldots i-1, \\ H_0(z) &= a_0^{(1)} \text{ and } G_0(z) = b_i^{(1)}. \end{aligned} \tag{25.22}$$

Before proceeding further, we consider an example.

Example 1 Let it be required to realize the transfer functions

$$H_3(z) = 1 + z^{-1} + z^{-2} + z^{-3}, \tag{25.23}$$

$$G_3(z) = 1 + 2z^{-1} + 3z^{-2} + 4z^{-3} \tag{25.24}$$

Using Eq. (25.22), we get

$$\begin{gathered} k_a^{(3)} = 1/4, k_b^{(3)} = 1, \\ a_0^{(2)} = 1, a_1^{(2)} = 2/3, a_2^{(2)} = 1/3, \\ b_0^{(2)} = 4/3, b_1^{(2)} = 8/3, b_2^{(2)} = 4, \\ k_a^{(2)} = 1/12, k_b^{(2)} = 4/3, \\ a_0^{(1)} = 1, a_1^{(1)} = 1/2, b_0^{(1)} = 2, b_1^{(1)} = 4, \\ k_a^{(1)} = 1/8, k_b^{(1)} = 2, H_0(z) = 1, G_0(z) = 4, \end{gathered} \tag{25.25}$$

The realization diagram can now be easily drawn, and is omitted for brevity.

Note, from this example, that $a_0^{(3)} = a_0^{(2)} = a_0^{(1)} = a_0^{(0)} = 1$. Is this in general true? The answer is in the affirmative because $a_0^{(i)} = 1$, (25.22) gives $a_0^{(i-1)} = 1$. This is a simplifying feature, and can be used in actual design. As can easily be verified, this simplification is not valid for the $b_0^{(i)}$ coefficients.

25.3 Cases Requiring Modification of the Procedure

We shall now discuss some cases where the algorithm given in Eq. (23.22) cannot be applied as it is, but some simple modifications overcome the problems. In all these cases, we illustrate the modifications by considering the first step of the procedure, but it should be obvious that such modifications will also work if any of these arise at any intermediate stage.

Case 1 If $b_N^{(N)} = 0$, then $k_a^{(N)} \to \infty$ and the synthesis procedure breaks down. Obviously, we have to synthesize the pair $H_N(z)$ and $G_{N-1}(z)$. The solution to the problem is to synthesize $H_N(z)$ and $G_N(z) = H_N(z) \pm G_{N-1}(z)$, both of which are now of degree N. Finally, one recovers $G_{N-1}(z)$ as $\mp [H_N(z) - G_N(z)]$.

Case 2 If $a_0^{(N)} = 0$, then $k_b^{(N)} \to \infty$. In this case, synthesize $H'_N(z) = H_N(z) \pm G_N(z)$ and $G_N(z)$, and recover $H_N(z)$ as $H'_N(z) \mp G_N(z)$.

Case 3 If $b_N^{(N)} = a_0^{(N)} = 0$, then realize $H'_N(z) = 1 + H_N(z)$ and $G_{N-1}(z)$, as in Case 1, and recover $H_N(z)$ as $H'_N(z) - 1$.

Case 4 If $k_a^{(N)} k_b^{(N)} = 1$, i.e. from (18), $a_N^{(N)} b_0^{(N)} = b_N^{(N)} a_0^{(N)}$, then one cannot proceed to the next stage. One remedy is to realize $H'_N(z) = 1 + H_N(z)$ and $G_N(z)$ as a pair and then recover $H_N(z)$ as $H'_N(z) - 1$.

25.4 Realization of a Single Transfer Function

When it is required to realize a single transfer function of the form Eq. (25.1) by a lattice, the existing procedure, as already mentioned, is wasteful in terms of multipliers. A canonical realization can be obtained by splitting the transfer function into two parts as follows. First, let N be odd; then we decompose $H_N(z)$ as

$$H_N(z) = H_{(N-1)/2}(z) + z^{-(N+1)/2} G_{(N-1)/2}(z) \tag{25.26}$$

where

$$H_{(N-1)/2}(z) = 1 + a_1^{(N)} z^{-1} + \cdots + a_{(N-1)/2}^{(N)} z^{-(N-1)/2}, \tag{25.27}$$

$$G_{(N-1)/2}(z) = a_{(N+1)/2}^{(N)} + a_{(N+3)/2}^{(N)} z^{-1} + \cdots + a_{(N)}^{(N)} z^{-(N-1)/2}. \tag{25.28}$$

Now realize $H_{(N-1)/2}(z)$ and $G_{(N-1)/2}(z)$ as a pair by the new procedure, and combine them as per Eq. (25.26). This realization will be canonic in delays as well as multipliers.

If N is even, then decompose $H_N(z)$ as

$$H_N(z) = H_{(N-2)/2}(z) + z^{-N/2} G_{N/2}(z) \tag{25.29}$$

where

$$H_{(N-2)/2}(z) = 1 + a_1^{(N)} z^{-1} + \cdots + a_{(N-2)/2}^{(N)} z^{-(N-2)/2}, \tag{25.30}$$

$$G_{N/2}(z) = a_{N/2}^{(N)} + a_{(N+2)/2}^{(N)} z^{-1} + \cdots + a_N^{(N)} z^{-N/2} \tag{25.31}$$

We then follow the same procedure as in the odd N case. Here we have exploited the fact that the last term in Eq. (25.9) can be zero but not that in Eq. (25.10). Alternatively, in this case, one could have written $H_N(z)$ as $H_{N/2}(z) + z^{-N/2} G_{N/2}(z)$ by assigning half of the term $a_{N/2}^{(N)} z^{-N/2}$ to each of $H_{N/2}(z)$ and $G_{N/2}(z)$. However, the decomposition given by Eqs. (25.29)–(25.31) is a better one because $k_a^{(N/2)} = 0$. We now illustrate the procedure by two examples.

Example 2 As an example of odd N, let it be required to realize

$$H_3(z) = 1 + z^{-1} + 2z^{-2} + 3z^{-3} \tag{25.32}$$

We first write

$$H_3(z) = (1 + z^{-1}) + z^{-2}(2 + 3z^{-1}) = H_1(z) + z^{-2} G_1(z) \tag{25.33}$$

and then use Eq. (25.22) on the pair $H_1(z)$ and $G_1(z)$; we get

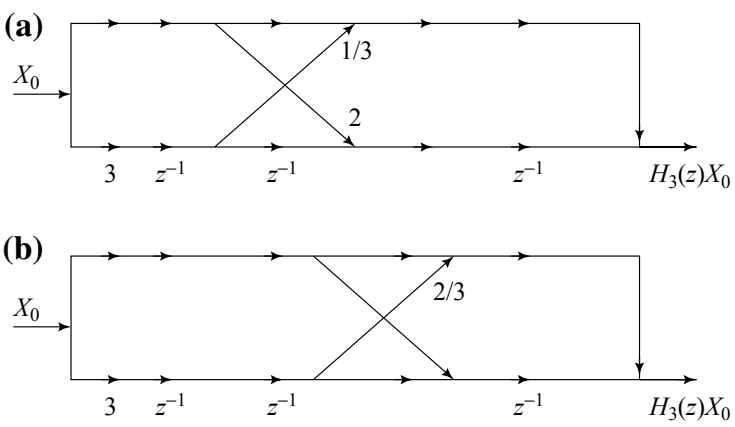


Fig. 25.2 **a** Realization of Eq. (25.32) as per the decomposition of Eq. (25.33) **b** alternative realization of Eq. (25.32) as per the decomposition of Eq. (25.35)

$$k_a^{(3)} = 1/3, k_b^{(3)} = 2, H_0(z) = 1, G_0(z) = 3 \tag{25.34}$$

The realization takes the form shown in Fig. 25.2a. It should be obvious that the decomposition is not unique. For example, we could also have taken

$$H_3(z) = (1 + 2z^{-2}) + z^{-1}(1 + 3z^{-2}) = H_2(z) + z^{-1}G_2(z) \tag{25.35}$$

Using Eq. (25.22) on the pair $H_2(z)$ and $G_2(z)$, we get

$$k_a^{(2)} = 2/3, k_b^{(2)} = 1, k_a^{(1)} = k_b^{(1)} = 0, H_0(z) = 1, G_0(z) = 3. \tag{25.36}$$

This realization is shown in Fig. 25.2b.

Example 3 We next consider the case of even N. Let the required transfer function be

$$H_4(z) = 1 + z^{-1} + 2z^{-2} + 3z^{-3} + 4z^{-4}. \tag{25.37}$$

One possible decomposition is

$$H_4(z) = (1 + z^{-1}) + z^{-2}(2 + 3z^{-1} + 4z^{-2}) = H_1(z) + z^{-2}G_2(z). \tag{25.38}$$

Applying Eq. (25.22) to the pair of transfer functions in Eq. (25.38), we get

$$k_a^{(2)} = 0, k_b^{(2)} = 2, k_a^{(1)} = 1/4, k_b^{(1)} = 1, \quad H_0(z) = 1, G_0(z) = 4. \tag{25.39}$$

The resulting realization is shown in Fig. 25.3a. Alternatively, we could have decomposed $H_4(z)$ as

$$H_4(z) = (1 + z^{-1} + z^{-2}) + z^{-2}(l + 3z^{-1} + 4z^{-2}) = H_2(z) + z^{-2}G_2(z) \tag{25.40}$$

Proceeding exactly as in the previous case, we get

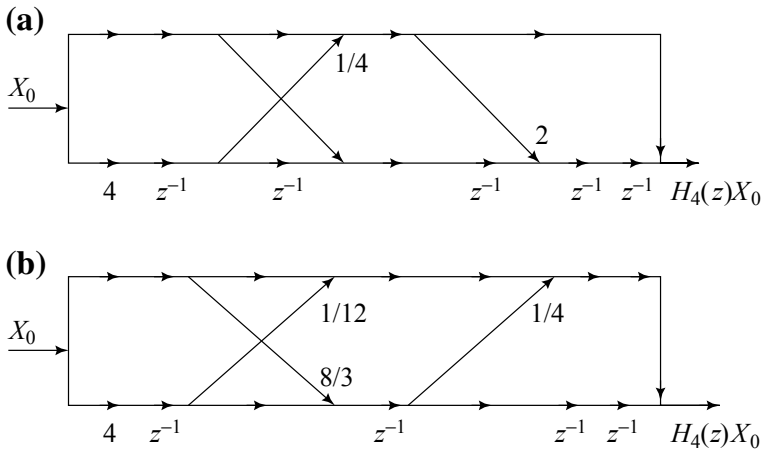


Fig. 25.3 **a** Realization of Eq. (25.37) as per the decomposition of Eq. (25.38) **b** alternative realization of Eq. (25.37) as per the decomposition of Eq. (25.40)

$$k_a^{(2)} = 1/4, k_b^{(2)} = 1, k_a^{(1)} = 1/12, k_b^{(1)} = 8/3, = H_0(z) = 1, G_0(z) = 4. \quad (25.41)$$

This realization is shown in Fig. 25.3b; compared to this, obviously, Fig. 25.3a realization is simpler.

25.5 Limitations of the New Procedures

Consider, first, the existing well-known case, in which $H_N(z)$ and $G_N(z)$ satisfy (2) so that $b_n^{(i)} = a_{i-n}^{(i)}$. Application of Eq. (25.22) then gives $k_a^{(i)} = k_b^{(i)} = a_i^{(i)}/a_0^{(i)} = k_i$. Next, consider the case in which $H_N(z)$ and $G_N(z)$ satisfy Eq. (25.8); this gives $b_n^{(i)} = (-1)^{i-n} a_{i-n}^{(i)}$ so that from Eq. (25.22), we get $k_a^{(i)} = (-1)^i k_b^{(i)}$. Thus the existing known cases follow as special cases of the new procedure. Another way of looking at these results is that in both of these cases, the new procedure fails to produce canonic realizations. This will be true for any pair of transfer functions, which are specially related to each other, so that the coefficients in one appear in the other also.

The same limitation arises in the case of a single linear phase transfer function. As is well known, the existing procedure, with two identical multipliers per lattice section, breaks down for a linear phase transfer function because $k_N = \pm 1$ and $a_N^{(N-1)}$ cannot be obtained from Eq. (25.6). The solution for this problem was given in [6] and a tutorial presentation of the solution was given in [7] through typical examples. However, that solution also utilized the conventional lattice structure. It is easily verified that the new procedure given in the previous section, when applied to a linear phase transfer function, gives the same realizations as in [6] and [7]. The underlying reason is that the two parts of the linear phase transfer function are specially related to each other. For this case, as well as for the case of Eq. (25.2),

the technique of [3] can be applied to obtain near canonic solutions. Whether such solutions can be obtained for other special relationships remains an open question.

25.6 Concluding Comments

A method has been presented for realizing two arbitrary FIR transfer functions of the same order by a canonic lattice structure. Problem cases have been considered, along with their solutions. A single transfer function is also shown to be realizable canonically by first splitting it into two parts, and then applying the new method. The limitations of the method have also been discussed.

A question that arises at this point is the following: Can an IIR transfer function be realized canonically by the same technique of using two non-identical lattice parameters in each section? A companion paper [8] answers this question in the affirmative.

References

1. S.K. Mitra, *Digital Signal Processing—A Computer Based Approach* (Tata McGraw Hill, New Delhi, 2001)
2. A.V. Oppenheim, R.W. Schafer, *Discrete Time Signal Processing* (Prentice Hall, Englewood Cliffs, 1989)
3. J. Makhoul, A class of all-zero lattice digital filters: properties and applications. IEEE Trans. Acoust. Speech Signal Process. ASSP **26**, 304–314 (1978)
4. S.C. Dutta Roy, R. Vishwanath, Derivation of the FIR lattice. IETE J. Educ. **45**, 211–212 (2004)
5. S.C. Dutta Roy, R. Vishwanath, Another FIR lattice structure. Int. J. Circ. Theor. Appl. **33**, 347–351 (2005)
6. S.C. Dutta Roy, Synthesis of FIR lattice structure. IEE Proc. Vis. Image Signal Process. **147**, 549–552 (2000)
7. S.C. Dutta Roy, Solution to a problem in FIR lattice synthesis. IETE J. Educ. **3**, 33–36 (January—March 2002) (Unfortunately, this paper contains a number of typographical and other mistakes. A corrected version is available from the author on request)
8. S.C. Dutta Roy, A New canonical realization of arbitrary IIR transfer functions. IETE J. Res. **53**(1), 19–23 (2007)

Chapter 26
A Note on Canonic Lattice Realization of Arbitrary FIR Transfer Functions

Abstract This chapter presents some simple alternative solutions to the cases requiring modification of the procedure given in (Dutta Roy in IETE J Res 53:13–18, 2007) [1] for canonic lattice realization of arbitrary transfer functions.

Keywords FIR lattice · Canonic realization · Digital filter

26.1 Introduction

In the paper [1], a canonical lattice realization of a pair of arbitrary FIR transfer functions was presented by this author. By a simple modification of the procedure, one could also obtain a canonical lattice realization of a single transfer function. The procedure was shown to require some modifications in a few cases. The purpose of this chapter is to present simple alternative solutions to these exceptional cases.

First, we give a summary of the new procedure as given in [1].

26.2 The New Procedure

The new procedure realizes the pair of transfer functions

$$H_N(z) = a_0^{(N)} + a_1^{(N)} z^{-1} + a_2^{(N)} z^{-2} + \cdots + a_N^{(N)} z^{-N} \tag{26.1}$$

and

$$G_N(z) = b_0^{(N)} + b_1^{(N)} z^{-1} + b_2^{(N)} z^{-2} + \cdots + b_N^{(N)} z^{-N} \tag{26.2}$$

in the form of a cascade of N sections as shown in Fig. 26.1a, where the i-th section, $i = N$ to 1, is of the form shown in Fig. 26.1b. The procedure is based on determining

Source: S.C. Dutta Roy, "A note on Canonic Lattice Realization of Arbitrary FIR Transfer Functions," *IETE Journal of Research*, vol 54, pp 71–72, January–February 2008.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_26

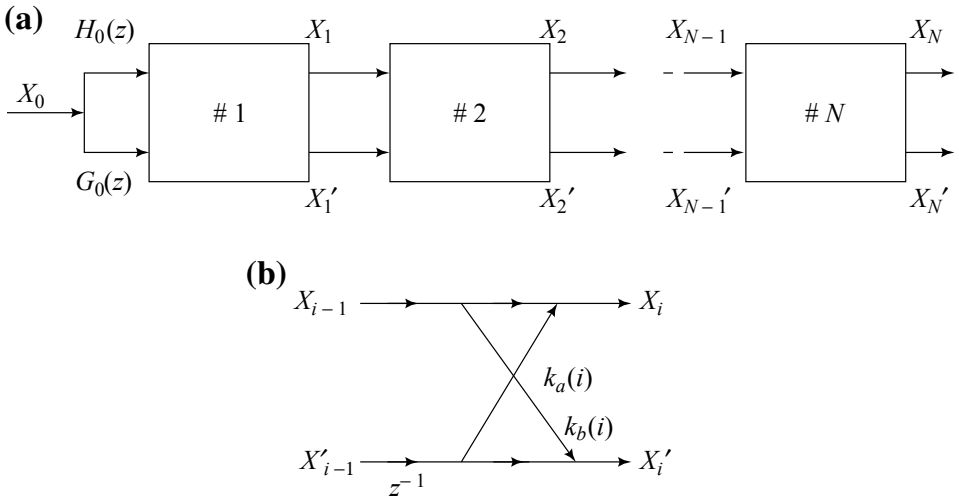


Fig. 26.1 **a** A generalized FIR lattice structure; **b** composition of the *i*th section of (**a**)

a series of N lower order transfer function pairs, of the same forms as Eqs. (26.1) and (26.2) with N replaced by i through the following algorithm:

$$k_a^{(i)} = a_i^{(i)}/b_i^{(i)}, k_b^{(i)} = b_0^{(i)}/a_0^{(i)}, i = N, N-1, \ldots 1;$$

$$\begin{aligned} a_n^{(i-1)} &= K_i[a_n^{(i)} - k_a^{(i)} b_n^{(i)}], b_n^{(i-1)} \\ &= K_i[b_{n+1}^{(i)} - k_b^{(i)} a_{n+1}^{(i)}], n = 0, 1, 2, \ldots i-1 \end{aligned} \tag{26.3}$$

$$K_i = 1/[1 - k_a^{(i)} k_b^{(i)}]; H_0(z) = a_0^{(1)} \text{and} G_0(z) = b_1^{(1)}$$

Obviously, if $b_i^{(i)} = 0$, then $k_a^{(i)} \to \infty, \infty$ and the procedure breaks down; we called this as Case 1 in [1]. Similarly, if $a_0^{(i)} = 0$, then $k_b^{(i)} \to \infty$ and one cannot proceed further; we designated this case as Case 2. If both $b_i^{(i)}$ and $a_0^{(i)}$ are zero, we get Case 3. Finally, if $k_a^{(i)} k_b^{(i)} = 1$, i.e. $a_i^{(i)} a_0^{(i)} = b_i^{(i)} b_0^{(i)}$, then one cannot proceed to the next stage; this was called Case 4. Modifications for each of these cases were presented in [1]. We present here simpler modifications for the first three cases.

Case 1 Let

$$H_i'(z) = G_i(z) = b_0^{(i)} + b_1^{(i)} z^{-1} + b_2^{(i)} z^{-2} + \cdots b_{i-1}^{(i)} z^{-(i-1)}, \tag{26.4}$$

and

$$G_i'(z) = H_i(z) = a_0^{(i)} + a_1^{(i)} z^{-1} + a_2^{(i)} z^{-2} + \cdots + a_i^{(i)} z^{-1}, \tag{26.5}$$

Realize the primed pair of transfer functions by the usual procedure. Here

$$k_a^{(i)'} = 0, \text{and} k_b^{(i)'} = a_0^{(i)}/b_0^{(i)} \tag{26.6}$$

and there is no breakdown.

Case 2 In this case, we let

$$H_i'(z) = G_i(z) = b_0^{(i)} + b_1^{(i)} z^{-1} + b_2^{(i)} z^{-2} + \cdots + b_i^{(i)} z^{-1} \tag{26.7}$$

and

$$G_i'(z) = H_i(z) = a_1^{(i)} z^{-1} + a_2^{(i)} z^{-2} + \cdots + a_i^{(i)} z^{-1} \tag{26.8}$$

and realize the primed pair of transfer functions by the usual procedure. Here

$$k_a^{(i)'} = b_i^{(i)} / a_i^{(i)} \text{ and } k_b^{(i)'} = 0 \tag{26.9}$$

and one can proceed further.

Case 3 In this case, we have to realize the transfer functions

$$H_i(z) = a_1^{(i)} z^{-1} + a_2^{(i)} z^{-2} + \cdots + a_i^{(i)} z^{-i} \tag{26.10}$$

and

$$G_i(z) = b_0^{(i)} + b_1^{(i)} z^{-1} + b_2^{(i)} z^{-2} + \cdots b_{i-1}^{(i)} z^{-(i-1)} \tag{26.11}$$

We can rewrite Eqs. (26.10) and (26.11) as

$$H_i(z) = z^{-1}(a_1^{(i)} + a_2^{(i)} z^{-1} + \cdots + a_i^{(i)} z^{-(i-1)} = z^{-1} H_{i-1}(z)) \tag{26.12}$$

and

$$G_i(z) = b_0^{(i)} + b_1^{(i)} z^{-1} + b_2^{(i)} z^{-2} + \cdots + b_{i-1}^{(i)} z^{-(i-1)} = G_{i-1}(z) \tag{26.13}$$

Now realize the pair $H_{i-1}(z)$ and $G_{i-1}(z)$ in the usual manner, and finally, add a delay at end of the upper line to get $H_i(z)$.

Case 4 Several alternative solutions can be worked out, but none of them appears to be simpler than that given in [1].

Reference

1. S.C. Dutta Roy, A new lattice realization of arbitrary FIR transfer functions. IETE J. Res. **53**, 13–18 (2007)

Chapter 27
A New Canonic Lattice Realization of Arbitrary IIR Transfer Functions

Abstract A new lattice realization is presented for any arbitrary stable IIR transfer function, which is always canonic in multipliers as well as delays, except for two degenerate cases, which require one (or more) additional delay(s), but no additional multipliers. In contrast to the existing lattice realization of such a transfer function, which uses an all-pass lattice combined with a ladder, the new structure is a truly lattice one. Also, for achieving multiplier canonicity in the conventional structure, one has to use special single multiplier lattice sections; in the new structure, canonicity evolves automatically.

Keywords Lattice structure · IIR filters · Canonic realization · Digital filters

27.1 Introduction

In realizing an arbitrary stable N-th order IIR transfer function

$$H_N(z) = \frac{A_N(z)}{B_N(z)} = \frac{a_0^{(N)} + a_1^{(N)} z^{-1} + \cdots a_N^{(N)} z^{-N}}{b_0^{(N)} + b_1^{(N)} z^{-1} + \cdots + b_N^{(N)} z^{-N}} \tag{27.1}$$

by the lattice structure, one first realizes an auxiliary all-pass lattice having the transfer function

$$H'_N(z) = \frac{z^{-N} B_N(z^{-1})}{B_N(z)} \tag{27.2}$$

In the next step, one combines a set of weighted output signals of this lattice to obtain $H_N(z)$, as illustrated in Fig. 27.1 [1, 2] for $N = 3$. This approach was originally proposed by Gray and Markel [3] and is the only one known so far for arbitrary IIR transfer functions. Obviously, the structure is delayed canonic but not so in terms of multipliers. However, it was also shown in [3] that by a simple manipulation, each

Source: S.C. Dutta Roy, "A New Canonic Lattice Realization of Arbitrary IIR Transfer Functions," *IETE Journal of Research*, vol 53, pp 19–23, January–February 2007.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_27

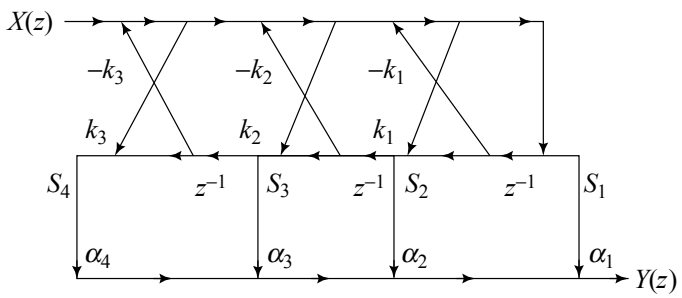


Fig. 27.1 The Gray-Markel structure for realizing a third order IIR transfer function

two multiplier lattice section can be converted to a single multiplier one, to achieve multiplier canonicity.

In this chapter, we present a new lattice structure for an arbitrary stable IIR transfer function, which does not require the additional ladder of Gray and Markel. In that sense, it is truly a modular lattice structure. It is always canonic in delays, as well as in multipliers, and the latter feature evolves automatically. In two degenerate cases, however, one (or more) extra delay(s) is (are) needed, but no extra multipliers.

The key to the new structure is very simple indeed. Instead of using two equal and opposite lattice parameters in each section ($\pm k_i$), as in Fig. 27.1, we use, in the new structure, two different lattice parameters to reduce the order of the transfer function by one.

27.2 Derivation of the New Structure

Since our aim is to realize Eq. (27.1) by a cascaded lattice structure, we explore the possibility of using the lattice section shown in Fig. 27.2, where $Y(z)/X(z)$ is required to be $H_N(z)$ and the terminating transfer function $G(z)$ is to be determined. If, by appropriate choice of $k_a^{(N)}$ and $k_b^{(N)}$, we can find a $G(z)$ of order $N-1$, then our purpose will be served and we can repeat the procedure till we get a terminating transfer function of order zero, which is, of course, a constant. By analysing Fig. 27.2, we get

$$\frac{Y(z)}{X(z)} = \frac{k_a^{(N)} + z^{-1}G(z)}{1 - k_b^{(N)} z^{-1} G(z)} \tag{27.3}$$

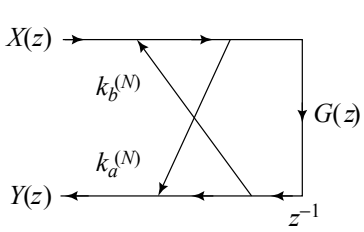


Fig. 27.2 New lattice section with unequal parameters

It can be verified that if $k_a^{(N)} = -k_b^{(N)}$ and $G(z) = 1$, then we get a first order all-pass function, as expected. Equating (27.3) to $A_N(z)/B_N(z)$, and simplifying, we get

$$G(z) = \frac{A_N(z) - k_a^{(N)} B_N(z)}{z^{-1}[B_N(z) + k_b^{(N)} A_N(z)]} = \frac{C(z)}{D(z)} \tag{27.4}$$

Substituting for $A_N(z)$ and $B_N(z)$ from Eq. (27.1), we get

$$\begin{aligned} C(z) = {} & (a_0^{(N)} - k_a^{(N)} b_0^{(N)}) + (a_1^{(N)} - k_a^{(N)} b_1^{(N)}) z^{-1} \\ & + (a_2^{(N)} - k_a^{(N)} b_2^{(N)}) z^{-2} + \cdots + (a_N^{(N)} - k_a^{(N)} b_N^{(N)}) z^{-N} \end{aligned} \tag{27.5}$$

and

$$\begin{aligned} D(z) = {} & z^{-1}[(b_0^{(N)}\theta + k_b^{(N)} a_0^{(N)}) + (b_1^{(N)} + k_b^{(N)} a_1^{(N)}) z^{-1} \\ & + (b_2^{(N)} + k_b^{(N)} a_2^{(N)}) z^{-2} + \cdots + (b_N^{(N)} + k_b^{(N)} a_N^{(N)}) z^{-N}] \end{aligned} \tag{27.6}$$

Now let

$$k_a^{(N)} = \frac{a_0^{(N)}}{b_0^{(N)}} \text{ and } k_b^{(N)} = -\frac{b_N^{(N)}}{a_N^{(N)}} \tag{27.7}$$

Then the constant term in $C(z)$ disappears, and so does the highest power term in $D(z)$. Thus $G(z)$ becomes a transfer function of order N—1, and can be re-designated as $H_{N-1}(z)$. It has the same form as Eq. (27.1) with N replaced by N—1, and its coefficients are related to those of $H_N(z)$ as follows:

$$a_i^{(N-1)} = a_{i+1}^{(N)} - k_a^{(N)} \theta b_{i+1}^{(N)},\ b_i^{(N-1)} = b_i^N + k_b^N a_i^N \qquad i = 0, 1, 2, \ldots N-1 \tag{27.8}$$

As mentioned earlier, this step can be repeated another N—1 times, to end up in a terminating transfer function, which will be a constant. The necessary design equations and the recurrence relations are obtained by replacing N in Eqs. (27.7) and (27.8) by a general index m, where $m = N - 1, N - 2, \ldots 1$. These will not be repeated here, for brevity. Using Eq. (27.8), we get the final terminating transfer function, $H_0(z)$ as

$$H_0(z) = \frac{a_0^{(0)}}{b_0^{(0)}} = \frac{a_1^{(1)}}{b_0^{(1)}} \tag{27.9}$$

It may be argued that we could have taken $b_0^{(N)} = 1$ in (27.1), without any loss of generality. However, note from Eq. (27.7), and the second recurrence relation in Eq. (27.8), that

$$b_0^{(N-1)} = b_0^{(N)} - b_N^{(N)} a_0^{(N)} / a_N^{(N)}, \tag{27.10}$$

which may not, in general, be unity, and we have to scale all the numerator and denominator coefficients of $H_{N-1}(z)$ by $b_0^{(N-1)}$ in order to be able to continue. This introduces another step of computation, and the recurrence relations will also get modified.

27.3 Examples

We shall now work out a number of examples, including some, in which the procedure, as presented, is not applicable as such, and some modifications are needed. As will be shown, these modifications are simple, and do not pose much of a problem. For simplicity, we shall confine ourselves only to the third order transfer function

$$H_3(z) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3}}{1 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}} \tag{27.11}$$

whose realization by the new procedure will, in general, be of the form shown in Fig. 27.3.

Example 1 Let $a_0 = 1$ and $a_1 = a_2 = a_3 = 0$; then from Eq. (27.7), we get $k_a^{(3)} = 1$ and $k_b^{(3)} = -\infty$. We cannot therefore proceed further. However, this difficulty can be removed by first realizing

$$H_3'(z) = 1 - H_3(z) = \frac{b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}}{1 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}} \tag{27.12}$$

By applying Eqs. (27.7) and (27.9) to Eq. (27.12), we get for $H_3'(z)$,

$$\begin{aligned} k_a^{(3)} &= 0 \text{ and } k_b^{(3)} = -1; \\ k_a^{(2)} &= b_1 \text{ and } k_b^{(2)} = 0; \\ k_a^{(1)} &= b_2 \text{ and } k_b^{(1)} = 0; \text{ and} \\ H_0(z) &= b_3 \end{aligned} \tag{27.13}$$

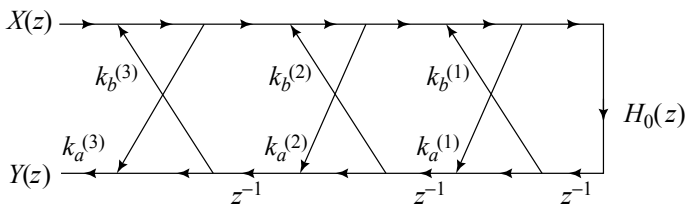


Fig. 27.3 General canonic lattice structure for a third order IIR transfer function

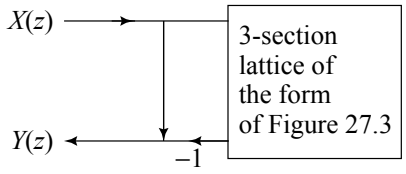


Fig. 27.4 Lattice structure for Examples 1–3. Parameter values are given in Eq. (27.13) for Example 1, Eq. (27.14) for Example 2 and Eq. (27.15) for Example 3

These results lead to the final realization shown in Fig. 27.4. Note that we require only three delays and three multipliers, viz. $k_a^{(2)}, k_a^{(1)}$ and $H_0(z)$, thus confirming canonicity.

Instead of Eq. (27.12), we could have also taken $H_3'(z) = 1 + H_3(z)$. The resulting structure is also canonic but, as may be verified, the structure obtained from Eq. (27.12) is somewhat simpler.

Example 2 Let $a_0 = 1$ and $a_2 = a_3 = 0$. Here also we face the same problem as in Example 1, and by proceeding in an exactly similar manner, we get the same realization as in Fig. 27.4, with the following changed parameters:

$$\begin{aligned} k_a^{(3)} &= 0 \text{ and } k_b^{(3)} = -1; \\ k_a^{(2)} &= b_1 - a_1 \text{ and } k_b^{(2)} = 0; \end{aligned} \tag{27.14}$$

$$\begin{aligned} k_a^{(1)} &= b_2 - a_1(b_1 - a_1) \text{ and } k_b^{(1)} = -a_1/b_3; \text{ and} \\ H_0(z) &= b_3. \end{aligned} \tag{27.15}$$

Here we require four multipliers and three delays, thus making the structure canonic.

Example 3 Let $a_0 = 1$ and $a_3 = 0$. Proceeding exactly as in the previous example, we get the realization of Fig. 27.4 with the following parameters:

$$\begin{aligned} k_a^{(3)} &= 0 \text{ and } k_b^{(3)} = -1; \\ k_a^{(2)} &= b_1 - a_1 \text{ and } k_b^{(2)} = -a_2/b_3; \\ k_a^{(1)} &= \frac{b_3[b_2 - a_2 - a_1(b_1 - a_1)]}{b_3 - a_2(b_1 - a_1)} \text{ and} \\ k_b^{(1)} &= \frac{a_1[b_3 - a_2(b_2 - a_2)]}{b_3[b_3 - a_2(b_1 - a_1)]}; \text{ and} \\ H_0(z) &= b_3. \end{aligned}$$

The requirement of five multipliers and three delays confirm canonicity,

Example 4 Let $a_0 = a_1 = a_2 = 0$. No modification of the transfer function is necessary in this case and application of Eqs. (27.7)–(27.9) leads to a realization with the following parameters:

$$
\begin{aligned}
k_a^{(3)} = 0 \text{ and } k_b^{(3)} &= -b_3/a_3; \\
k_a^{(2)} = 0 \text{ and } k_b^{(2)} &= -b_2/a_3; \\
k_a^{(1)} = 0 \text{ and } k_b^{(1)} &= -b_1/a_3; \text{ and} \\
H_0(z) &= a_3. \qquad (27.16)
\end{aligned}
$$

There are four multipliers and three delays, as needed for canonicity.

Our next example is a degenerate case, in which canonicity in multipliers can be assured but one (or more) extra delay(s) will be needed.

Example 5 Let $a_0 = a_3 = 0$. If we proceed as in Examples 1–3, we get, for $H_3'(z) = 1 - H_3(z)$, the following parameters:

$$k_a^{(3)} = 1 \text{ and } k_b^{(3)} = -1; \qquad (27.17)$$

and

$$k_a^{(2)} = a_0^{(2)}/b_0^{(2)} = -\infty;$$

We cannot therefore proceed further. In general, one can observe from Eq. (27.10) that this difficulty will occur in synthesizing Eq. (27.1) whenever

$$\frac{a_0^{(N)}}{b_0^{(N)}} = \frac{a_N^{(N)}}{b_N^{(N)}} \qquad (27.18)$$

We shall consider such a case in the next example.

For the present example, there are two alternatives. First, instead of $H_3(z)$, we realize $H_3'(z) = z^{-1}H_3(z)$, whose realization poses no problem. In most practical cases, this will be satisfactory because the magnitude function remains unaltered while the phase has an added linear term—ω. The resulting realization is multiplier canonic but requires an extra delay.

The second alternative is to rewrite the given transfer function as

$$
\begin{aligned}
H_3(z) &= \frac{(a_1 z^{-1} + a_2 z^{-2})(1 \pm 1/2z^{-1})}{1 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}} \cdot \frac{1}{(1 \pm 1/2z^{-1})} \\
&= H_3'(z) H_3''(z) \qquad (27.19)
\end{aligned}
$$

The realization of $H_3'(z)$ poses no problem; the resulting structure has to be cascaded to that of the first order transfer function $H_3''(z)$. Again, multiplier canonicity is maintained, but one extra delay is needed.

Example 6 Let $a_0 = a_2 = a_3 = 0$. We can follow the same procedure as in Example 5. In the first alternative, we realize $H_3'(z) = z^{-2}H_3(z)$, while in the second alternative, we multiply both numerator and denominator by $(1 \pm 1/2\ z^{-1})^2$. In either case, multiplier canonicity is achieved, but two extra delays are needed. Note that instead

of the constant 1/2 in the multiplying factor, in general, one can use 2^{-q} (where q is an integer greater than or equal to 1), which requires only shifting and no multiplication.

Example 7 As the last example, we consider a case which satisfies Eq. (27.18). Let, in Eq. (27.11), $a_3 = a_0 b_3$. In this case, we write

$$H_3'(z) = H_3(z) - a_0 \tag{27.20}$$

which gives the same transfer function as that of Example 5, and we apply either of the two techniques discussed there. The resulting structures are still multiplier canonic, but not so in delays.

27.4 Conclusions

In this chapter, we have presented a true lattice realization of an arbitrary N-th order IIR transfer function, which is canonic in multipliers, as well as in delays, except for two degenerate cases: one defined by Eq. (27.18) and the other by $a_0^{(N)} = a_N^{(N)} = 0$. Even in these cases, methods have been formulated for achieving multiplier canonicity.

While designing the Gray-Markel structure, the stability of the given transfer function is automatically tested for; this is not so in the procedure presented here. It is also not possible to incorporate such a test in the new procedure, because stability involves only the coefficients of $B_N(z)$, while our lattice parameters involve the coefficients of $A_N(z)$ as well. One, therefore, has to test for stability before proceeding with the design. Also, the Gray-Markel structure for a stable transfer function remains stable, even under finite word length implementation because the basic recursive unit is an all-pass filter, which remains all-pass even under coefficient quantization. This is not so in the new lattice structure.

A natural question that arises at this point is whether an arbitrary FIR transfer function can also be realized by a lattice, which is canonic in multipliers as well as delays. A companion paper [4] answers this question in the affirmative.

References

1. S.K. Mitra, *Digital Signal Processing—A Computer-based Approach* (Tata McGraw Hill, New Delhi, 2001)
2. A.V. Oppenheim, R.W. Schafer, *Discrete Time Signal Processing* (Prentice Hall, New Jersey, 1989)
3. A.H. Gray Jr., J.D. Markel, Digital lattice and ladder filter synthesis. IEEE Trans. Audio Electroacoust. AU-**21**, 491–500 (1973)
4. S.C. Dutta Roy, A New canonical realization of arbitrary FIR transfer functions. IETE J. Res. **53**, 13–18 (2007)

Chapter 28
A New Feedback Configuration for Canonical Realization of IIR Filters and Its Application to Lattice Realizations

Abstract A feedback configuration is presented for canonical realization of an arbitrary IIR transfer function $H(z) = P(z)/Q(z)$ with $M = P(z)^0 < N = Q(z)^0$, which does not appear to have been discussed in the literature. It consists of an IIR transfer function $F(z)$ of degree R. In the forward path, where $M \leq R < N$, and an FIR transfer function $G(z)$ of degree $N - M$, in which the first $R - M + 1$ terms are absent in the feedback path. $F(z)$ and $G(z)$ can be realized by any structure including lattice, if applicable. Application of this method is shown to solve the problem of canonical lattice realization of any IIR transfer function $H(z) = P(z)/Q(z)$ in which $M < N$ and the constant term in $P(z)$ is absent.

Keywords IIR filters · Feedback configuration · Canonical realization · Lattice structures

28.1 Introduction

Many different structures exist for canonical realization of an arbitrary IIR transfer function, examples being direct, cascade, parallel, ladder and lattice [1]. We present here another method, which does not appear to have been discussed in the literature. It is particularly suitable for transfer functions of the form $H(z) = P(z)/Q(z)$ with $M = P(z)^0 < N = Q(z)^0$, and consists of an IIR transfer function $F(z)$ of degree R in the forward path, where $M \leq R < N$, and an FIR transfer function $G(z)$ of degree $N - M$, in which the first $R - M + 1$ terms are absent, in the feedback path, $F(z)$ and $G(z)$ can be realized by any structure, including lattice, if applicable. Even in the case $M = N$, the numerator degree can be reduced by one by subtracting a constant term, and the new method can then be applied to the reduced transfer function.

In [2], a canonical lattice structure was presented for the arbitrary transfer function.

Source: S.C. Dutta Roy & Phani Bhushan Tholeti, "a New Feedback Configuration for Canonical Realization of IIR Filters and its Application to Lattice Realizations," *IETE Journal of Research*, vol 54, pp 45–50, January–February 2008.

© Springer Nature Singapore Pte Ltd. 2020

S. C. Dutta Roy, *Topics in Signal Processing*,
https://doi.org/10.1007/978-981-13-9532-1_28

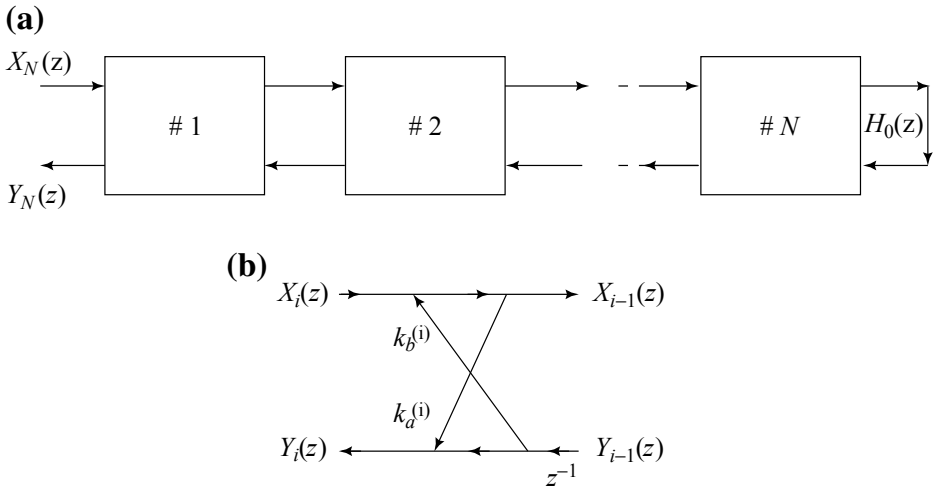


Fig. 28.1 **a** A general cascade of N sections **b** architecture of a typical section

$$H_N(z) = \frac{Y_N(z)}{X_N(z)} = \frac{a_0^{(N)} + a_1^{(N)} z^{-1} + \cdots + a_N^{(N)} z^{-N}}{1 + b_1^{(N)} z^{-1} + \cdots b_N^{(N)} z^{-N}} = \frac{P(z)}{Q(z)} \tag{28.1}$$

in the form shown in Fig. 28.1a, where the ith section, $i = N$ to 1, has the lattice architecture shown in Fig. 28.1b. The design equations are as follows:

$$k_a^{(i)} = \frac{a_0^{(i)}}{b_0^{(i)}} \text{ and } k_b^{(i)} = -\frac{b_i^{(i)}}{a_i^{(i)}}, \tag{28.2}$$

$$a_n^{(i-1)} = a_{n+1}^{(i)} - k_a^{(i)} b_{n+1}^{(i)}, \; b_n^{(i-1)} = b_n^{(i)} + k_b^{(i)} a_n^{(i)},$$
$$n = 0, 1, 2, \ldots i-1, \qquad i = N, (N-1), \ldots 1; \tag{28.3}$$

and

$$H_0(z) = \frac{a_1^{(1)}}{b_0^{(1)}} \tag{28.4}$$

Problem arises when one of the k parameters tends to infinity. Several illustrative cases and their solutions were presented in [2], except for those cases in which the first and the last terms in (28.1) are both absent, and a case in which

$$\frac{a_0^{(N)}}{b_0^{(N)}} = \frac{a_N^{(N)}}{b_N^{(N)}} \tag{28.5}$$

In each of these exceptional cases, one or more extra delays were needed for the solution. The feedback configuration presented here will be shown to solve the problem with $F(z)$ realized canonically by the method of [2], and $G(z)$ realized canonically by an FIR lattice by the method of [3].

Examples will be given to illustrate the applications of the method.

28.2 The Method

Consider the transfer function given by Eq. (28.1) with $a_i^{(N)} = 0$, $i = M+1, M+2, \ldots, N$, $M < N$, and the feedback configuration shown in Fig. 28.2 for which

$$H(z) = \frac{P(z)}{Q(z)} = \frac{F(z)}{1 + F(z)G(z)} \tag{28.6}$$

Let

$$F(z) = \frac{P(z)}{1 + c_1 z^{-1} + c_2 z^{-2} + \cdots c_R^{z^{-R}}} = \frac{P(z)}{D_R(z)} \tag{28.7}$$

where $M \leq R < N$. Then Eq. (28.6) becomes

$$H(z) = \frac{P(z)}{D_R(z) + P(z)G(z)} \tag{28.8}$$

Since $P(z)$ and $D_p(z)$ are both polynomials, of degrees M and R, respectively, it is clear that $G(z)$ is required to be a polynomial of degree $N - M$ i.e. $G(z)$ is an FIR transfer function. Further, since there are N known coefficients in $Q(z)$ and R unknown coefficients in $D_R(z)$, it is necessary to have only $N - R$ coefficients in $G(z)$, i.e. the first $R - M + 1$ coefficients will be missing in $G(z)$. Thus $G(z)$ is of the form

$$G(z) = g_{R-M+1} z^{-(R-M+1)} + g_{R-M+2} z^{-(R-M+2)} + \cdots + g_{N-M} z^{-(N-M)} \tag{28.9}$$

Thus from Eqs. (28.1), (28.8) and (28.9), we get

$$\begin{aligned} &1 + c_1 z^{-1} + c_2 z^{-2} + \cdots + c_R z^{-R} + [a_0^{(N)} + a_1^{(N)} z^{-1} + \cdots + a_M^{(N)} z^{-M}] \\ &\quad \times [g_{R-M+1} z^{-(R-M+1)} + g_{R-M+2} z^{-(R-M+2)} + \cdots + g_{N-M} z^{-(N-M)})] \\ &\quad = 1 + b_1^{(N)} z^{-1} + \cdots + b_N^{(N)} z^{-N} \end{aligned} \tag{28.10}$$

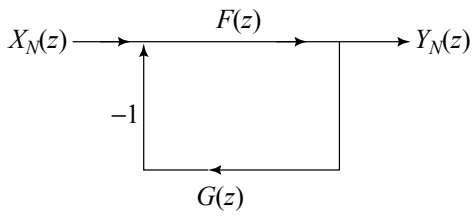


Fig. 28.2 The feedback configuration

Equating coefficients on both sides of (28.10) and solving the set of linear equations, we get the required values of c_i's and g_i's. Obviously, $F(z)$ and $G(z)$ are simpler transfer functions and can be realized by any existing method, including lattice, if applicable.

We now illustrate the method with an example.

Example 1 Let it be required to realize

$$H_3(z) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2}}{1 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}} \tag{28.11}$$

The only choice for $F(z)$ is

$$F(z) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2}}{1 + c_1 z^{-1} + c_2 z^{-2}} \tag{28.12}$$

and, correspondingly,

$$G(z) = g_1 z^{-1}. \tag{28.13}$$

Thus, from Eq. (28.10),

$$1 + c_1 z^{-1} + c_2 z^{-2} + (a_0 + a_1 z^{-1} + a_2 z^{-2}) g_1 z^{-1} = 1 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3} \tag{28.14}$$

Equating coefficients on both sides of Eq. (28.14), we get

$$c_1 + a_0 g_1 = b_1,\ c_2 + a_1 g_1 = b_2 \text{ and } a_2 g_1 = b_3 \tag{28.15}$$

Solution of Eq. (28.15) gives

$$g_1 = b_3/a_2,\ c_1 = b_1 - a_0 b_3/a_2 \text{ and } c_2 = b_2 - a_1 b_3/a_2 \tag{28.16}$$

The complete realization is shown in Fig. 28.3 with $F(z)$ realized in direct canonic form.

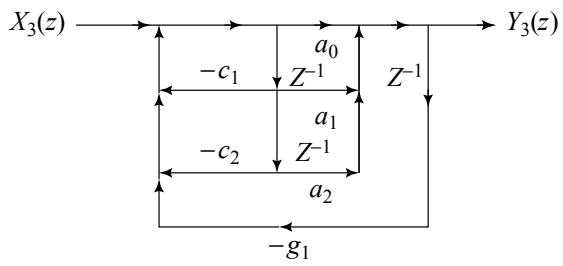


Fig. 28.3 Realization of Eq. (28.11) by the feedback configuration. Parameter values are given in Eq. (28.16)

Even if $M = N$, one can write the transfer function as the sum of a constant and another transfer function whose numerator is one degree less than the denominator. The next example illustrates this modification.

Example 2 Let

$$H_3(z) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2} + a_3 z^{-3}}{1 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}} \tag{28.17}$$

We can write

$$H_3(z) = \frac{a_3}{b_3} + \frac{a_0' + a_1' z^{-1} + a_2' z^{-2}}{1 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}} = \frac{a_3}{b_3} + H_3'(z) \tag{28.18}$$

where

$$\begin{aligned} &a_0' = a_0 - a_3/b_3,\ a_1' = a_1 - b_1 a_3/b_3, \\ &\text{and} \\ &a_2' = a_2 - b_2 a_3/b_3 \end{aligned} \tag{28.19}$$

$H_3'(z)$ can be realized as in Example 1, and the complete realization takes the form shown in Fig. 28.4, where c_1, c_2 and g_1 are given by Eq. (28.16) with a_i replaced by a_i', $i = 0, 1, 2$. The next example illustrates the multiplicity of choice possible for $D_R(z)$.

Example 3 Let

$$H_5(z) = \frac{a_0 + a_1 z^{-1}}{1 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3} + b_4 z^{-4} + b_3 z^{-5}} \tag{28.20}$$

The denominator polynomial of $F(z)$ can be any of the following:

$$D_1(z) = 1 + c_1 z^{-1}, \tag{28.21}$$

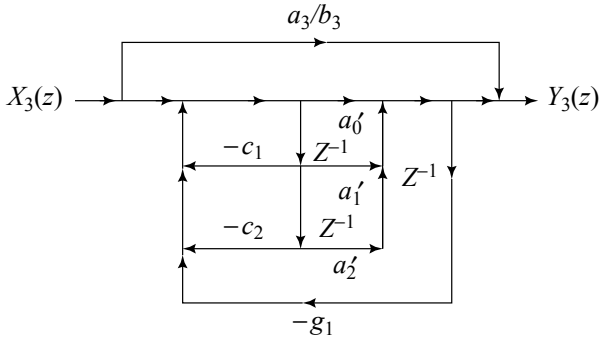


Fig. 28.4 Realization of Eq. (28.17) by the feedback configuration. Parameter values are given in Eq. (28.19)

$$D_2(z) = 1 + c_1 z^{-1} + c_2 z^{-2} \tag{28.22}$$

$$D_3(z) = 1 + c_1 z^{-1} + c_2 z^{-2} + c_3 z^{-3} \text{ and} \tag{28.23}$$

$$D_4(z) = 1 + c_1 z^{-1} + c_2 z^{-2} + c_3 z^{-3} + c_4 z^{-4}. \tag{28.24}$$

the corresponding $G(z)$ will be, respectively,

$$G_1(z) = g_1 z^{-1} + g_2 z^{-2} + g_3 z^{-3} + g_4 z^{-4}, \tag{28.25}$$

$$G_2(z) = g_2 z^{-2} + g_3 z^{-3} + g_4 z^{-4}, \tag{28.26}$$

$$G_3(z) = g_3 z^{-3} + g_4 z^{-4} \tag{28.27}$$

and

$$G_4(z) = g_4 z^{-4}. \tag{28.28}$$

In each of the last three variations, the feedback configuration can be applied to $F(z)$ as well.

28.3 Applications to Canonical IIR Lattice Realizations

In addition to $P(z)$ being of degree $M < N$, let, in Eq. (28.1), $a_0^{(N)} = 0$. As shown in [2], subtracting unity or any other constant from the transfer function creates the situation given by Eq. (28.5), and the procedure for lattice realization breaks down in the next step. We show here that the problem is solved by application of the new method presented here. We shall illustrate the solution with the same problems for which extra delays were required in [2].

Example 4 Let

$$H_3(z) = \frac{a_1 z^{-1} + a_2 z^{-2}}{1 + b_1 z^{-1} b_2 z^{-2} + b_3 z^{-3}} \tag{28.29}$$

Proceeding as in Example 1, we get

$$F(z) = \frac{a_1 z^{-1} + a_2 z^{-2}}{1 + b_1 z^{-1} + [b_2 - (a_1 b_3 / a_2) z^{-2}]} \tag{28.30}$$

and

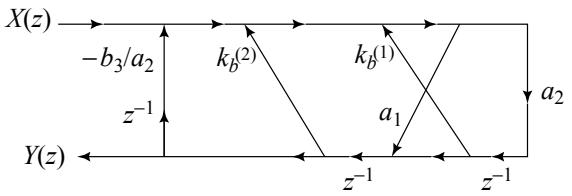


Fig. 28.5 Canonical lattice realization of Eq. (28.29) using the feedback configuration. Parameters are given in Eq. (28.31)

$$G(z) = (b_3/a_2)z^{-1}.$$

$F(z)$ can be realized by a lattice, following the procedure given in [2]. The required parameters are calculated from the design equations summarized in Eqs. (28.2)–(28.4) as:

$$\begin{aligned} k_a^{(2)} &= 0, k_b^{(2)} = (a_1b_3 - a_2b_2)/a_2^2, k_a^{(1)} = a_1 \\ k_b^{(1)} &= -\{(b_1/a_1) + [(a_1b_3 - a_2b_2)/a_2^2]\} \\ \text{and } H_0(z) &= a_2. \end{aligned} \tag{28.31}$$

The resulting complete realization is shown in Fig. 28.5.

Example 5 Let

$$H_3(z) = \frac{a_1z^{-1}}{1 + b_1z^{-1} + b_2z^{-2} + b_3z^{-3}} \tag{28.32}$$

Following the same procedure as in Example 1, we get

$$F(z) = \frac{a_1z^{-1}}{1 + b_1z^{-1}} \text{ and } G(z) = (b_2/a_1)z^{-1} + (b_3/a_1)z^{-2} \tag{28.33}$$

For the lattice realization of $F(z)$, the parameters are

$$k_a^{(1)} + 0,\ k_b^{(1)} = -(b_1/a_1) \text{ and } H_0(z) = a_1 \tag{28.34}$$

The complete realization is shown in Fig. 28.6.

Example 6 Let

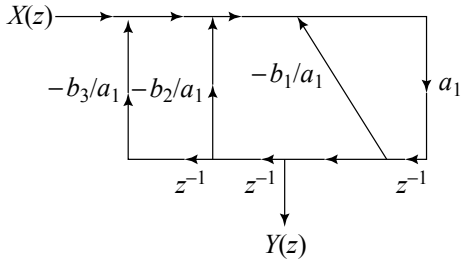


Fig. 28.6 Canonical lattice realization of Eq. (28.32) using the feedback configuration

$$H_3(z) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2} + a_0 b_3 z^{-3}}{1 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}} \tag{28.35}$$

This is an example of the case in which Eq. (28.5) is valid. Subtracting a_0 from $H_3(z)$, we get the new transfer function

$$H_3'(z) = \frac{(a_1 - a_0 b_1) z^{-1} + (a_2 - a_0 b_2) z^{-2}}{1 + b_1 z^{-1} + b_2 z^{-2} + b_3 z^{-3}} \tag{28.36}$$

This is of the same form as Eq. (28.29) and can be handled in the same manner as in Example 4. In the three examples considered in this section, the realization of $G(z)$ was simple. For a higher order $G(z)$ having many terms, one can use the canonical FIR lattice realization method given in [3], except for the case in which $G(z)$ is a linear phase polynomial. In the latter case, any other canonic FIR realization will have to be resorted to.

28.4 Conclusions

A new method has been presented for realizing an arbitrary IIR transfer function $H(z) = P(z)/Q(z)$ with $M = P(z)^0 < N = Q(z)^0$ in the form of two blocks, one of which is IIR and the other is FIR, in a feedback configuration. Both transfer functions are of lower order and can be realized by any known method, including lattice, if applicable. It is shown that the method is particularly useful in providing a canonical solution to the problem cases in IIR lattice synthesis, as reported in [2].

Acknowledgements After the publication of [2], it was pointed out by Professor S. K. Mitra of the University of California, Santa Barbara, that the idea of using two different parameters in a lattice section was exploited earlier in [4]. However, the approach, algorithms and the final realizations are, in general, different.

References

1. S.K. Mitra, *Digital Signal Processing. A Computer Based Approach* (Tata McGraw-HilI, New Delhi, 2001)
2. S.C. Dutta Roy, A new canonic lattice realization of arbitrary IIR transfer functions. IETE J. Res. **53**, 19–23 (2007)
3. S.C. Dutta Roy, A new canonic lattice realization of arbitrary FIR transfer functions. IETE J. Res. **53**, 13–18 (2007)
4. S.K. Mitra, P.S. Kamat, D.C. Huey, Cascaded lattice realization of digital filters. Int. J. Circuit Theory Appl. **5**, 3–11 (1977)